Information available at the Journal of Asia Business and Service Innovation (http://www.absi.kr/)

Asia Business and Service Innovation

ISSN: 3092-0558 (Print) | eISSN: 3092-1112 (Online)

A Study on Online Perception of Greenwashing Using Big Data Analysis

Arman Ahmed[1], Tae-Kyung Lee[2*], Sae-Beul Lee[3]

[1] *Department of Global Business, Kyungsung University, Busan, Republic of Korea*

[2] *Department of Business, Pukyong National University, Republic of Korea*

[3] *School of Global Business, Kyungsung University, Busan, Republic of Korea*

[*] Corresponding author: leetae3575@hanmail.net

ARTICLE INFO

Keywords:
greenwashing;
ESG management;
sustainability;
CSR;
consumer perception;
text mining,
big data analysis

ABSTRACT

This study aims to analyze public perceptions of greenwashing and to provide managerial implications from a business perspective. To achieve this objective, online text data related to greenwashing were collected from the period following the end of COVID-19 social distancing measures to the present. After data collection and preprocessing, frequency analysis, network analysis, centrality analysis, and CONCOR analysis were conducted to identify the characteristics of public discourse on greenwashing.

The results indicate that ESG-related keywords, including eco-friendliness, ESG management, sustainability, and corporate social responsibility (CSR), appeared most frequently. These highly frequent keywords also occupied central positions within the semantic network, indicating their importance in public discourse. In the centrality analysis, however, some frequently mentioned keywords, such as uncertainty and trust, exhibited relatively low centrality values, suggesting that their structural roles within the overall network were limited. Furthermore, the CONCOR analysis identified four thematic clusters: environmental policy and regulation, sustainability practices, greenwashing risks and controversies, and consumer and market responses.

Overall, the findings suggest that public discourse on greenwashing is multidimensional and organized around several key thematic areas. This study contributes to a better understanding of public perceptions of greenwashing in the Korean context and provides managerial implications for organizations seeking to develop effective ESG management and sustainability strategies.

1. Introduction

Nearly 8 out of 10 consumers (78%) indicate that sustainability is important when selecting a brand, and 61% say they are willing to alter their habits to shop more sustainably (IBM Institute for Business Value, 2024). Yet this consumer demand has a dark counterpart: when sustainability claims prove deceptive, trust collapses. The European Commission's (2020) Union-wide screening of environmental claims found that 53.3% of examined claims provided vague, misleading, or unfounded information. These findings underscore that greenwashing is structurally embedded in contemporary service competition rather than geographically confined. The erosion of trust is mirrored in the executive suite, where 58% of business leaders admit their own organizations have overstated their sustainability efforts and engaged in greenwashing (The Harris Poll, 2022). Together, these statistics reveal a widening chasm between symbolic environmental commitment and verifiable corporate action—a divergence defined as an environmental claim made by a private sector organization that cannot be substantiated, is made with deceptive intent, and aims to establish a competitive advantage (Spaniol et al., 2024).

In the era of global marketing, sustainability has become an increasingly important criterion in communication strategies. Consumers evaluate corporate and brand images with sustainability in mind, and their behavioral decisions are often shaped accordingly (White, Habib, & Hardisty, 2019). As public interest in sustainability grows, concerns about the credibility of corporate ESG management have also emerged as a critical issue. When ESG-related activi-

ties are perceived as misleading or exaggerated, they may foster negative attitudes toward companies and brands among the public, potentially leading to adverse behavioral responses (Nyilasy et al., 2014). In this context, greenwashing has come to be recognized as a key factor in global marketing, drawing attention from both academia and practitioners in the era of global business (Su, Zhu, & Albitar, 2026).

Greenwashing refers to the practice whereby companies construct an outwardly eco-friendly image through exaggerated environmental claims, despite the fact that their actual environmental performance yields little benefit or even causes harm to the environment. More specifically, it describes the act of selectively promoting positive environmental claims while concealing negative environmental information, thereby packaging poor performance as an overall achievement and creating a misleading impression (Delmas & Burbano, 2011). As ESG communication has emerged as a central element in the marketing domain, many companies have increasingly emphasized sustainability in their strategies. However, as prior research has noted, such communication efforts rarely translate into actual environmental outcomes. In this regard, ESG decoupling is broadly defined as the misalignment between ESG-related marketing disclosures and actual performance (Zhan, Zhan, & Li, 2025).

As such, a gap may exist between what companies communicate externally and their actual performance, a phenomenon that can be explained through the concept of decoupling. This has contributed to the spread of public skepticism toward corporations, which operates not merely as an individual-level response, but as a standard by which the legitimacy and validity of corporate marketing activities are reassessed and evaluated (Bagatini, Perin, & Medeiros, 2026). Furthermore, this situation has been described as a "trust gap," in which heightened public interest in ESG coexists with low trust in corporate environmental claims (Nguyen Thu Hoai, 2026).

These challenges have become further complicated in the context of the COVID-19 pandemic. The unprecedented pandemic has reshaped the social, cultural, and economic landscape (Lee & Yhang, 2025), and has had wide-ranging effects on how the public perceives sustainability and assesses risk (Nguyen Thu Hoai, 2026). In the post-pandemic era, the public has shown a growing tendency to be more sensitive to ethical and environmental issues, including corporate behavior. This suggests that marketing efforts must become more strategic and specific, and that deriving implications for strategic marketing activities based on online public discourse is a consideration of increasing importance (Lee, 2023). The pandemic heightened risk awareness across social, cultural, and economic domains, and as sustainable investment products grew during this period, demands for transparency in ESG-related marketing activities have also increased (Adams & Abhayawansa, 2022). Moreover, the fact that perceptions of greenwashing and ESG can shift in response to the prevailing socio-temporal environment underscores the need to examine public perception through a data-driven lens.

Nevertheless, prior research on greenwashing has largely relied on survey-based or experimental methods (Schmuck et al., 2018). While such approaches can yield valuable insights through rigorous scientific design, they are limited in their ability to capture large-scale, real-time, and naturally occurring public perceptions, given that the research subjects tend to be predetermined. Although some recent studies have begun applying text mining and big data methodologies, understanding of how concepts related to public perception are connected and structured remains limited.

This limitation is particularly pronounced in the Korean context. In South Korea, online portals such as Naver and Daum play a central role in information consumption and public opinion formation (Lee & Yhang, 2025). Korean users actively exchange opinions through a variety of community activities on portal-based platforms, including news comment sections and blogs (Lee, 2023). These platforms can thus serve as rich data sources that reflect public perception. However, research that draws on such data to analyze public perceptions of greenwashing in greater depth remains scarce.

To address this gap, rooted in both the social and cultural backdrop surrounding greenwashing and the limitations of prior research, this study aims to analyze public perceptions of greenwashing using large-scale textual data collected from major Korean portal sites. In order to develop more concrete, globally oriented business strategies, the study period is centered on the post-social distancing era in South Korea, reflecting the possibility that public perceptions of greenwashing may have shifted in the aftermath of the pandemic. Furthermore, to more closely examine the structural relationships among key terms derived from textual data, this study employs a text mining approach to analyze the frequency, network relationships, and structural patterns of core terms.

The study pursues three objectives. First, it identifies the key issues related to greenwashing in online textual data. Second, it analyzes the network structure and relative importance of core concepts associated with greenwashing-related terms. Third, it examines the cluster structure of core terms to uncover the structural dimensions of public perception regarding greenwashing.

This study contributes to the existing literature in several meaningful ways. First, by moving beyond the survey-based and experimental methods that have dominated prior greenwashing research, this study leverages large-scale, naturally occurring online text data to examine how public perceptions of greenwashing are structured within the Ko-

rean digital context. Second, the application of a multi-method analytical framework combining frequency analysis, network analysis, centrality analysis, and CONCOR analysis enable a more comprehensive understanding of the structural patterns underlying greenwashing discourse than single-method approaches have previously allowed. Together, these contributions provide both scholars and practitioners with a data-driven basis for understanding public perceptions of greenwashing and evaluating the effectiveness of corporate ESG communication in the Korean context.

2. Literature Review

2.1. Greenwashing in The Business Context

Greenwashing has become an increasingly important issue in contemporary business environments, particularly as firms place greater emphasis on sustainability communication. Earlier studies define greenwashing as the selective disclosure of positive environmental information while concealing negative aspects of corporate practices (Delmas and Burbano, 2011; Lyon and Maxwell, 2011). In this sense, greenwashing is not simply a mistake in communication, but rather a strategic practice embedded in corporate behavior.

From a theoretical perspective, this phenomenon can be explained through several frameworks. Signaling theory suggests that firms attempt to reduce information asymmetry by communicating sustainability-related information (Connelly et al., 2011). However, when such signals are not supported by actual performance, they may instead distort market understanding. Similarly, legitimacy theory explains that firms respond to social expectations by aligning their communication with dominant norms, even when their practices do not fully reflect those expectations (Suchman, 1995). Stakeholder theory also provides insight, as firms face increasing pressure from various stakeholders, including investors, regulators, and the broader public (Freeman, 1984). A recent concept analysis synthesizing 79 scholarly definitions establishes greenwashing as an unsubstantiated claim on environmental performance, made with deceptive intent, and aimed at securing a competitive advantage (Spaniol et al., 2024). This definitional consensus underscores that greenwashing is not merely an accidental misstatement, but a calculated corporate communication tactic. The analytical focus has also expanded beyond the focal firm to its institutional network; for instance, the concept of "vicarious greenwashing" describes how stakeholders attribute blame to lead firms for the deceptive environmental practices of supply-chain partners, thereby extending accountability from brand-level messaging to network-level governance (Pizzetti et al., 2021).

As sustainability becomes more central to corporate reputation and valuation, firms have stronger incentives to highlight environmental responsibility. At the same time, however, actual environmental performance does not always keep pace with communication efforts. This gap has contributed to growing skepticism and a decline in trust toward corporate environmental claims.

Previous research has shown that greenwashing leads to negative outcomes such as confusion, perceived risk, and reduced trust (Chen and Chang, 2013; Nyilasy et al., 2014). Furthermore, the damage of perceived greenwashing on customer satisfaction is heavily moderated by the firm's prior operational credibility; once skepticism is activated, even operationally competent firms suffer reputational decay, indicating that prior brand equity offers limited insulation against suspicion (Ioannou et al., 2023). A recent systematic bibliometric analysis of 419 peer-reviewed articles published between 2004 and 2024 confirms that consumer trust and corporate transparency have emerged as the most frequently investigated outcomes in the greenwashing literature, underscoring their centrality to business strategy research (Persakis et al., 2025). These effects are not limited to specific claims but often extend to overall perceptions of firms. In some cases, they may even weaken confidence in genuinely sustainable initiatives.

Despite these insights, most existing studies rely on survey or experimental methods. While useful, these approaches are limited in their ability to capture how perceptions form and evolve in real-world contexts. In particular, they do not fully explain how perceptions are structured or how different concepts are connected in public discourse.

This limitation suggests the need for approaches that can capture large-scale, naturally occurring data. In this regard, big data and text mining methods offer a useful alternative. By analyzing user-generated content, researchers can observe how perceptions emerge, interact, and change over time.

Building on this perspective, the present study examines public perceptions of greenwashing using large-scale textual data. By focusing on naturally generated online discourse, this study aims to provide a more realistic understanding of how greenwashing is perceived and structured in contemporary digital environments.

2.2. Greenwashing in The Business Context

Existing greenwashing studies have primarily relied on surveys, experiments, and structured interviews (Schmuck et al., 2018). While these approaches have contributed to understanding public perceptions, they are often limited by respondents' subjective interpretations, making it difficult to capture spontaneous public opinions as they naturally emerge.

In recent years, big data analytics has emerged as an alternative approach for examining public perceptions in real-world settings (Gandomi & Haider, 2015). Digital platforms such as online news media, blogs, and online communities provide large amounts of user-generated content (UGC) that

reflect naturally occurring public discourse. In the context of greenwashing, such data are especially valuable. The public freely expresses opinions about corporate environmental claims on online platforms, and these online discussions reflect authentic public responses to corporate greenwashing activities, making it possible to access opinions that surveys typically cannot reach. In this regard, user-generated content functions as a reliable indicator for understanding public evaluations (Tirunillai & Tellis, 2012). Furthermore, text mining techniques enable the systematic extraction of structural patterns in public perceptions from large-scale unstructured text data, going beyond the limitations of traditional survey-based methods (Humphreys & Wang, 2018).

Among text mining techniques, network-based analysis has gained attention as an effective methodology for examining the structure of public perceptions. Semantic network analysis identifies key keywords from textual data and visualizes co-occurrence relationships among them as a network, revealing how the public understands a given topic. This methodology is particularly useful for public perception research in that it goes beyond simple keyword frequency to capture relationships among concepts and broader patterns in discourse (Lee & Yhang, 2025). Prior studies have used this approach to examine online public perceptions on various social issues and found it to be effective.

In the Korean context, major portal platforms such as Naver and Daum function as key channels for information sharing and public opinion formation (Lee & Yhang, 2025), and textual data collected from these platforms can serve as rich data sources reflecting the perceptions of the Korean public. Given that public perceptions of greenwashing shift continuously in response to regulatory, social, and environmental changes, a network-based text mining approach offers a practical way to track how greenwashing discourse develops over time.

Based on this perspective, the present study applies a network-based text mining approach to analyze public perceptions of greenwashing as reflected in online discourse. Specifically, by combining frequency analysis, centrality analysis, network analysis, and CONCOR analysis, the study examines the frequency and structural relationships among greenwashing-related keywords, to better understand how greenwashing is discussed in the Korean online environment.

3. Methodology

3.1. Research Design

This study adopts a text mining-based research design to analyze public perceptions of greenwashing. The primary objective of this study is to identify key concepts related to greenwashing and examine the relationships and major themes among these concepts based on large-scale online public discourse.

Research Question1: What are the key issues related to greenwashing as expressed in online public discourse?

Research Question2: What are the structural relationships among the key greenwashing-related keywords?

Research Question3: What thematic clusters can be identified from the key greenwashing-related keywords?

3.2. Analysis Subjects

This study aims to exploratively identify the overall public perception of greenwashing formed in the Korean online space, rather than focusing on a specific industry or business sector. Data were collected from major Korean portal platforms, specifically Naver and Daum, which serve as key channels for information sharing and public opinion formation in South Korea, making them well suited for the comprehensive capture of public perceptions within the Korean context.

The data collection period was set from April 18, 2022 to April 23, 2026, spanning approximately four years. This period was set based on the end of social distancing measures in South Korea following the COVID-19 pandemic, with the aim of examining how public discourse on greenwashing is structured within the specific social context of the post-pandemic era. The pandemic brought unprecedented changes across social and economic environments, and it was considered that these changes would have had a meaningful impact on public perceptions of ESG and greenwashing.

Textom was used as the data collection tool. Textom is a big data analytics platform that collects large-scale textual data from diverse channels including news articles, blog posts, and online communities across major portal platforms. Data were collected using keywords such as "greenwashing," "ESG," "sustainability," and "eco-friendly." This study is not a sample survey targeting a specific group, but an exploratory study that analyzes the full scope of naturally occurring online public discourse, with the aim of identifying structural patterns in discourse rather than achieving statistical representativeness.

The collected data underwent a preprocessing procedure to enhance data quality. As a result of preprocessing, a total of 5,977 unique terms were extracted, comprising 4,376 nouns, 62 adjectives, 323 verbs, and 1,216 foreign terms. It should be noted that this figure represents the number of unique terms extracted after preprocessing, not the number of original posts or documents collected. To ensure contextual accuracy, subsequent analysis was restricted to nouns.

For the final analysis, the top 50 most frequently occurring words were selected. Single-syllable words with limited semantic value were removed, and semantically similar words were consolidated. For example, carbon-related terms such as "greenhouse gas" and "carbon

emission" were merged into a single category, and "sustainable" was standardized as "sustainability."

4. Results

4.1. Frequency Analysis of Key Terms

Among the 4,376 nouns extracted from the dataset, frequency analysis was conducted to identify the most salient themes in public discourse. As shown in Table 1, the results indicate that eco-friendliness (2,846), ESG management (2,482), and sustainability (1,524) are the most frequently occurring terms, suggesting that discussions are primarily centered on environmental responsibility and ESG-related issues.

In addition, terms such as corporate social responsibility (CSR) (948) and carbon emissions (592) further highlight the strong association between greenwashing and corporate environmental practices. At the same time, the presence of terms such as deception (362) and uncertainty (249) reflects a notable level of skepticism in public discourse, indicating concerns about the credibility of corporate environmental claims.

Overall, the results suggest that public perceptions of greenwashing are shaped by a combination of sustainability-related expectations and critical evaluations of corporate behavior. The high frequency of ESG-related keywords, including ESG management and sustainability, suggests that the Korean online public perceives ESG not merely as a marketing concept but as a key criterion for evaluating corporate values and behavior. This is consistent with Adams & Abhayawansa (2022), who argue that ESG communication functions as a central mechanism in corporate management. At the same time, the high frequency of negative keywords such as deception, uncertainty, and confusion suggests that while public expectations regarding ESG have increased, trust in corporate environmental claims remains fragile. In other words, the coexistence of high ESG expectations and persistent skepticism indicates that the more fundamental challenge for firms lies not simply in communicating ESG, but in establishing the credibility of that communication.

Table 1. Frequency analysis result

No.	Keyword	Frequency
1	Eco-friendliness	2846
2	ESG management	2482
3	Sustainability	1524
4	Corporate social responsibility	948
5	Carbon emissions	592
6	Environmental pollution	405
7	Consumer	389
8	Deception	362
9	Plastic	277
10	Uncertainty	249
11	Brand image	237
12	Misperception	162
13	Strategy	158
14	Promotion	157
15	Climate	125
16	Waste	124
17	Negative impact	109
18	Regulation	106
19	Packaging advertising	106
20	Confusion	104
21	Energy	104
22	Investment	100
23	Government	99
24	Certification	91
25	Circulation	87
26	Policy	81
27	Environmental protection	80
28	Trust	58
29	Solution	57
30	Evaluation	56
31	Profitability	54
32	Reduction	53
33	Risk	51
34	Campaign	51

35	Fairness	45
36	Distortion	41
37	Claim	36
38	Controversy	33
39	Ethics	31
40	Purchase intention	31
41	Misuse	26
42	Compliance	25
43	Accountability	23
44	Exposure	23
45	Transparency	22
46	Exaggeration	22
47	Acceptance	21
48	Company	20
49	Eco-label	20
50	Deceptiveness	20

Furthermore, the network analysis of the collected keywords reveals that the overall network structure exhibits a highly interconnected pattern, in which most keywords are directly linked to one another. Rather than appearing as isolated groups, the keywords are closely interconnected around environmental themes.

As shown in Figure 1, the blue nodes represent the frequency of each keyword, with larger nodes indicating higher frequency. According to the figure, among the keywords with relatively high frequency, terms such as sustainability, ESG management, and corporate social responsibility (CSR) are positioned at or near the center of the network and maintain a comparatively large number of connections with other nodes. This suggests that these concepts are frequently mentioned and discussed alongside other keywords in online discourse, functioning as central reference points within the network. In addition, carbon emissions and environmental pollution are also connected to multiple keywords and appear near the central area of the network. This indicates that environmental issues are not discussed as independent concepts are not discussed as independent concepts but are consistently mentioned in conjunction with broader ESG-related themes.

Keywords such as transparency, accountability, ethics, and compliance are distributed across the network while remaining closely interconnected with one another. This finding suggests that governance-related discussions tend to co-occur in the context of greenwashing, reflecting their close associations within the network. Furthermore, consumer-related keywords are connected to a wide range of ESG and environmental terms. This reflects the fact that consumer-related discussions are not confined to a single topic but are linked to multiple dimensions of ESG. Taken together, the results suggest that greenwashing-related keywords form a closely interconnected semantic network encompassing ESG-related environmental issues, corporate social responsibility, and consumer-related perspectives. This interconnected network suggests that discussions of greenwashing are closely associated with multiple dimensions of ESG, including corporate social responsibility, consumer-related issues, and governance.

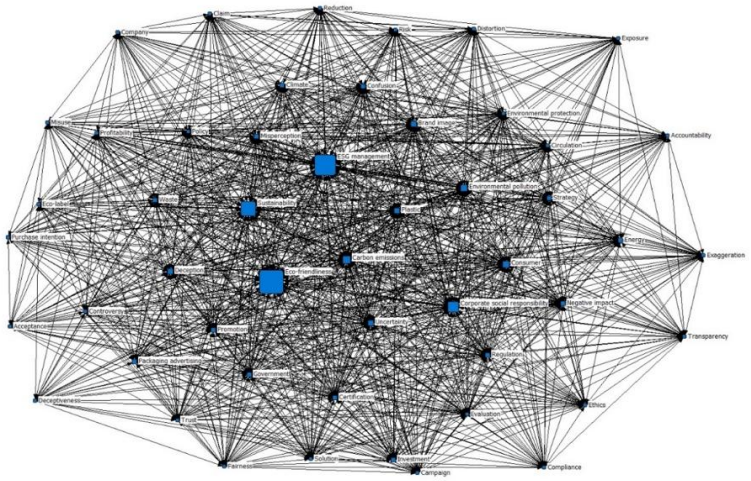

Figure 1. Network analysis result.

4.2. Centrality Analysis of Key Terms

This study examined the structural importance of keywords by applying Degree Centrality and Eigenvector Centrality. As shown in Table 2, the results indicate that eco-friendliness (0.157), ESG management (0.147), sustainability (0.119), and corporate social responsibility (CSR) (0.112) are positioned at the center of the network. These keywords are not only frequently mentioned but also connected to a wide range of other terms, suggesting that they play a central role in organizing the overall structure of the network.

Another notable finding is the large gap between the highest and lowest centrality values. For instance, the centrality value of eco-friendliness (0.157) is more than fifteen times higher than that of trust (0.010). This indicates that connections within the network are not evenly distributed but instead concentrated around a small number of key concepts. In other words, ESG-related discussions appear to revolve around a limited set of dominant themes.

On the other hand, keywords such as uncertainty, misperception, strategy, investment, certification, and trust show relatively lower centrality values despite appearing with a certain level of frequency. This suggests that these terms tend to appear in more specific contexts and are less involved in linking different parts of the network. As a result, they do not contribute as strongly to the overall structure.

The results from eigenvector centrality follow a similar pattern. Keywords that rank highly in degree centrality also tend to maintain their positions in eigenvector centrality, which implies that they are connected not only to many other terms, but also to other influential keywords. This reflects a structure in which key concepts are closely interconnected with one another.

Taken together, the findings suggest that ESG-related discourse is shaped not simply by how often keywords appear, but by how they are connected within the network. A small number of central keywords appear to play a dominant role in organizing the relationships among concepts.

Keywords with high centrality, such as eco-friendliness, ESG management, and sustainability, suggest that these concepts serve as central themes in Korean online discussions of greenwashing. This suggests that corporate ESG communication strategies that fail to address these central concepts in a substantive and credible manner are unlikely to resonate with the public.

Particularly noteworthy is the finding that trust, despite appearing with a certain level of frequency, exhibits markedly low centrality within the network. This indicates that while trust is referenced in greenwashing discourse, it has not yet established itself as a core concept in organizing public perceptions. This finding can be explained by the coexistence of high public interest in ESG and persistently low trust in corporate environmental claims, which Nguyen (2026) describes as a "trust gap." From a managerial perspective, this reflects the vulnerability of trust as a concept that remains structurally peripheral in greenwashing discourse. Paradoxically, however, this may represent a strategic opportunity for firms. Those that are able to substantively address trust-related concerns and demonstrate transparent ESG performance may be well positioned to secure a distinctive competitive advantage in a discursive environment where trust remains structurally marginalized.

Table 2. Centrality analysis result

Keyword	Frequency		Degree Centrality		Eigen Vector Centrality	
	Freq.	Rank	Coef.	Rank	Coef.	Rank
Eco-friendliness	2846	1	0.157	1	0.460	1
ESG management	2482	2	0.147	2	0.441	2
Sustainability	1524	3	0.119	3	0.372	3
Corporate social responsibility	948	4	0.112	4	0.359	4
Carbon emissions	592	5	0.072	5	0.237	5
Environmental pollution	405	6	0.064	6	0.210	6
Consumer	389	7	0.055	7	0.187	7
Deception	362	8	0.051	8	0.180	8
Plastic	277	9	0.045	10	0.151	10
Uncertainty	249	10	0.047	9	0.161	9
Brand image	237	11	0.037	11	0.127	11
Misperception	162	12	0.029	13	0.093	14
Strategy	158	13	0.029	14	0.099	13

Promotion	157	14	0.030	12	0.106	12
Climate	125	15	0.026	15	0.084	15
Waste	124	16	0.021	17	0.067	17
Negative impact	109	17	0.022	16	0.075	16
Regulation	106	18	0.018	22	0.057	23
Packaging advertising	106	19	0.019	19	0.066	18
Confusion	104	20	0.020	18	0.066	19
Energy	104	21	0.018	23	0.058	21
Investment	100	22	0.014	26	0.048	26
Government	99	23	0.019	20	0.063	20
Certification	91	24	0.013	27	0.045	27
Circulation	87	25	0.019	21	0.058	22
Policy	81	26	0.016	24	0.051	24
Environmental protection	80	27	0.015	25	0.050	25
Trust	58	28	0.010	32	0.036	32
Solution	57	29	0.010	33	0.034	33
Evaluation	56	30	0.012	28	0.039	28
Profitability	54	31	0.011	29	0.039	29
Reduction	53	32	0.011	30	0.039	30
Risk	51	33	0.011	31	0.037	31
Campaign	51	34	0.009	35	0.030	35
Fairness	45	35	0.010	34	0.033	34
Distortion	41	36	0.008	36	0.027	36
Claim	36	37	0.008	37	0.027	37
Controversy	33	38	0.007	38	0.025	38
Ethics	31	39	0.006	39	0.022	39
Purchase intention	31	40	0.006	40	0.019	40
Misuse	26	41	0.005	41	0.017	41
Compliance	25	42	0.004	42	0.014	42
Accountability	23	43	0.004	43	0.014	43
Exposure	23	44	0.002	50	0.007	50
Transparency	22	45	0.004	44	0.014	44
Exaggeration	22	46	0.004	45	0.014	45
Acceptance	21	47	0.004	46	0.013	47
Company	20	48	0.004	47	0.012	49
Eco-label	20	49	0.004	48	0.013	48
Deceptiveness	20	50	0.004	49	0.014	46

4.3. CONCOR (CONvergence of Iterated CORrelations) Analysis

This study conducted CONCOR analysis to examine the structure of the keyword network. As shown in Figure 2, the results indicate that greenwashing-related themes are divided into four distinct clusters. This suggests that greenwashing-related discourse is not organized around a single topic, but rather structured across four domains: environmental policy and regulation, sustainability practices, greenwashing risks and controversies, and consumer and market responses.

The first cluster, Environmental Policy and Regulation, includes keywords such as circulation, energy, waste, regulation, carbon emissions, and compliance. This cluster reflects discussions related to policy responses, regulatory frameworks, and resource management in addressing environmental issues in the context of greenwashing. This suggests that greenwashing discourse in the Korean online context is closely associated with the regulatory and policy environment.

The second cluster, Sustainability Practices, consists of keywords such as solution, eco-label, ESG management, eco-friendliness, environmental protection, corporate social responsibility (CSR), and certification. This cluster represents the practical dimensions of corporate sustainability, encompassing environmentally friendly activities and the institutional mechanisms used to signal such efforts. This suggests that discussions of corporate sustainability are centered not on general claims, but on specific and verifiable mechanisms such as eco-friendly certification and eco-labels. This further suggests that firms relying on vague or unsubstantiated environmental claims may be more vulnerable to greenwashing perceptions.

The third cluster, Greenwashing Risks and Controversies, includes keywords such as environmental pollution, uncertainty, plastic, exposure, accountability, policy, government, deception, exaggeration, confusion, risk, and negative impact. This cluster captures negative perceptions, risks, and controversies associated with greenwashing, particularly issues related to misinformation and exaggeration. The coexistence of keywords such as deception, exaggeration, and confusion within a single cluster suggests that greenwashing discourse encompasses multiple forms of risk, including not only deliberate misinformation but also unintentional confusion and accountability-related concerns.

The fourth cluster, Consumer and Market Responses, includes keywords such as strategy, purchase intention, consumer, profitability, sustainability, company, acceptance, promotion, trust, campaign, and brand image. This cluster reflects how ESG activities are linked to consumer perceptions, market responses, and corporate performance. The clustering of keywords such as purchase intention, brand image, trust, and profitability suggests that greenwashing discourse is closely linked to consumer behavior and corporate performance. This is consistent with prior research showing that greenwashing leads to negative consumer outcomes, including reduced trust and diminished purchase intention (Nyilasy et al., 2014; Chen & Chang, 2013).

Taken together, the four-cluster structure suggests that greenwashing discourse in the Korean online context constitutes a multidimensional evaluative framework that simultaneously encompasses regulatory, practical, risk-related, and market dimensions.

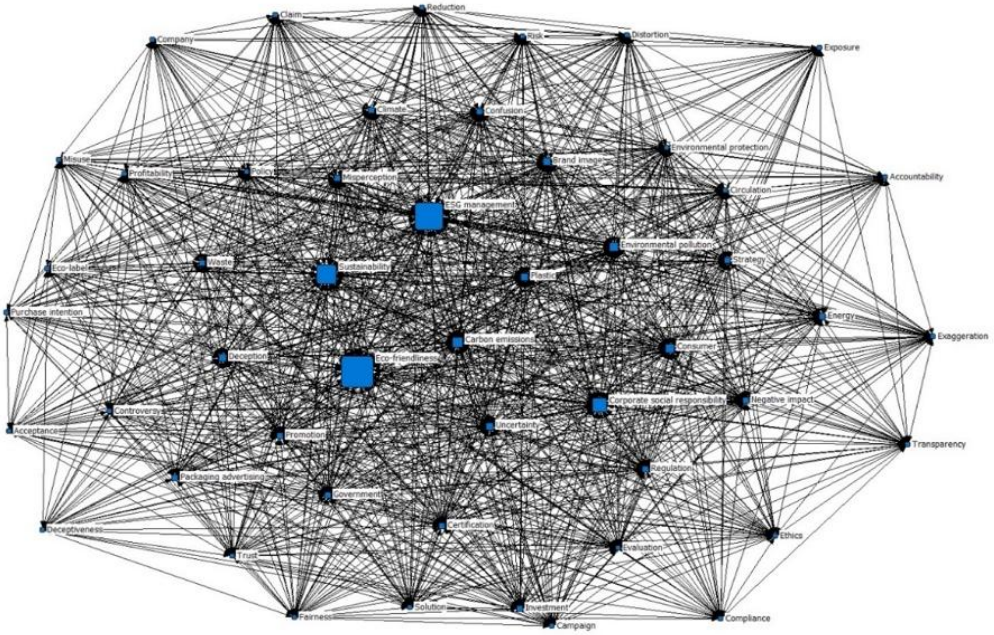

Figure 2. CONCOR analysis results.

5. Conclusion

5.1. Summary

This study aimed to analyze the perceptions of the Korean online public regarding greenwashing using online text data. To achieve this, prior research on greenwashing was reviewed and the limitations of existing studies were identified, after which public perceptions within online discourse were analyzed through big data and text mining approaches. Given the possibility that public perceptions of greenwashing-related risks and environmental awareness may have

shifted in the aftermath of the COVID-19 pandemic, data were collected from the period following the end of social distancing measures in South Korea. Data were collected using Textom and refined through text mining techniques. Frequency analysis, centrality analysis, network analysis, and CONCOR analysis were then conducted sequentially to examine the data from multiple perspectives.

The results of the frequency analysis indicate that ESG-related keywords such as eco-friendliness, ESG management, and sustainability appeared most frequently, reflecting the centrality of environmental responsibility in public discourse. At the same time, the presence of terms such as deception and uncertainty confirms that a degree of public skepticism exists toward corporate environmental claims.

The network and centrality analyses reveal that eco-friendliness, ESG management, sustainability, and CSR are positioned at the center of the network. A substantial gap in centrality values was observed, suggesting that network connections are concentrated around a limited number of key concepts. In contrast, keywords such as uncertainty and trust, despite appearing with a certain level of frequency, exhibit relatively low centrality, indicating that they function as peripheral concepts with weaker structural connectivity within the network. The eigenvector centrality results follow a similar pattern, confirming strong interconnections among the central concepts.

The CONCOR analysis divides the network into four clusters: environmental policy and regulation, sustainability practices, greenwashing risks and controversies, and consumer and market responses. This indicates that greenwashing-related discourse is not organized around a single topic but can be structured across multiple thematic dimensions.

5.2. Theoretical Implications

First, prior greenwashing research has largely relied on survey-based or experimental designs, which has limited its ability to capture large-scale public discourse in a naturally occurring manner (Schmuck et al., 2018). To address this limitation, this study adopted a big data approach utilizing online text data and analyzed public perceptions of greenwashing from multiple angles by combining frequency analysis, network analysis, centrality analysis, and CONCOR analysis. Compared to prior studies that have depended on single analytical methods, this study goes beyond simple keyword frequency to examine the structural relationships among concepts and the overall architecture of greenwashing discourse, representing a meaningful methodological contribution (Lee & Yhang, 2025).

Second, this study contributes to the theoretical literature by revisiting and extending existing theoretical frameworks at the level of online discourse structure. According to legitimacy theory, firms attempt to maintain legitimacy by communicating in ways that align with social expectations (Suchman, 1995). However, the high frequency of keywords such as deception, uncertainty, and confusion in the frequency analysis suggests that corporate ESG communication may in fact be reinforcing public skepticism rather than alleviating it. The finding that trust occupies a structurally peripheral position in the centrality analysis further indicates that corporate ESG signals are not translating into public trust, a result that is consistent with Nguyen's (2026) concept of the trust gap.

Furthermore, the four clusters derived from the CONCOR analysis reveal that public perceptions of greenwashing are structured not along a single dimension, but across multiple dimensions encompassing regulation, practice, risk, and market response. Consistent with stakeholder theory (Freeman, 1984), this finding confirms that greenwashing is evaluated simultaneously from the perspectives of diverse stakeholder groups including consumers, regulators, and investors, thereby extending the discussion of prior research that has largely examined greenwashing along a single dimension (Delmas & Burbano, 2011).

Finally, this study contributes to the literature by exploratively examining the structure of greenwashing discourse within the specific socio-temporal context of the post-pandemic era. As social demands for ESG transparency have intensified in the aftermath of the pandemic (Adams & Abhayawansa, 2022), this study empirically demonstrates how such changes are reflected in the structure of online public discourse, offering a timely perspective for understanding the rapidly evolving landscape of greenwashing.

5.3. Practical Implications

First, the frequency and centrality analyses identified eco-friendliness, ESG management, sustainability, and CSR as the most frequently occurring and structurally central keywords, suggesting that Korean consumers and stakeholders place considerable emphasis on these ESG-related concepts when evaluating corporate legitimacy. This finding serves as an important warning signal for firms engaged in greenwashing. Firms should move beyond simply emphasizing a green image and transition toward substantive communication that incorporates ESG management systems, sustainability goals, and measurable performance indicators. Firms that fail to provide specific and verifiable information, such as carbon emission reduction figures, ESG ratings, and certification status, must recognize that they face a growing risk of losing the trust of consumers and investors in the Korean market and beyond.

Second, the high frequency of keywords such as deception, uncertainty, and confusion in the frequency analysis reflects the negative perceptions that the Korean online public holds alongside their expectations of corporate environmental responsibility. Firms engaged in greenwashing may be able to generate a degree of positive image in the short term but are likely to face negative consequences in

the long run, including erosion of consumer trust. Therefore, firms should immediately correct exaggerated environmental claims and establish institutionalized systems that mandate internal verification processes and legal review before any ESG-related messages are communicated externally. Given the increasingly stringent regulatory trend surrounding greenwashing, developing clear guidelines and standards for advertising and promotional communications is becoming not a matter of choice but an essential requirement.

Third, the Environmental Policy and Regulation cluster identified through the CONCOR analysis suggests that the Korean online public strongly associates greenwashing with the regulatory and policy environment. This indicates that firms should integrate ESG not merely as a marketing tool, but as a core element of regulatory compliance and risk management strategy. Environmental indicators such as waste, energy, and carbon emissions have already become key evaluation criteria for major regulatory bodies and investors, and firms that fail to set these as key performance indicators (KPIs) and pursue genuine improvements face a significantly heightened risk of serious business consequences, including exclusion from supply chains and regulatory sanctions. Furthermore, the Sustainability Practices cluster demonstrates that the public evaluates corporate sustainability not through general claims, but through specific and verifiable mechanisms such as eco-friendly certification and eco-labels. Building a sustainable management system grounded in substantive environmental performance, rather than relying on eco-labels and certification schemes in a formalistic manner, is therefore a fundamental requirement for securing long-term competitiveness.

Finally, the Consumer and Market Responses cluster, in which keywords such as purchase intention, brand image, trust, and profitability form a single cluster, demonstrates that greenwashing is ultimately evaluated in terms of its consequences for consumer behavior and corporate performance. This is consistent with prior research showing that greenwashing leads to reduced consumer trust and diminished purchase intention (Chen & Chang, 2013). When greenwashing is exposed, corporate image and consumer trust can deteriorate rapidly, potentially causing severe damage to brand value. Therefore, rather than pursuing short-term marketing gains, firms should seek a strategic transition toward building long-term trust relationships with consumers through authentic communication grounded in substantive ESG performance.

5.4. Limitations and Future Studies

This study has the following limitations. First, as this study relies solely on online text data, it has inherent limitations in capturing a deeper and more nuanced understanding of public perceptions. Furthermore, although sentiment plays a critical role in shaping perceptions of greenwashing, this study primarily focuses on the structural aspects of textual data, resulting in relatively limited analysis of the emotional and attitudinal dimensions of public discourse.

Second, the key perception factors identified in this study have not been empirically examined in relation to behavioral outcomes such as satisfaction and purchase intention. As a result, the practical implications for consumer behavior remain somewhat limited.

Third, the positive and negative effects of greenwashing may vary depending on the level, timing, and contextual environment in which it occurs. However, this study was unable to fully account for the actual degree of greenwashing practices or the diverse environmental factors that may influence its effects.

Fourth, the findings of this study are based on Korean online public discourse and should therefore be interpreted within the Korean context. The extent to which these findings can be generalized to other national or cultural contexts remains limited, and cross-national comparative research utilizing data from multiple countries would be needed to assess the broader applicability of the results.

Fifth, this study did not restrict its scope to a specific industry or business sector, instead examining public perceptions of greenwashing across society as a whole. As a result, industry-specific patterns and differences in greenwashing perceptions could not be examined in detail. Future research targeting specific industries such as hospitality and tourism, fashion, or energy would enable the derivation of more specific and actionable practical implications.

Based on these limitations, the following directions are proposed for future research. First, incorporating semi-structured surveys and interviews alongside text-based analysis would enable the derivation of more in-depth insights and strengthen the practical relevance of the findings. In addition, the integration of sentiment analysis is recommended to capture the emotional and attitudinal dimensions embedded in online text data, which the present study was unable to fully address. Furthermore, empirically examining the causal relationships between key greenwashing-related variables identified in this study, such as deception, trust, and consumer responses, and behavioral outcomes such as purchase intention would provide more meaningful and robust contributions to the existing literature. Finally, cross-national comparative studies and industry-specific analyses of greenwashing perceptions would enable the proposal of more specific and actionable implications for both practitioners and policymakers.

Funding: This research received no external funding.

Institutional Review Board Statement: This study followed ethical standards for human research. Participants were informed about the study purpose, voluntary participation, confidentiality, and their right to withdraw at any time, and all data were used solely for academic purposes.

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

Conflicts of Interest: The author declares no conflict of interest.

References

Adams, C. A., & Abhayawansa, S. (2022). Connecting the COVID-19 pandemic, environmental, social and governance (ESG) investing and calls for 'harmonisation' of sustainability reporting. *Critical Perspectives on Accounting, 82*, Article 102309. https://doi.org/10.1016/j.cpa.2021.102309

Bagatini, F. Z., Perin, M. G., & de Medeiros, J. F. (2026). How consumers contest legitimacy: Skepticism toward corporate social responsibility. *Corporate Social Responsibility and Environmental Management.* https://doi.org/10.1002/csr.70532

Borgatti, S. P. (2005). Centrality and network flow. *Social NetWorks, 27*(1), 55–71. https://doi.org/10.1016/j.socnet.2004.11.008

Chen, Y. S., & Chang, C. H. (2013). Greenwash and green advertising: The roles of perceived credibility and perceived green image. *Journal of Business Ethics, 114*(3), 489–505. https://doi.org/10.1007/s10551-012-1453-7

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management, 37*(1), 39–67. https://doi.org/10.1177/0149206310388419

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*

European Commission. (2020). *Screening of websites for 'green' claims*. *https://ec.europa.eu/info/sites/default/files/ec_green_claims_report_final.pdf*

Freeman, R. E. (1984). *Strategic management: A stakeholder approach.* Pitman.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Greg, N., Gangadharbatla, H., & Paladino, A. (2014). Perceived greenwashing: The interactive effects of green advertising and corporate environmental performance on consumer reactions. *Journal of Business Ethics, 125*(4), 693706. https://doi.org/10.1007/s10551-013-1944-3

Humphreys, A., & Wang, R. J. H. (2018). Automated text analysis for consumer research. *Journal of Consumer Research, 44*(6), 1274–1306. https://doi.org/10.1093/jcr/ucx104

IBM Institute for Business Value. (2024). *Consumer sustainability survey*. IBM. https://www.ibm.com/thought-leadership/institute-business-value/report/consumer-sustainability

Ioannou, I., Kassinis, G., & Vafeas, N. (2023). The greenwashing trap: Do consumers' perceptions of greenwashing affect their trust in green brands? *Journal of Business Ethics, 185*(2), 333–352. https://doi.org/10.1007/s10551-022-05131-5

Lee, S.-B. (2023). The impact of value co-creation on performance through engagement in a non-face-to-face service environment: Focusing on live streaming [Doctoral dissertation, Pusan National University].

Lee, S.-B. (2025). The effect of genre-structured themes of film scenario as cultural contents on determinants of viewing intention: Mixed research methods [Doctoral dissertation, Pukyong National University].

Lee, S.-B., Yhang, W.-J., & Kim, K.-M. (2024). A study on audience perception on the identity trend of 'Jeonju International Film Festival' using big data analysis. *Culture & Convergence, 46*(2), 261–273.

Lee, S.-B., & Yhang, W.-J. (2024). Analysis of trends in linguistic perception of text-based complex cultural spaces - Focusing on Busan Opera House. *Journal of Industrial Innovation, 40*(2), 383–397.

Lee, S.-B., & Yhang, W.-J. (2025a). Big data-based text mining of the Tongyeong International Music Festival for cultourism strategies. *Journal of Convergence Tourism Contents, 11*(4), 111–128.

Lee, S.-B., & Yhang, W.-J. (2025b). Language analysis of film scenarios using big data and strategies for cultural tourism. *Journal of Convergence Tourism Contents, 11*(2), 73–88.

Luchs, M. G., Naylor, R. W., & Raghunathan, R. (2010). The sustainability liability: Potential negative effects of ethicality on product preference. *Journal of Marketing, 74*(5), 18-31. https://doi.org/10.1509/jmkg.74.5.018

Lyon, T. P., & Maxwell, J. W. (2011). Greenwash: Corporate environmental disclosure under threat of audit. *Journal of Economics & Management Strategy, 20(*1), 3–41. https://doi.org/10.1111/j.1530-9134.2010.00282.x

Nguyen, T. H. (2026). The mediating role of trust in bridging the ESG assurance credibility chasm. *Discover Sustainability, 7,* Article 369. https://doi.org/10.1007/s43621-026-02747-3

Persakis, A., et al. (2025). Greenwashing in the literature: A bibliometric analysis. *Journal of Cleaner Production, 435*, Article 140123. https://doi.org/10.1016/j.jclepro.2024.140123

Pizzetti, M., Gatti, L., & Seele, P. (2021). Firms talk, suppliers walk: Analyzing the locus of greenwashing in the blame game and introducing 'vicarious greenwashing'. *Journal of Business Ethics, 170*(1), 21–38. https://doi.org/10.1007/s10551-019-04406-2

Schmuck, D., Matthes, J., & Naderer, B. (2018). Misleading consumers with green advertising? An affect–reason–involvement account of greenwashing effects in environmental advertising. *Journal of Advertising*, *47*(2), 127–145. https://doi.org/10.1080/00913367.2018.1452652

Spaniol, M. J., Danilova-Jensen, E., Nielsen, M., Rosdahl, C. G., & Schmidt, C. J. (2024). Defining greenwashing: A concept analysis. *Sustainability, 16*(20), Article 9055. https://doi.org/10.3390/su16209055

Su, Y., Zhu, C., & Albitar, K. (2026). Market competition and corporate ESG greenwashing: A perspective from new entrants. *Business Strategy and the Environment, 35*(2), 2266–2294. https://doi.org/10.1002/bse.70282

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review, 20*(3), 571–610. https://doi.org/10.2307/258788

The Harris Poll. (2022). *Executive survey on sustainability.* https://theharrispoll.com/

Tirunillai, S., & Tellis, G. J. (2012). Does chatter really matter? Dynamics of user-generated content and stock performance. *Marketing Science, 31*(2), 198–215. https://doi.org/10.1287/mksc.1110.0682

White, K., Habib, R., & Hardisty, D. J. (2019). How to SHIFT consumer behaviors to be more sustainable: A literature re-

view and guiding framework. *Journal of Marketing, 83*(3), 22–49. https://doi.org/10.1177/0022242919825649

Zhan, X., Zhan, Y., & Li, G. (2025). The cost of ESG decoupling: How misaligned disclosures undermine corporate sustainability? *SAGE Open.* https://doi.org/10.1177/21582440251369170

Received 2026.04.12 - Revised 2026.06.18 - Accepted 2026.06.25