



# “Institutional differences in CEO sustainability messages: Text mining of Korean and global firms’ reports”

## AUTHORS

Tae-Il Yoon  

## ARTICLE INFO

Tae-Il Yoon (2026). Institutional differences in CEO sustainability messages: Text mining of Korean and global firms’ reports. *Problems and Perspectives in Management*, 24(1), 129-147. doi: [10.21511/ppm.24\(1\).2026.10](https://doi.org/10.21511/ppm.24(1).2026.10)

## DOI

[http://dx.doi.org/10.21511/ppm.24\(1\).2026.10](http://dx.doi.org/10.21511/ppm.24(1).2026.10)

## RELEASED ON

Thursday, 22 January 2026

## RECEIVED ON

Monday, 15 September 2025

## ACCEPTED ON

Tuesday, 13 January 2026

## LICENSE



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

## JOURNAL

"Problems and Perspectives in Management"

## ISSN PRINT

1727-7051

## ISSN ONLINE

1810-5467

## PUBLISHER

LLC “Consulting Publishing Company “Business Perspectives”

## FOUNDER

LLC “Consulting Publishing Company “Business Perspectives”



NUMBER OF REFERENCES

**37**



NUMBER OF FIGURES

**11**



NUMBER OF TABLES

**4**

© The author(s) 2026. This publication is an open access article.



BUSINESS PERSPECTIVES



LLC "CPC "Business Perspectives"  
Hryhorii Skovoroda lane, 10,  
Sumy, 40022, Ukraine  
[www.businessperspectives.org](http://www.businessperspectives.org)

**Type of the article:** Research Article

**Received on:** 15<sup>th</sup> of September, 2025

**Accepted on:** 13<sup>th</sup> of January, 2026

**Published on:** 22<sup>nd</sup> of January, 2026

© Tae-Il Yoon, 2026

Tae-Il Yoon, Ph.D., Professor,  
Department of Advertising & Public  
Relations, Hallym University, South  
Korea.

Tae-Il Yoon (South Korea)

# INSTITUTIONAL DIFFERENCES IN CEO SUSTAINABILITY MESSAGES: TEXT MINING OF KOREAN AND GLOBAL FIRMS' REPORTS

## Abstract

Sustainable management and corporate sustainability communication (CSC) have become central themes in both research and practice. Within sustainability reports, CEO messages serve as strategic communication channels that reveal how firms construct legitimacy and align with environmental, social, and governance (ESG) imperatives. Despite their importance, limited research has compared these messages across institutional contexts. This study aims to uncover institutional differences in CEO messages within sustainability reports of Korean and global firms. For this purpose, this study analyzes CEO messages from 343 Korean and 323 global sustainability reports (from Fortune Global 500 firms following GRI standards) using text mining techniques (e.g., semantic network analysis, word embedding, and topic modeling). Results of semantic network analysis show that Korean firms emphasize management (centrality = 0.345922), growth (0.32013), and market (0.152733), while global firms highlight community (0.362553), customer (0.331332), and commitment (0.182617). Word embedding indicates that the terms "sustainability" and "stakeholder" are linked to symbolic gratitude and external recognition in Korean messages, but to citizenship and responsibility in global ones. Topic modeling reveals four performance- and evaluation-oriented themes in Korean firms versus three execution- and long-term value-oriented themes in global firms. These findings show how different levels of institutionalization shape CEO discourse and suggest that Korean firms rely more on symbolic legitimation, while global firms stress substantive sustainability commitments.

## Keywords

sustainable management, corporate sustainability communication, CEO messages, legitimacy theory, institutional theory, text mining

## JEL Classification

M14, Q56, D22

## INTRODUCTION

Sustainability has evolved from a peripheral concern into a strategic imperative that shapes how firms create value and respond to societal expectations. Regulatory changes, global initiatives such as the Sustainable Development Goals (SDGs), and the growing influence of investors and civil society have pushed companies to institutionalize sustainable management as part of their corporate strategy. In this context, corporate sustainability communication has become a key arena in which firms articulate their commitments, justify their actions, and negotiate legitimacy with multiple stakeholders.

Academic research has shown that sustainability communication is not merely informational but deeply embedded in legitimacy-seeking processes. From a legitimacy theory perspective, companies use sustainability discourse to signal that their practices are desirable, proper, and appropriate within a socially constructed system of norms and values. Institutional theory complements this view by emphasizing how organizations respond to coercive, normative, and mimetic pressures, which can lead to both substantive change and symbolic con-



This is an Open Access article, distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.



### Conflict of interest statement:

Author(s) reported no conflict of interest

formity. Together, these perspectives highlight the interplay between strategic intent and institutional constraints in shaping the language and content of sustainability communication.

CEO messages, typically located at the beginning of sustainability reports, occupy a distinctive position in this communication landscape. They provide an authoritative narrative about how the company understands sustainable management, how it views its stakeholders, and how it situates its performance and commitments in a broader institutional environment. Prior studies suggest that CEO messages can be used for impression management, to reinforce corporate identity, and to frame corporate responsibility and future orientation. However, most existing papers have focused on single-country contexts or small samples, and seldom compared CEO messages across institutional settings using large-scale text mining techniques.

Against this backdrop, the scientific problem guiding this study is how institutional pressures influence the framing of CEO messages in sustainability reports and how these messages contribute to different legitimation strategies. Specifically, this study examines institutional differences in CEO sustainability messages between Korean and global firms. Addressing this problem contributes to an empirical explanation of how the structure, associations, and thematic content of CEO discourse reflect differing levels and forms of sustainability institutionalization.

---

## 1. LITERATURE REVIEW

### 1.1. Theoretical background of sustainable management

Sustainable management encompasses activities that promote economic, social, and environmental accountability (Kang, 2022). Prior studies identify two perspectives: resource-based and institutional (Branco & Rodrigues, 2008). From a resource-based perspective, signaling theory posits that companies engage in sustainable management and actively disclose related information to signal superior performance to shareholders and potential investors, thereby enhancing economic returns. Supporting studies (Lai et al., 2016; Mahoney et al., 2013) have found that high-performance firms tend to be more proactive in sustainability disclosures.

However, empirical evidence also shows that low-performing firms sometimes disclose sustainability information more aggressively (Cho et al., 2012), suggesting that disclosure can serve purposes beyond signaling performance. This is where legitimacy theory and institutional theory provide alternative explanations (Bengtson & Mossberg, 2023).

According to legitimacy theory, sustainable management is often a mandatory activity undertaken in response to stakeholder legitimacy pressures

(Fernando & Lawrence, 2014; Marschlich, 2022). Suchman (1995) categorizes legitimacy into three types: pragmatic legitimacy, which emphasizes delivering economic and social benefits; moral legitimacy, which emphasizes conformity with societal norms and values; and cognitive legitimacy, in which corporate activities are understood and accepted as appropriate without question. Legitimacy pressures may come from diverse stakeholders (e.g., shareholders, consumers, employees, media, NGOs, etc.) and take various forms (e.g., internal vs. external, direct vs. indirect, top-down vs. bottom-up, etc.) (Yoon, 2022). Companies respond with substantive legitimation, which involves actual changes to align with sustainable principles, or symbolic legitimation, which emphasizes public communication of such efforts (Eliwa et al., 2021).

Institutional theory (DiMaggio & Powell, 1983) complements legitimacy theory by focusing on how external structures influence organizational behavior. Institutional pressures, such as coercive (regulatory), mimetic (imitation of peers), and normative (professional expectations) pressures, affect how companies disclose sustainability practices. For example, adopting standardized reporting frameworks, such as the Global Reporting Initiative (GRI), constitutes a normative response to institutional expectations. Previous research demonstrated how industry regulations and gov-

ernance structures shape both the frequency and content of disclosures (Aguilera et al., 2021; Herbert & Graham, 2021; Pasamar et al., 2023). Furthermore, companies in highly institutionalized environments tend to incorporate codified ESG frameworks into communications. As sustainability becomes a global imperative, firms are under increasing pressure to converge in managerial practices, with CEO messages aligning more closely to international standards (e.g., GRI), reflecting both mimetic isomorphism and stakeholder demand for comparability (Bengtson & Mossberg, 2023; Pasamar et al., 2023).

## 1.2. Corporate sustainability communication and CEO messages

Corporate sustainability communication refers to strategic corporate communication aimed at sharing and engaging with stakeholders about a company's sustainable management practices and achievements, thereby contributing to sustainable societal development (Godemann & Michelsen, 2011; Signitzer & Prexl, 2007). Among CSC channels, sustainability reports are the most prominent. Prepared in line with the UN-established GRI guidelines, these reports serve as formal disclosures of a company's sustainability efforts.

CEO messages, typically placed at the beginning of sustainability reports, have evolved from ceremonial prefaces to strategic communication tools. Acting as "adaptive frames" (Arvidsson & Sabelfeld, 2023), these messages respond dynamically to changing institutional contexts, regulatory developments, industry norms, and stakeholder expectations. They provide a rhetorical and symbolic platform through which firms articulate their ESG commitments and align their corporate identity with societal values (Hong & Cho, 2024).

CEO messages have significant strategic implications. During economic crises, for example, CEOs of surviving firms explicitly integrated survival strategies into their communications (D'Aveni & MacMillan, 1990). Similarly, forward-looking CEO language in annual reports has been linked to enhanced organizational innovativeness in sensing and developing new technologies (Yadav et al., 2007).

Early studies found that CEO messages often seek to minimize negative performance, diffuse responsibility, or elicit favorable reactions from investors (Abrahamson & Park, 1994). In the sustainability context, they serve as legitimacy narratives that explicitly link company strategy, performance, and values to broader environmental and social goals (Al Amosh, 2025; Berkin et al., 2025; Braga et al., 2024). However, longitudinal analysis of CEO messages in sustainability and annual reports found them to be often unsystematic and more focused on impression management than on accountability (Barkemeyer et al., 2014).

Both content and language style matter. Messages may be strategically timed and formatted to enhance persuasive effect (Jegadeesh & Wu, 2013; Davis & Tama-Sweet, 2012). Linguistic analysis reveals that firms with weaker sustainability performance use more positive or optimistic expressions, highlight future achievements, and include more environmental terminology (Herbert & Graham, 2021). Placement of positive content at the beginning or end of messages has also been observed as a strategic choice (Boudt & Thewissen, 2019). Na et al. (2020) criticized CEO messages for overemphasizing positive language and failing to reflect financial realities.

In response to institutional pressures surrounding sustainable management, corporate strategies aimed at securing legitimacy have become increasingly important in a globalized business environment. Corporate behaviors deemed acceptable in a domestic context may not be tolerated in culturally different global settings; thus, sustainable management that strengthens corporate citizenship serves as a vital means of securing legitimacy (Bachmann & Ingenhoff, 2016).

Although there has been a trend toward standardization of sustainability reports globally, CEO messages still vary depending on industry environment, cultural context, stage of institutionalization, and the level of legitimacy pressures. Aguilera et al. (2021) suggest that corporate governance structures influence the depth and consistency of sustainability disclosures. For example, firms with dedicated sustainability committees or ESG leadership tend to produce more robust

CEO narratives that integrate long-term goals and stakeholder dialogue (Berkin et al., 2025; Braga et al., 2024).

Even among multinational corporations, national cultural norms and regulatory environments shape differences in how CEO messages in sustainability reports are framed. In particular, Korean companies, subject to higher legitimacy pressures than their global counterparts, tend to refrain from overt emotional expression in their CEO messages (Yoon & Byun, 2023).

In sum, prior research shows that sustainable management and CSC are key mechanisms for securing legitimacy, and that institutional contexts shape how firms adopt and communicate sustainability practices. Yet there is a lack of large-scale, comparative research on CEO messages in sustainability reports across different institutional environments. This study addresses this gap by applying text-mining methods to the CEO messages in sustainability reports from Korean and global companies.

To achieve this aim, the study adopts a theory-informed, exploratory text-mining design. Although legitimacy and institutional theory provide a conceptual basis for expecting cross-contextual differences in sustainability framing, the study's main analytical techniques are unsupervised and aim to identify latent structures in large-scale corpora. Given the limited prior comparative evidence that would justify precise directional predictions, the study is framed through research questions rather than confirmatory hypotheses. Thus, three research questions are proposed as follows:

*RQ1: How do the semantic networks and centrality of keywords in CEO messages of sustainability reports differ between Korean and global companies?*

*RQ2: How do the associated words of core concepts (e.g., sustainability and stakeholder) in CEO messages of sustainability reports differ between Korean and global companies?*

*RQ3: How do the topics covered in the CEO messages of sustainability reports differ between Korean and global companies?*

## 2. METHODOLOGY

### 2.1. Research design

This study adopts a comparative, text-mining research design to examine institutional differences in CEO messages across sustainability reports. CEO messages from Korean and global firms are analyzed by employing a four-step text mining approach: 1) data collection, 2) preprocessing, 3) text mining analysis, and 4) interpretation. Three text-mining techniques are applied to examine three research questions (RQ1–RQ3): semantic network analysis (SNA) addresses RQ1 by mapping keyword structures and centrality patterns; word embedding addresses RQ2 by examining associative word patterns around core concepts; and topic modeling addresses RQ3 by identifying underlying themes in CEO discourse.

### 2.2. Data

The data analyzed in this study consist of CEO messages from English-language sustainability reports of Korean and global companies. During data collection, efforts were made to ensure equivalence by maintaining as much similarity as possible between the two groups' reports. To achieve this, only English-language sustainability reports from both Korean and global companies were selected, and only those that adhered to standardized reporting guidelines (e.g., GRI) were included.

To maintain a similar industry distribution across the two groups, English-language reports from Korean companies were collected first, and then matching reports from global companies in the same industries were selected accordingly.

Korean companies' reports were collected from the Sustainability Report Database of the Korean Standards Association (KSA) and the GRI Sustainability Disclosure Database. The GRI, an UN-affiliated organization, provides internationally recognized guidelines for sustainability reporting and regularly updates these standards. Companies around the world follow these guidelines when producing their sustainability reports, which are in turn used as fundamental resources for international sustainability evaluations. Each report in the GRI database also specifies which of the 38 GRI sectors it belongs to.

When collecting the Korean reports, the industry classification of each report was identified, resulting in coverage across 33 sectors, as shown in Appendix A. Only reports from the past seven years (2013–2019) that complied with GRI 3.1 or later guidelines were included.

Corresponding reports from global companies were also collected from the GRI database. These global firms were selected from the Fortune Global 500, an annual list of the world's top 500 companies published by Fortune magazine. Many studies (e.g., Legendre & Coderre, 2013) analyze this group due to their high level of corporate disclosure.

From the 500 companies, firms in the five sectors not represented by Korean companies were excluded. For each of the 33 sectors represented by Korean companies, the top-ranking global firms were selected in numbers matching those of the Korean companies in each sector. For example, if there were two Korean companies in the chemical sector (Sector 4 by GRI classification) in 2013, the top two global companies in that sector were selected accordingly. As with the Korean reports, any global report not adhering to GRI 3.1 or later was excluded. Additionally, even if a Korean company was listed among the Fortune Global 500, it was still classified as a Korean company.

Through this process, a total of 666 English-language sustainability reports were collected: 343 from Korean companies and 323 from global companies. From these reports, only the CEO message sections were extracted, organized into a separate Excel file, and analyzed using Python-based text mining techniques to examine the semantic structure, the construction of core concepts, and the main issues in the text.

### 2.3. Text mining techniques for analysis

Text mining, one of the big data analysis methods in managerial research, involves summarizing, classifying, clustering, and identifying features from unstructured data. Among the various Python-based text mining techniques, this study used the following three methods (Bengfort et al., 2018; Ignatow & Mihalcea, 2016).

Semantic network analysis was used to address RQ1 by examining the semantic structure inherent in corpora. To assess the structure, centralities of the network (e.g., degree, betweenness, closeness, and eigenvector centralities) were analyzed. Since these centralities tend to correlate highly, this study focused on eigenvector centrality. For this study, the top 20 keywords were extracted based on simple frequency, and their semantic network was visualized using Python's Network X library.

To examine RQ2, word embedding (Word2Vec) was used to analyze the associative words of core concepts. There are two main approaches: CBOW (Continuous Bag of Words) and Skip-gram. This study used the Skip-gram method, which predicts nearby words based on a target word. All words were mapped onto a two-dimensional plane using Python's t-SNE library. Each word was assigned a coordinate (x, y), and the semantic similarity between words was calculated based on proximity. This allowed analysis of words most closely associated with core terms such as sustainability or stakeholder.

Topic modeling was used to address RQ3 by identifying latent themes in the CEO-message corpora. The study employed Latent Dirichlet Allocation (LDA) using Python's gensim library (Blei et al., 2003). The number of candidate topics was evaluated using coherence scores, and the final topic solution was selected by balancing coherence, parsimony, and interpretability. For each retained topic, the study reports the most distinctive keywords and their reflection rates (topic-word probabilities) after removing common high-frequency terms shared across topics. These results are visualized using bar charts, and figure notes summarize key preprocessing and visualization settings for reproducibility. Each figure includes an explicit caption and a brief note describing what is plotted (e.g., units, what node size/edges represent, and model-selection logic).

## 3. RESULTS

To examine research question 1 (RQ1), the sustainability reports of 343 Korean companies and 323 global companies were analyzed. First, in

both the Korean and global company datasets, all words (excluding stop words) were processed, and the simple frequency of occurrence for each word was calculated. The top 20 keywords were then extracted based on frequency. Figure 1 visualizes the semantic network of the 20 keywords extracted from the Korean companies' reports, while Figure 2 shows the semantic network for the 20 keywords extracted from the global companies' reports. In both the Korean and global companies' semantic networks, almost all keywords were interconnected, meaning that each of the 20 keywords appeared together at least once within the same sentence. This is likely due to the vast volume of text data comprising the entire report corpus.

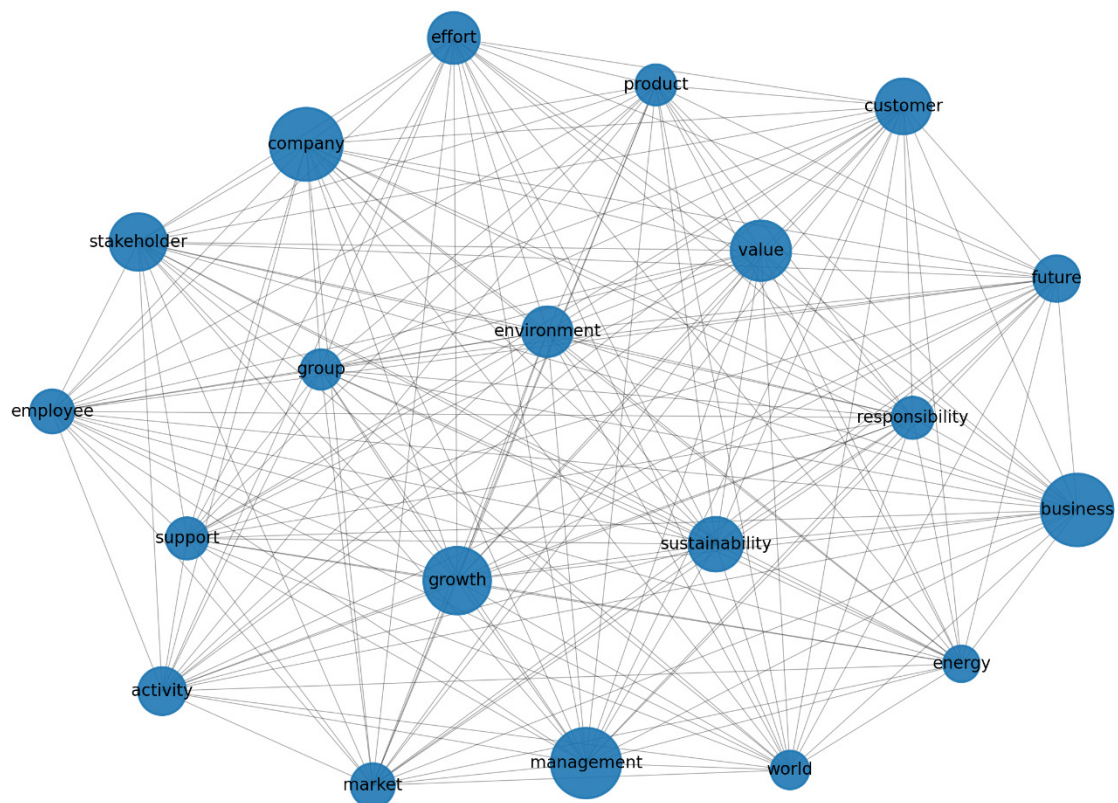
Based on the semantic network analysis, centrality indices were calculated for the 20 keywords. While simple frequency reflects the quantitative weight of how often a given word appears in the

text, the centrality index indicates how effectively a word connects with other words within the semantic network. In other words, a high centrality score reflects high qualitative importance in the semantic structure. Table 1 presents, in descending order, the keywords and their centrality indices for both the Korean and global companies.

Among the keywords unique to Korean companies (i.e., absent in the global companies' list) were management, growth, stakeholders, efforts, future, etc. Compared to global companies, Korean companies also recorded relatively higher centrality scores for management, growth, future, market, etc.

Among the keywords unique to global companies were community, employee, people, commitment, etc. Compared to Korean companies, global companies recorded higher centrality scores

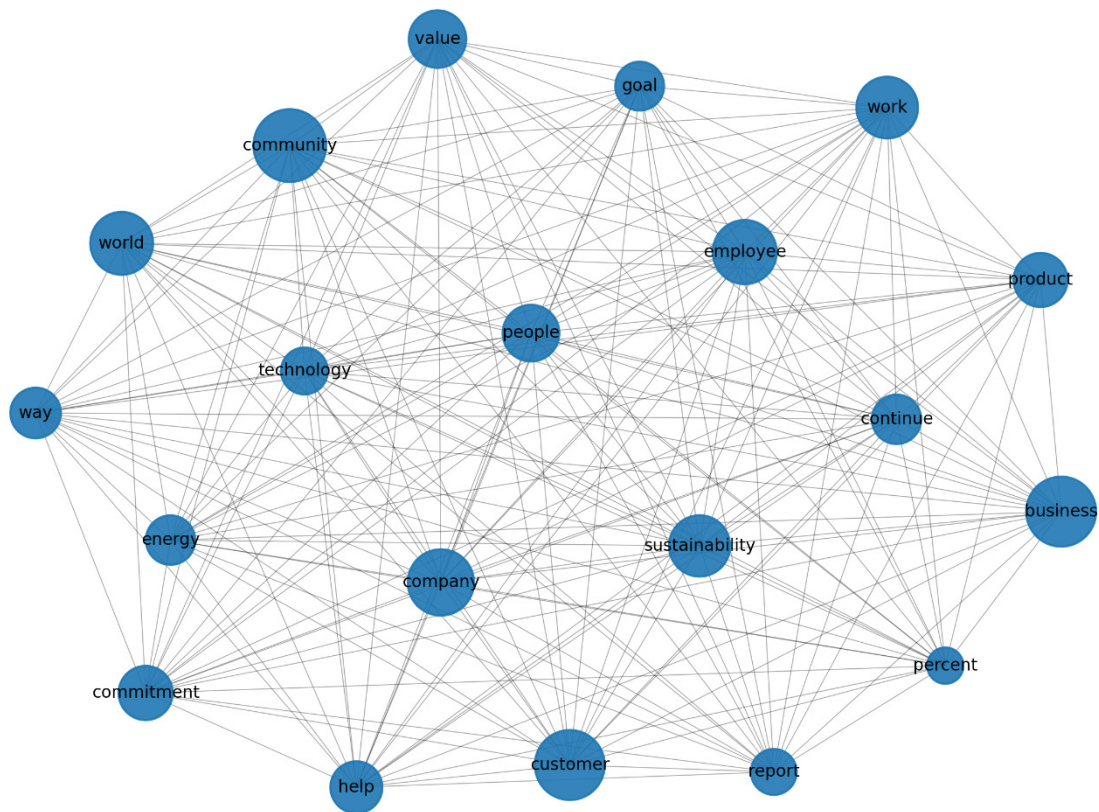
Semantic co-occurrence network (Top 20 keywords): Korean firms



Note: Node size represents eigenvector centrality; edges represent co-occurrence of keywords within the same sentence. The layout uses a force-directed algorithm (spring layout). Visualization is based on the top 20 keywords by frequency after stopword removal.

Figure 1. Semantic co-occurrence network (Top 20 keywords): Korean firms

## Semantic co-occurrence network (Top 20 keywords): Global firms



*Note:* Node size represents eigenvector centrality; edges represent co-occurrence of keywords within the same sentence. The layout uses a force-directed algorithm (spring layout). Visualization is based on the top 20 keywords by frequency after stopword removal.

**Figure 2.** Semantic co-occurrence network (Top 20 keywords): Global firms

for community, customer, employee, world, and commitment.

These patterns show that Korean firms' central keywords emphasize management, growth, and market, whereas global firms emphasize community, customers, and employees.

To address the research question 2 (RQ2), all words appearing in the corpora of Korean and global companies' sustainability reports were mapped into two dimensions using Python's t-SNE library. Words positioned close together on the two-dimensional plane are considered to be highly similar and thus identified as semantically associated terms. Based on this similarity analysis, the study examined the associated words of key concepts.

Table 2 presents the top 10 associated words, in order of similarity, to the core concept "sustain-

ability" for the Korean and global company corpora, respectively. As shown in Table 2, there was almost no overlap in the associated words between the two groups. This indicates that the meaning of sustainability is constructed within different contexts in the CEO messages of Korean and global companies' sustainability reports. In the Korean company corpus, notable associated words to sustainability include thank, express, appreciate, gratitude, support, etc. In the global company corpus, prominent associated words include success, business, serve, responsibility, etc.

Table 3 presents the top 10 associated words, in order of similarity, to the core concept "stakeholder" for the Korean and global company corpora, respectively. As shown in Table 3, the only common associated word across the two groups was 'report'; all other words differed. This also demonstrates that the contextual structuring of

**Table 1.** Keyword centrality values for Korean and global firms

Rank	Korea firm		Global firm	
	Word	Centrality	Word	Centrality
1	company	0.368285	community	0.362553
2	business	0.365257	business	0.331819
3	management	0.345922	customer	0.331332
4	growth	0.32013	company	0.293489
5	value	0.261279	employee	0.272189
6	stakeholder	0.239385	world	0.260113
7	customer	0.227219	work	0.248678
8	sustainability	0.217688	sustainability	0.244761
9	effort	0.198875	value	0.210946
10	environment	0.191567	people	0.203805
11	activity	0.174618	commitment	0.182617
12	future	0.166081	product	0.181845
13	market	0.152733	help	0.164816
14	employee	0.150812	way	0.154116
15	support	0.141547	energy	0.145142
16	responsibility	0.141361	continue	0.144811
17	product	0.135745	goal	0.141014
18	group	0.131949	technology	0.12638
19	world	0.123188	report	0.11927
20	energy	0.113108	percent	0.059614

stakeholder meaning differs between Korean and global companies. In the Korean company corpus, notable associated words to stakeholders include

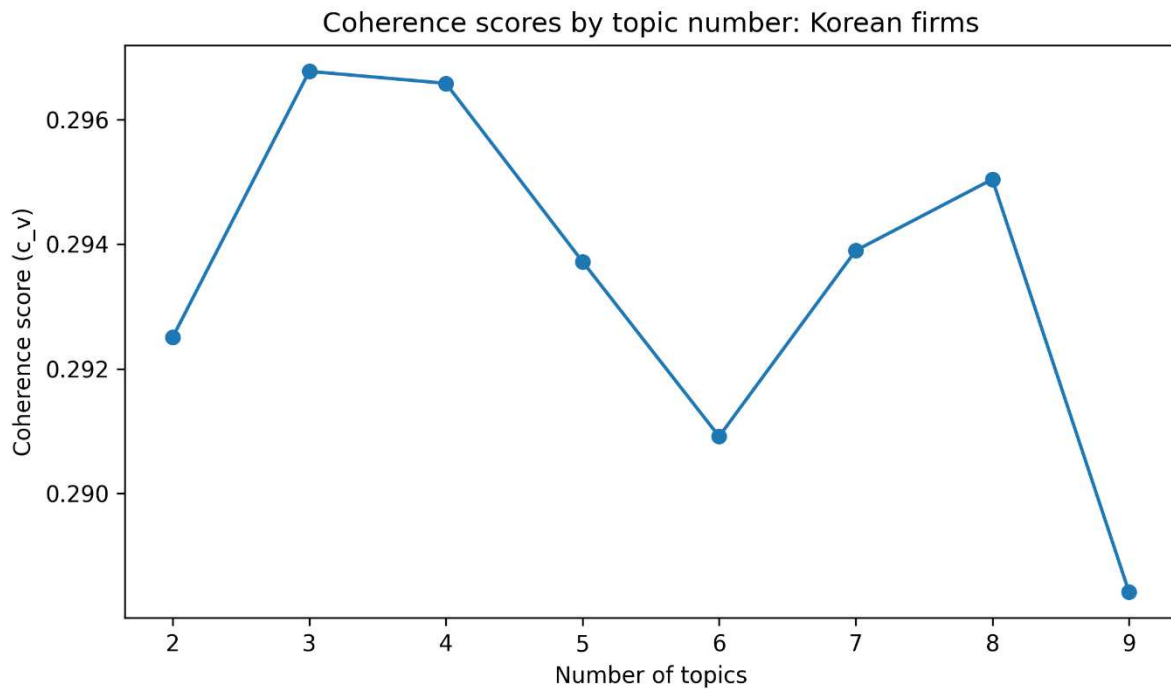
report, DJSI, DOW, JOHNS, Index, publication, etc. In the global company corpus, key associated words include citizenship, CSR, responsibility, etc.

**Table 2.** Associated words of “sustainability” for Korean and global firms

Rank	Korean firm		Global firm	
	Word	Similarity	Word	Similarity
1	thank	0.9991087317466736	success	0.9998946785926819
2	support	0.9990710020065308	business	0.9998894929885864
3	express	0.9989477396011353	work	0.9998655915260315
4	encouragement	0.9989175796508789	serve	0.9998590350151062
5	interest	0.9987897276878357	responsibility	0.9998540878295898
6	appreciate	0.9987293481826782	company	0.999846339225769
7	ask	0.9985685348510742	core	0.999817967414856
8	look	0.9984631538391113	people	0.9998157620429993
9	gratitude	0.9983395934104919	value	0.9998058080673218
10	shareholder	0.9979007244110107	employee	0.9998001456260681

**Table 3.** Associated words of “stakeholder” for Korean and global firms

Rank	Korean firm		Global firm	
	Word	Similarity	Word	Similarity
1	report	0.9974967241287231	report	0.9998564720153809
2	DJSI	0.9960989952087402	citizenship	0.9998285174369812
3	DOW	0.9957627654075623	strategy	0.9998277425765991
4	index	0.9955856800079346	CSR	0.9998225569725037
5	publishing	0.9955546855926514	progress	0.9998198747634888
6	pacific	0.9941987991333008	responsibility	0.9998124837875366
7	publication	0.9936447143554688	detail	0.9997998476028442
8	seventh	0.9931953549385071	principle	0.999796687889099
9	7th	0.9930688142776489	key	0.9997937679290771
10	JONES	0.99294114112854	effort	0.9997913241386414



Note: Coherence scores ( $c_v$ ) are computed for candidate topic numbers. The selected topic number balances coherence, parsimony, and interpretability.

**Figure 3.** Coherence scores by topic number: Korean firms

Overall, the terms “sustainability” and “stakeholder” in Korean firms were associated with symbolic gratitude and external recognition, whereas in global firms they were associated with substantive commitment.

To verify the research question 3 (RQ3), topic modeling (LDA) was applied separately to each corpus. To determine the optimal number of topics within the CEO messages, the coherence score was calculated for varying numbers of topics in both the Korean and global company datasets.

For Korean companies’ CEO messages, coherence scores suggested that a 3–4 topic solution best fit the corpus (Figure 3). Although the maximum coherence score ( $c_v$ ) was observed at three topics ( $n = 3$ ,  $c_v = 0.29677$ ), the four-topic solution ( $n = 4$ ,  $c_v = 0.29658$ ) provided clearer thematic separation and interpretability while maintaining a comparable coherence level. Therefore, the optimal number of topics was set at four for subsequent interpretation.

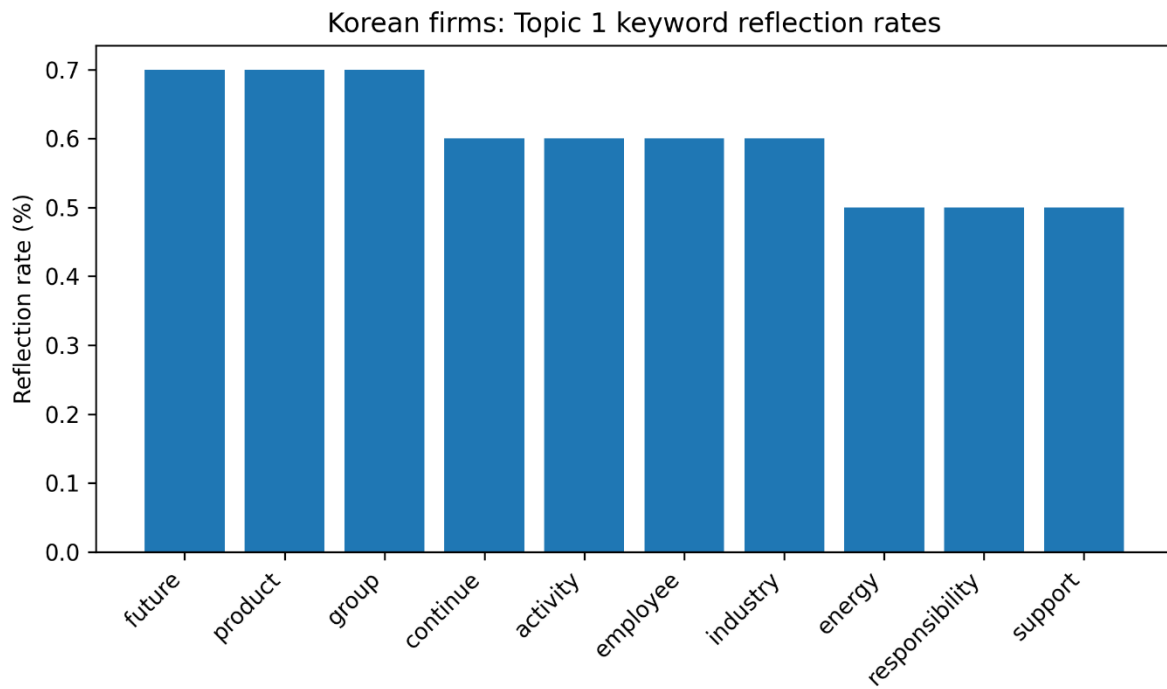
For further analysis of the individual topics in Korean companies’ CEO messages, high-frequency keywords common to all topics (e.g., business,

company, management, growth, stakeholder, etc.) were excluded, and distinctive keywords for each topic were extracted. To enhance interpretability, the reflection rates (topic–word probabilities) of the distinctive keywords were summarized and visualized using bar charts (Figures 4–7).

Four topics emerged in the CEO messages of Korean companies’ sustainability reports: (1) future product industries, (2) global community support, (3) energy projects, and (4) employee development.

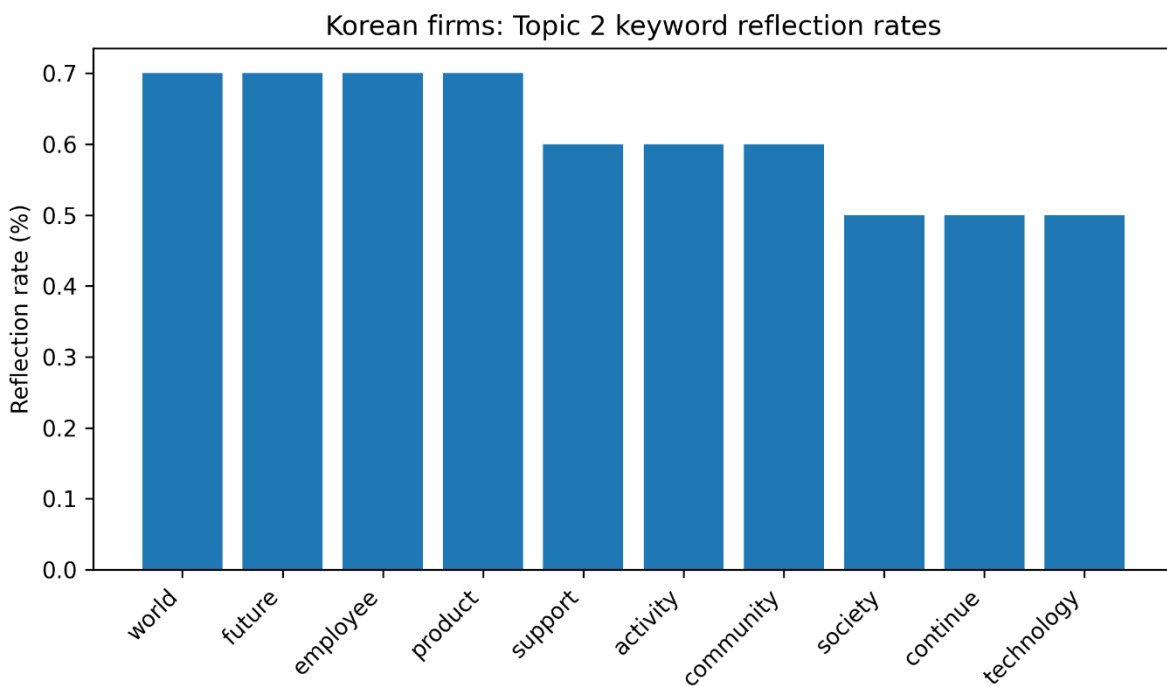
Topic 1 is characterized as “future product industries.” Unique keywords such as future, product, group, continue, activity, and responsibility strongly represent this topic (see Figure 4). The content primarily addresses industrial efforts aimed at the development of future products.

Topic 2 is identified as “global community support.” Unique keywords such as world, employee, community, support, and activity constitute this topic (see Figure 5). The emphasis is placed on activities supporting various stakeholder communities, including employees.



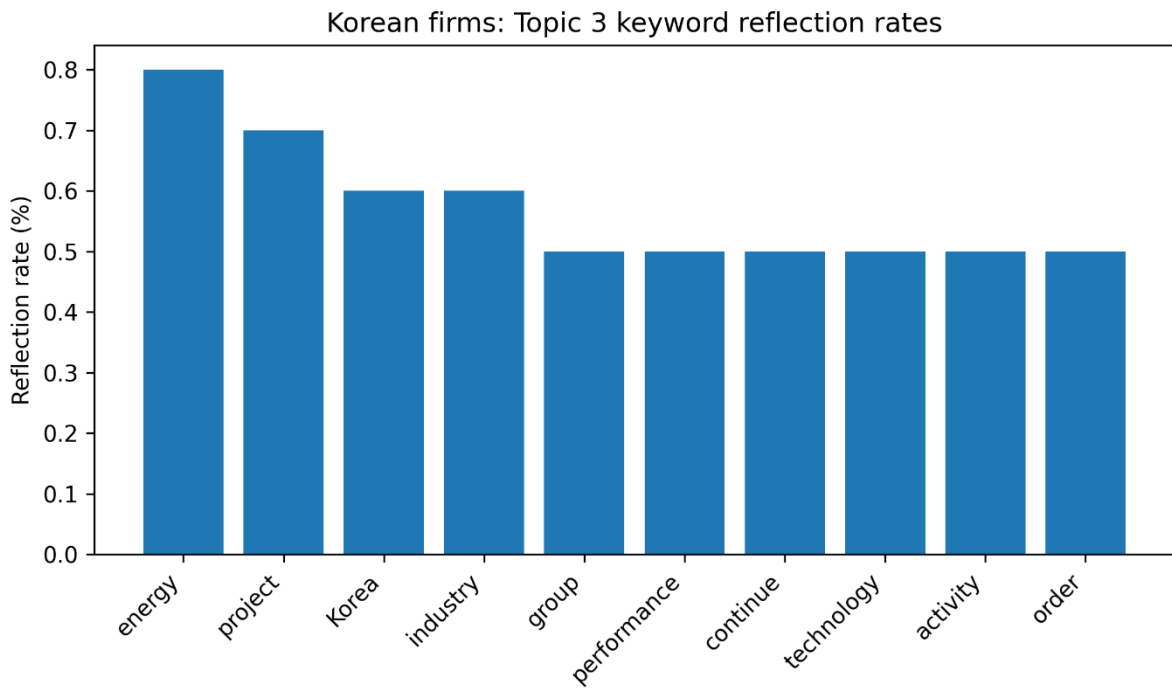
Note: Bars show the reflection rates (%) of the top distinctive keywords for Topic 1 after removing common high-frequency terms shared across topics.

Figure 4. Keyword reflection rates of topic 1 in Korean firms



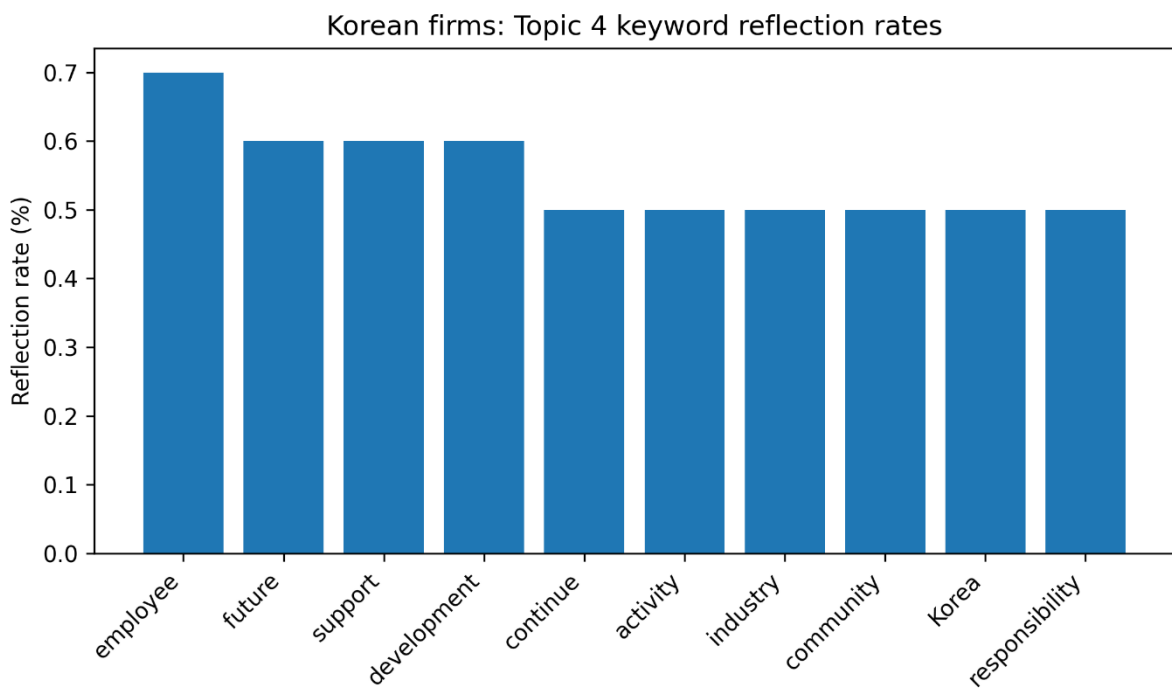
Note: Bars show the reflection rates (%) of the top distinctive keywords for Topic 2 after removing common high-frequency terms shared across topics.

Figure 5. Keyword reflection rates of topic 2 in Korean firms



Note: Bars show the reflection rates (%) of the top distinctive keywords for Topic 3 after removing common high-frequency terms shared across topics.

Figure 6. Keyword reflection rates of topic 3 in Korean firms



Note: Bars show the reflection rates (%) of the top distinctive keywords for Topic 4 after removing common high-frequency terms shared across topics.

Figure 7. Keyword reflection rates of topic 4 in Korean firms

Topic 3 is described as “energy projects.” This topic is defined by unique keywords such as energy, industry, project, Korean, and technology (see Figure 6). The content focuses primarily on achievements and technological developments within Korea’s energy industry.

Topic 4 is titled “employee development.” Unique keywords such as employee, future, support, development, and activity capture the essence of this topic (see Figure 7). It centers on responsibilities and support activities aimed at employees’ future prospects and career development.

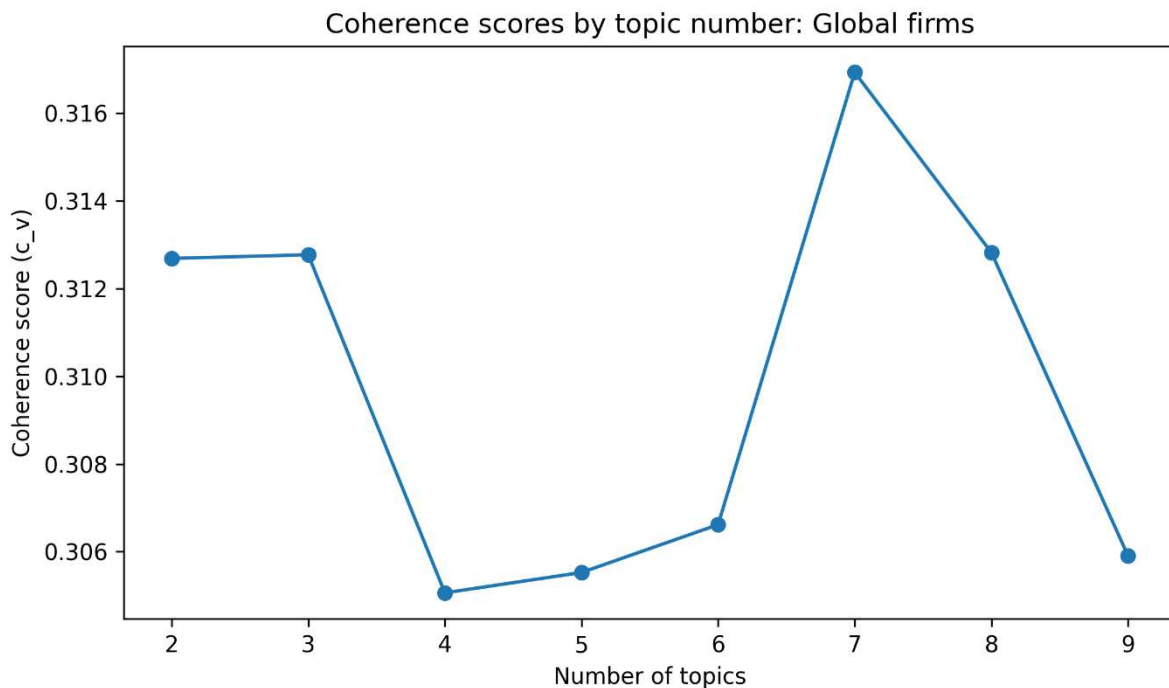
For the global companies’ CEO messages, coherence scores were relatively stable across candidate topic numbers (Figure 8). Although the highest coherence score appeared at seven topics ( $n = 7$ ,  $c_v = 0.31693$ ), the study selected a parsimonious three-topic model ( $n = 3$ ,  $c_v = 0.31277$ ) because it produced coherent, non-overlapping themes and facilitated a direct comparison with the Korean results. Therefore, the optimal number of topics was set at three.

For further analysis of the individual topics in global companies’ CEO messages, high-frequency keywords common to all topics (e.g., business, company, community, customer, world, sustainability) were excluded, and distinctive keywords for each topic were extracted. To enhance interpretability, the reflection rates (topic–word probabilities) of the distinctive keywords were summarized and visualized using bar charts (Figures 9–11).

Three topics emerged in the CEO messages of global companies: (1) goal-oriented commitment, (2) technological growth, and (3) energy value realization.

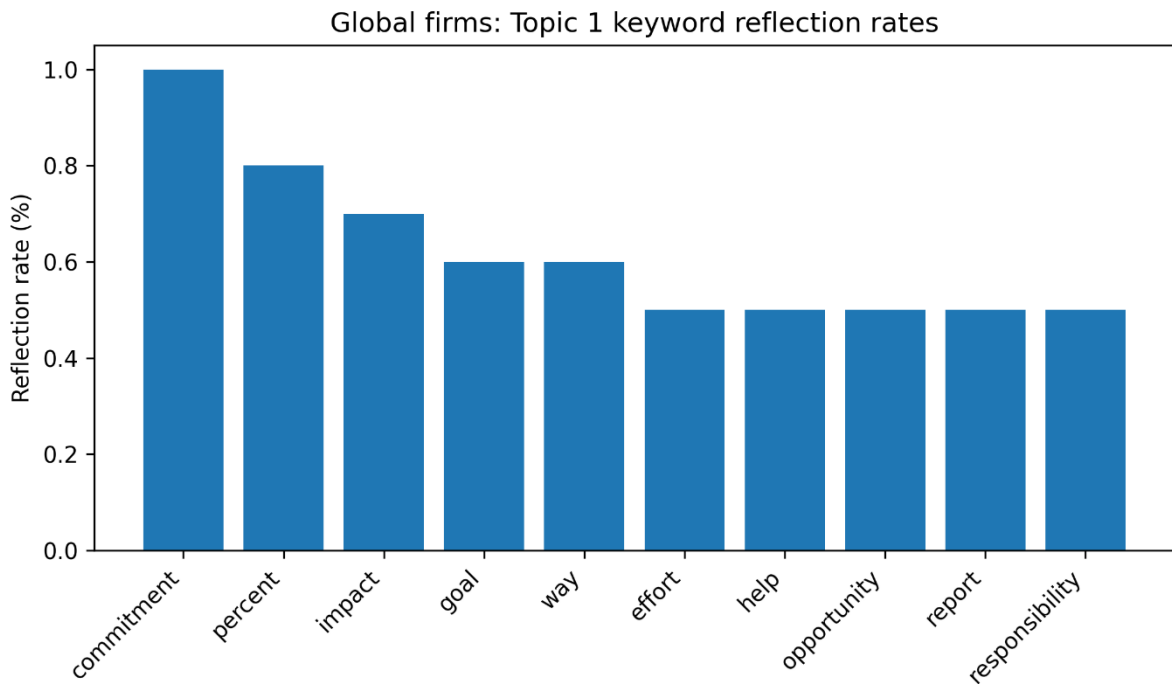
Topic 1 is described as “goal-oriented commitment.” Unique keywords such as commitment, impact, goal, effort, and way capture the essence of this topic (see Figure 9). The focus is on opportunities, dedicated efforts, and responsibilities related to achieving strategic goals.

Topic 2 is identified as “technological growth.” Unique keywords such as energy, technology, continue, growth, and progress clearly represent this



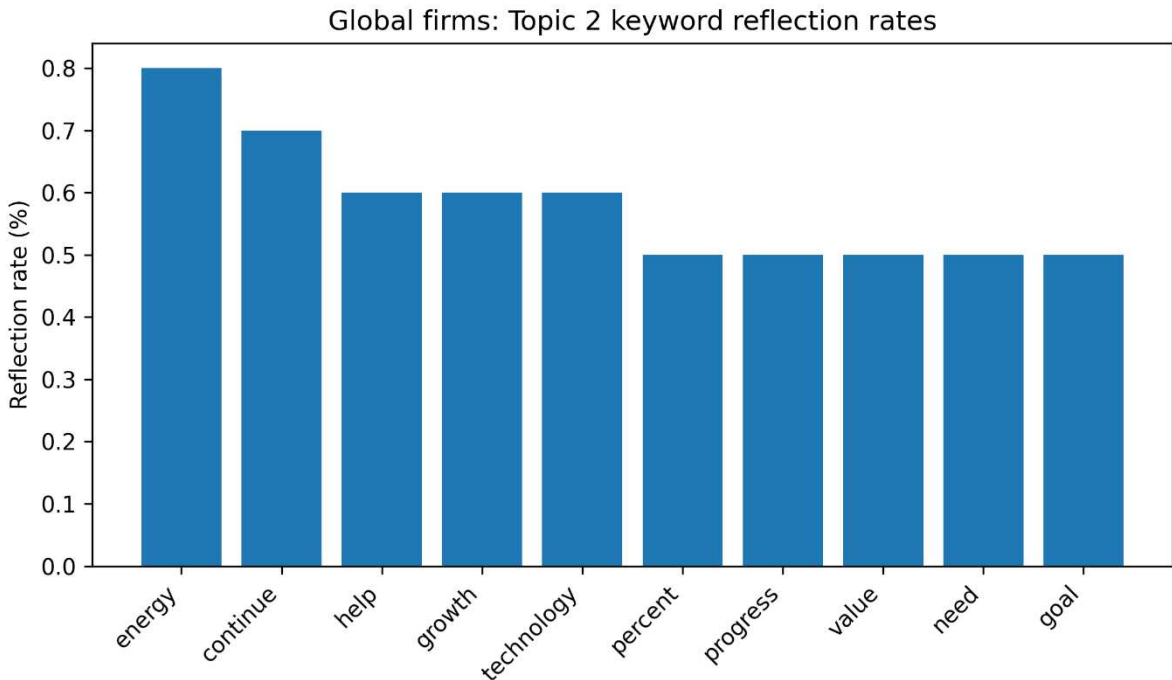
Note: Coherence scores ( $c_v$ ) are computed for candidate topic numbers. The selected topic number balances coherence, parsimony, and interpretability.

Figure 8. Coherence scores by topic number: Global firms



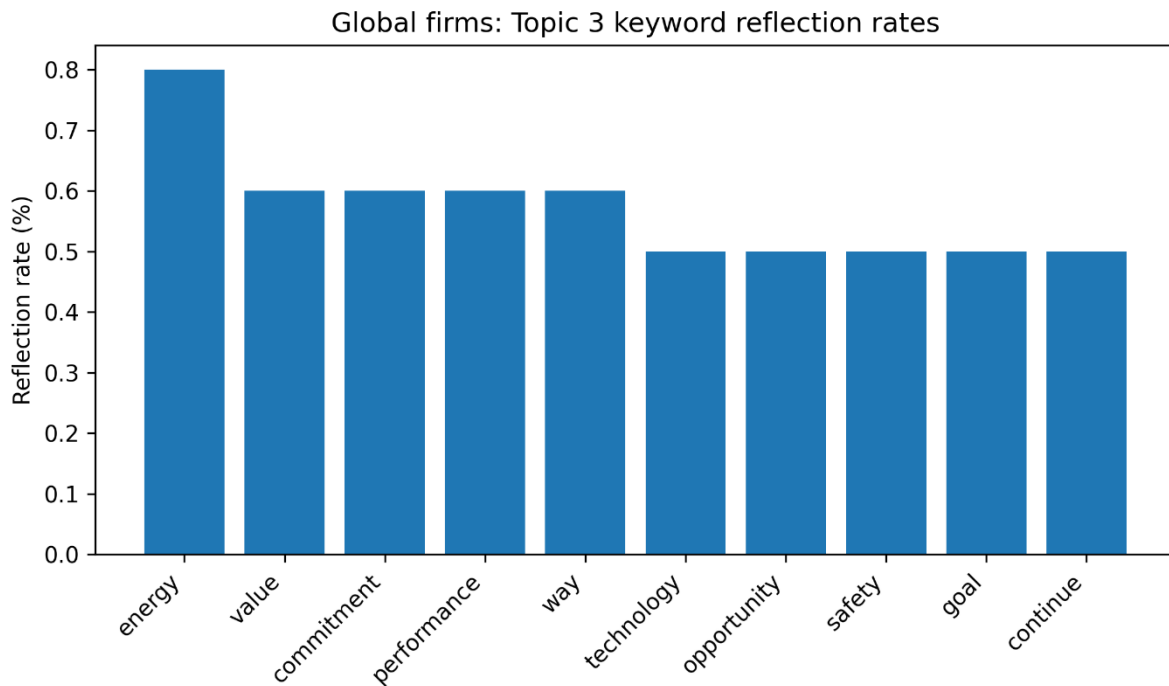
Note: Bars show the reflection rates (%) of the top distinctive keywords for Topic 1 after removing common high-frequency terms shared across topics.

Figure 9. Keyword reflection rates of topic 1 in global firms



Note: Bars show the reflection rates (%) of the top distinctive keywords for Topic 2 after removing common high-frequency terms shared across topics.

Figure 10. Keyword reflection rates of topic 2 in global firms



Note: Bars show the reflection rates (%) of the top distinctive keywords for Topic 3 after removing common high-frequency terms shared across topics.

**Figure 11.** Keyword reflection rates of topic 3 in global firms

topic (see Figure 10). The content primarily emphasizes efforts and activities aimed at technological development within the energy industry.

Topic 3 is characterized as “energy value realization.” Unique keywords such as energy, value, commitment, performance, and opportunity strongly reflect this topic (see Figure 11). The content largely concerns opportunities and commitments directed toward realizing the value of energy technology.

To summarize the topics of Korean and global companies, Korean topics reflected institutional aspiration, while global topics demonstrated a stronger execution focus.

## 4. DISCUSSION

### 4.1. Interpretation of results

This study examined differences in CEO messages across sustainability reports of Korean and global firms using three text-mining techniques. Across 666 reports, the findings reveal systematic

differences in how sustainable communication is framed.

First, semantic network analysis revealed that while the top 20 keywords in both Korean and global companies’ CEO messages were highly interconnected, their centrality rankings diverged. Korean companies emphasized terms such as management, growth, future, and market, reflecting a future-oriented and performance-driven framing. Global companies, in contrast, prioritized community, customer, employee, and commitment, indicating an emphasis on relational engagement and responsibility toward stakeholders. These differences suggest that Korean and global firms construct the meaning of sustainable management through distinct thematic emphases. More specifically, Korean companies highlight business growth and market value, whereas global companies focus on service and commitment to stakeholder communities. This finding aligns with Morsing and Schultz’s (2006) model of sustainability communication, where firms balance stakeholder information, response, and involvement strategies depending on institutional and cultural contexts.

Next, word embedding analysis showed distinct associated words for the core concepts of sustainability and stakeholder. In Korean companies, these terms were closely linked to references such as Dow Jones and the DJSI, as well as expressions of gratitude (thank, appreciate), reflecting an emphasis on external recognition and awards. In contrast, global companies associated these concepts with citizenship, responsibility, CSR, and serve, signaling a deeper orientation toward ethical responsibility and social contribution. Such differences reinforce the argument by Pasamar et al. (2023) that institutional pressures (coercive, mimetic, and normative) affect the tone, framing, and substantive content of sustainability disclosures.

Finally, the topic modeling analysis identified four dominant themes in Korean companies' CEO messages: "future product industries," "global community support," "energy projects," and "employee development." In global companies, three themes emerged: "goal-oriented commitment," "technological growth," and "energy value realization." Both groups addressed innovation, community, and energy-related initiatives, thus displaying institutional isomorphism by superficially advocating for sustainable management. However, while Korean firms framed these topics in ways that foreground managerial achievement and external evaluation, global firms highlighted execution, measurable achievements, and long-term value creation. These results echo findings by Arvidsson and Sabelfeld (2023) that CEO letters adapt framing to institutional expectations, with industry- and country-specific differences shaping legitimacy strategies.

Synthesizing across the three analyses, Korean companies still use CEO messages in sustainability reports mainly as ceremonial prefaces, whereas global companies treat them as strategic communication tools. This pattern can be interpreted through legitimacy theory (Suchman, 1995) and institutional theory (DiMaggio & Powell, 1983). Korean CEO messages stress economic and social contributions, external recognition, and gratitude, reflecting pragmatic legitimacy and symbolic legitimacy strategies (Mahoney et al., 2013). Global firms emphasize service, responsibility, citizenship, and concrete achievements, aiming at moral or cognitive legitimacy through substantive

legitimation. This divergence reflects different institutionalization levels, consistent with Aguilera et al. (2021) on symbolic conformity in emerging sustainability contexts.

## 4.2. Implications, limitations, and directions for future research

The findings contribute to research on sustainability communication by demonstrating how institutional context shapes the micro-level structure of CEO discourse. By applying text mining analysis, this study empirically supports the proposition that communication strategies are shaped by a firm's stage of institutionalization and type of legitimacy pressure (Suchman, 1995; Morsing & Schultz, 2006; Pasamar et al., 2023). This suggests a need to refine theoretical models of sustainability communication to account for context-specific legitimacy pathways and the symbolic-substantive continuum in CEO discourse.

The results also have practical implications for managers and communication professionals. Especially for Korean companies that want to be competitive by meeting global standards of sustainable management, moving from symbolic alignment toward substantive action and performance-based disclosure could enhance long-term legitimacy and stakeholder trust. Communication strategists should also recognize that core concepts like sustainability and stakeholder carry different connotations across contexts, requiring nuanced framing to avoid misalignment or perceived greenwashing (Eliwa et al., 2021).

Several limitations of this study point to avenues for future research. Data analyzed in this study include only English-language sustainability reports, which may bias linguistic and thematic patterns, particularly for Korean companies whose original reports may differ in tone or emphasis. The dataset was limited to large firms adhering to GRI standards, potentially excluding diverse communication strategies from smaller or non-GRI reporting companies. Future research could address these limitations by incorporating multilingual corpora and expanding to non-GRI reporting firms. Additionally, linking message content to actual ESG performance outcomes could help assess the effectiveness of symbolic versus substantive strategies in shaping legitimacy and stakeholder trust.

## CONCLUSION

This study aimed to uncover institutional differences in CEO messages within sustainability reports of Korean and global firms by applying text mining techniques to 666 reports. The results show that Korean firms frame sustainable management primarily in terms of management, growth, future, and market performance, whereas global firms emphasize community, employees, customers, and commitment to stakeholders. The concepts of sustainability and stakeholder are closely linked to external awards and expressions of gratitude in Korean CEO messages, but to citizenship, responsibility, and service in global ones. Topic modeling further reveals four aspiration- and evaluation-oriented themes in Korean messages compared with three execution- and long-term value-oriented themes in global messages.

These findings indicate that Korean and global firms pursue sustainability legitimacy through different communicative routes. Korean CEO messages place a greater emphasis on demonstrating conformity and securing pragmatic approval, whereas global CEO messages reflect a more embedded and substantive orientation toward sustainability.

In conclusion, this study provides empirical evidence that differences in institutionalization and legitimacy pressures not only manifest in but actively shape the structure, framing, and content of CEO messages in sustainability reports. This reinforces the view that sustainability communication is both a mirror of organizational values and a strategic adaptation to institutional contexts.

## AUTHOR CONTRIBUTIONS

Conceptualization: Tae-Il Yoon.

Data curation: Tae-Il Yoon.

Formal analysis: Tae-Il Yoon.

Funding acquisition: Tae-Il Yoon.

Investigation: Tae-Il Yoon.

Methodology: Tae-Il Yoon.

Project administration: Tae-Il Yoon.

Resources: Tae-Il Yoon.

Software: Tae-Il Yoon.

Supervision: Tae-Il Yoon.

Validation: Tae-Il Yoon.

Visualization: Tae-Il Yoon.

Writing – original draft: Tae-Il Yoon.

Writing – review & editing: Tae-Il Yoon.

## ACKNOWLEDGMENT

This research was supported by Hallym University Research Fund, 2025 (HRF-202503-001).

## REFERENCES

1. Abrahamson, E., & Park, C. (1994). Concealment of negative organizational outcomes: An agency theory perspective. *Academy of Management Journal*, 37(5), 1302-1334. Retrieved from <https://journals.aom.org/doi/abs/10.5465/256674>
2. Aguilera, R. V., Aragón-Correa, J. A., Marano, V., & Tashman, P. A. (2021). The corporate governance of environmental sustainability: A review and proposal for more integrated research. *Journal of Management*, 47(6), 1468-1497. <https://doi.org/10.1177/0149206321991212>
3. Al Amosh, H. (2025). The power of words: How CEO speech length and frequency shape ESG transparency. *In-*

- ternational Journal of Business Communication*. <https://doi.org/10.1177/23294884251361462>
4. Arvidsson, S., & Sabelfeld, S. (2023). Adaptive framing of sustainability in CEO letters. *Accounting, Auditing & Accountability Journal*, 36(9), 161-199. <https://doi.org/10.1108/AAAJ-11-2019-4274>
  5. Bachmann, P., & Ingenhoff, D. (2016). Legitimacy through CSR disclosures? The advantage outweighs the disadvantages. *Public Relations Review*, 42(3), 386-394. <https://doi.org/10.1016/j.pubrev.2016.02.008>
  6. Barkemeyer, R., Comyns, B., Figge, F., & Napolitano, G. (2014). CEO statements in sustainability reports: Substantive information or background noise? *Accounting Forum*, 38(4), 241-257. <https://doi.org/10.1016/j.accfor.2014.07.002>
  7. Bengfort, B., Bilbro, R., Danielsen, M., & Ojeda, T. (2018). *Applied text analysis with Python: Enabling language-aware data products with machine learning*. O'Reilly Media.
  8. Bengtson, E., & Mossberg, O. (2023). Institutional theory and green legitimation. In *The virtues of green marketing: A constructive talk on corporate rhetoric* (pp. 19-29). Cham: Palgrave Macmillan. [https://doi.org/10.1007/978-3-031-32979-1\\_3](https://doi.org/10.1007/978-3-031-32979-1_3)
  9. Berkin, A., Aerts, W., Roszkowska-Menkes, M., & Van Caneghem, T. (2025). CEO talk under scrutiny: Visibility and rhetorical impression management in sustainability narratives. *Accounting, Auditing & Accountability Journal*, 38(8), 1949-1978. <https://doi.org/10.1108/AAAJ-08-2024-7297>
  10. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022. Retrieved from <https://dl.acm.org/doi/10.5555/944919.944937>
  11. Boudt, K., & Thewissen, J. (2019). Jockeying for position in CEO letters: Impression management and sentiment analytics. *Financial Management*, 48(1), 77-115. <https://doi.org/10.1111/fima.12219>
  12. Braga, L. D., Tardin, M. G., Perin, M. G., & Boaventura, P. (2024). Sustainability communication in marketing: A literature review. *RAUSP Management Journal*, 59(3), 293-311. Retrieved from <https://ideas.repec.org/a/eme/rauspp/rausp-10-2023-0205.html>
  13. Branco, M. C., & Rodrigues, L. L. (2008). Factors influencing social responsibility disclosure by Portuguese companies. *Journal of Business Ethics*, 83(4), 685-701. <https://doi.org/10.1007/s10551-007-9658-z>
  14. Cho, C. H., Freedman, M., & Patten, D. M. (2012). Corporate disclosure of environmental capital expenditures: A test of alternative theories. *Accounting, Auditing & Accountability Journal*, 25(3), 486-507. <https://doi.org/10.1108/09513571211209617>
  15. D'Aveni, R. A., & MacMillan, I. C. (1990). Crisis and the content of managerial communications: A study of the focus of attention of top managers in surviving and failing firms. *Administrative Science Quarterly*, 35(4), 634-657. <https://doi.org/10.2307/2393512>
  16. Davis, A. K., & Tama-Sweet, I. (2012). Managers' use of language across alternative disclosure outlets: Earnings press releases versus MD&A. *Contemporary Accounting Research*, 29(3), 804-837. <https://doi.org/10.1111/j.1911-3846.2011.01125.x>
  17. DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147-160. <https://doi.org/10.2307/2095101>
  18. Eliwa, Y., Aboud, A., & Saleh, A. (2021). ESG practices and the cost of debt: Evidence from EU countries. *Critical Perspectives on Accounting*, 79, Article 102097. <https://doi.org/10.1016/j.cpa.2019.102097>
  19. Fernando, S., & Lawrence, S. (2014). A theoretical framework for CSR practices: Integrating legitimacy theory, stakeholder theory, and institutional theory. *Journal of Theoretical Accounting Research*, 10(1), 149-178. Retrieved from <https://www.scirp.org/reference/referencespapers?referenceid=3715238>
  20. Godemann, J., & Michelsen, G. (2011). Sustainability communication – An introduction. In J. Godemann & G. Michelsen (Eds.), *Sustainability communication: Interdisciplinary perspectives and theoretical foundations* (pp. 3-11). Springer. [https://doi.org/10.1007/978-94-007-1697-1\\_1](https://doi.org/10.1007/978-94-007-1697-1_1)
  21. Herbert, S., & Graham, M. (2021). Applying legitimacy theory to understand sustainability reporting behaviour within South African integrated reports. *South African Journal of Accounting Research*, 36(2), 147-169. <https://doi.org/10.1080/10291954.2021.1918481>
  22. Hong, Y., & Cho, K. (2024). Differences in CEO communication strategies between high- and low-performing firms in the global auto parts industry. *Sustainability*, 16(8), Article 3100. <https://doi.org/10.3390/su16083100>
  23. Ignatow, G., & Mihalcea, R. (2016). *Text mining: A guidebook for the social sciences*. Sage Publications.
  24. Jegadeesh, N., & Wu, D. (2013). Word power: A new approach for content analysis. *Journal of Financial Economics*, 110(3), 712-729. <https://doi.org/10.1016/j.jfineco.2013.08.018>
  25. Kang, M. (2022). Sustainable management of major online platforms in the Republic of Korea. *Commerce Education*, 36(1), 75-125. (In Korean) Retrieved from <https://doi.org/10.34274/krabe.2022.36.1.004>
  26. Lai, A., Melloni, G., & Stacchez-zini, R. (2016). Corporate sustainable development: Is 'integrated reporting' a legitimation strategy? *Business Strategy and the Environment*, 25(3), 165-177. <https://doi.org/10.1002/bse.1863>
  27. Legendre, S., & Coderre, F. (2013). Determinants of GRI G3 application levels: The case of the Fortune Global 500. *Corporate Social Responsibility and Environmental Management*, 20(3), 182-192. <https://doi.org/10.1002/csr.1285>

28. Mahoney, L. S., Thorne, L., Cecil, L., & LaGore, W. (2013). A research note on standalone corporate social responsibility reports: Signaling or greenwashing? *Critical Perspectives on Accounting*, 24(4-5), 350-359. <https://doi.org/10.1016/j.cpa.2012.09.008>
29. Marschlich, S. (2022). Conceptual framework: Sociological neo-institutionalism, legitimacy, and public relations. In *Corporate diplomacy: How multinational corporations gain organizational legitimacy* (pp. 13-50). Wiesbaden: Springer VS. [https://doi.org/10.1007/978-3-658-36818-0\\_2](https://doi.org/10.1007/978-3-658-36818-0_2)
30. Morsing, M., & Schultz, M. (2006). Corporate social responsibility communication: Stakeholder information, response and involvement strategies. *Business Ethics: A European Review*, 15(4), 323-338. Retrieved from <https://doi.org/10.1111/j.1467-8608.2006.00460.x>
31. Na, H. J., Lee, K. C., Choi, S. U., & Kim, S. T. (2020). Exploring CEO messages in sustainability management reports: Applying sentiment mining and Sustainability Balanced Scorecard methods. *Sustainability*, 12(2), Article 590. <https://doi.org/10.3390/su12020590>
32. Pasamar, S., Bornay-Barrachina, M., & Morales-Sánchez, R. (2023). Institutional pressures for sustainability: A triple bottom line approach. *European Journal of Management and Business Economics*, 34(4), 460-484. <https://doi.org/10.1108/EJMBE-07-2022-0241>
33. Signitzer, B., & Prexl, A. (2007). Corporate sustainability communications: Aspects of theory and professionalization. *Journal of Public Relations Research*, 20(1), 1-19. <https://doi.org/10.1080/10627260701726996>
34. Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, 20(3), 571-610. <https://doi.org/10.5465/amr.1995.9508080331>
35. Yadav, M. S., Prabhu, J. C., & Chandy, R. K. (2007). Managing the future: CEO attention and innovation outcomes. *Journal of Marketing*, 71(4), 84-101. <https://doi.org/10.1509/jmkg.71.4.084>
36. Yoon, T. (2022). How does a company secure legitimacy through ESG?: A qualitative interview exploring perceptions of ESG practitioners. *Communication Theories*, 18(3), 135-177. (In Korean). <https://doi.org/10.20879/ct.2022.18.3.135>
37. Yoon, T., & Byun, J. (2023). A sentiment analysis of CEO message in sustainability report. *Journal of Public Relations*, 27(1), 153-181. (In Korean). <https://doi.org/10.15814/jpr.2023.27.1.153>

## APPENDIX A

**Table A1.** GRI industry sector and coding number

<b>Coding number</b>	<b>Sector</b>	<b>Coding number</b>	<b>Sector</b>
1	Agriculture	18	Health care product
2	Automotive	19	Healthcare service
3	Aviation	20	Household and personal products
4	Chemical	21	Logistics
5	Commercial services	22	Media
6	Other	23	Metal product
7	Computer	24	Mining
8	Conglomerate	25	Real estate
9	Construction	26	Retailers
10	Construction material	27	Technology hardware
11	Consumer durable	28	Telecommunications
12	Energy	29	Textiles and apparel
13	Energy utilities	30	Tourism/leisure
14	Equipment	31	Toys
15	Financial service	32	Non-profit/services
16	Food and beverage product	33	Public agency
17	Forest and paper product		