한국경영과학회지
제50권 제3호
2025년 8월
https://doi.org/10.7737/JKORMS.2025.50.3.051

AI 기반 소셜 미디어 감성 분석을 활용한 ESG 사회지표 개발

장 혁 수†
SK(주) AX 애커튼파트너스 ESG/SHE 본부

Toward Real-Time ESG Evaluation: A Public Opinion-Based Approach Using Social Media Data

Hyuk Soo Jang
SK AX Inc. Ackerton Partners ESG/SHE Division

■ Abstract ■

This study proposes a novel Social Index for evaluating ESG (Environmental, Social, and Governance) performance by leveraging deep learning-based text analysis of unstructured social media data. Traditional ESG assessment methods primarily rely on structured data sources, such as corporate disclosures or surveys, which fail to capture real-time public sentiment. To address this limitation, the proposed methodology utilizes the GloVe (Global Vectors for Word Representation) model to analyze social media texts in relation to the Sustainable Development Goals (SDGs), thereby quantifying public perception of corporate ESG activities in real time.

The findings reveal three key contributions. First, the proposed Social Index significantly narrows the gap between reported ESG performance and public sentiment by incorporating dynamic, real-time data. Second, distinct patterns in public perception across ESG and SDG dimensions suggest the need for companies to adopt customized ESG strategies sensitive to regional and sectoral contexts. Third, this index offers a transparent and objective tool for external stakeholders to evaluate ESG performance, enhancing corporate accountability and alignment with societal expectations.

Overall, the study demonstrates that integrating public opinion into ESG assessment provides a more responsive and socially attuned evaluation framework, supporting the development of sustainable and socially responsible corporate strategies.

Keywords : ESG, Social Indicators, Public Opinion, Deep Learning, Text Mining

논문접수일 : 2025년 06월 24일 논문게재확정일 : 2025년 07월 16일
논문수정일 : 2025년 07월 13일
† 교신저자, hsjang9317@gmail.com

1. Introduction

Global challenges such as pandemics, climate change, and natural disasters have underscored the urgency of building a sustainable future. In response, Environmental, Social, and Governance (ESG) strategies have gained prominence worldwide, with governments and corporations strengthening ESG evaluation standards and increasing the publication of sustainability reports. Notably, the International Sustainability Standards Board (ISSB) introduced global sustainability disclosure standards in 2023, to be applied from 2024, establishing a unified framework for climate-related reporting. The growing importance of ESG is reflected in projections by Bloomberg Intelligence, which estimate ESG assets will reach USD 50 trillion by 2025.

High-profile cases such as Nike's child labor scandal in the 1990s demonstrate the reputational and financial risks companies face when failing to meet ESG expectations. In response to public backlash, Nike implemented robust corporate social responsibility initiatives, including environmental and diversity programs [18]. The global spread of ESG investing was further accelerated by the launch of the UN Principles for Responsible Investment in 2006, which saw signatories rise from 63 to 2,372 by 2019 [25]. In Europe, regulatory efforts such as the EU Taxonomy and the proposed carbon border tax reflect a trend toward mandatory ESG disclosures and climate action [26].

As ESG considerations become essential for corporate survival, companies have increasingly adopted ESG metrics to evaluate their performance. In South Korea, indices such as the Korea Corporate Governance Service (KCGS) and the Korea Economic Justice Institute Index (KEJI) are widely used [20]. However, discrepancies in ESG rating methodologies across domestic and international agencies raise concerns over standardization and transparency [23]. The K-ESG Guidelines emphasize labor-related dimensions of ESG—working conditions, human rights, and diversity—and highlight stakeholder engagement as critical for fostering sustainable labor ecosystems [35]. Internal job satisfaction has also been positively linked to financial performance, such as return on assets [21], supporting the view that ESG must address the interests of diverse stakeholders, including employees, consumers, and local communities—not just investors.

In South Korea, ESG management has evolved into a strategic imperative for both large corporations and SMEs, driven in part by public procurement and supplier selection processes [39]. Although over 600 ESG rating systems exist globally, their inconsistent methodologies have led companies to seek evaluations from multiple agencies, underscoring the lack of a unified standard [10, 28]. This inconsistency can result in divergent assessments of the same corporate activities and highlights the need for ESG evaluation models tailored to national and industrial contexts. Without robust ESG systems, companies risk consumer boycotts, exclusion from certifications, and limited access to public projects.

Despite its importance, ESG evaluation research remains underdeveloped, with most models primarily investor-focused. Scholars have noted that ESG scores influence stakeholder perceptions of corporate risks and opportunities [1, 3]. However, the growing role of public opinion

and the increasing demand for customized, context-specific ESG models call for innovative evaluation approaches. In this context, this study proposes the development of an ESG Social Index that integrates public opinion into ESG performance evaluation using social media text data and natural language processing (NLP). By incorporating the perspectives of key stakeholders—especially the public—this index aims to enhance the responsiveness and reliability of ESG evaluations and address the limitations of traditional, structured-data approaches.

This research seeks to address two key questions: (1) How can real-time public sentiment improve the responsiveness of ESG performance assessments? (2) How can the ESG Social Index reflect regional and industrial differences and reduce the gap between corporate ESG efforts and public perception?

The study, while contributing to ESG assessment methodology, acknowledges several limitations: potential bias in social media data, accuracy issues in machine translation, and the constraints of using the GloVe model and a limited data timeframe. Additionally, as the analysis focuses on a single corporate case, broader applicability may be limited.

This thesis is organized as follows: Chapter 2 reviews prior studies on ESG evaluation and public sentiment analysis. Chapter 3 outlines the methodology, including data collection and NLP techniques. Chapter 4 presents analysis results based on SDG keyword similarity using the GloVe model. Chapter 5 concludes with key findings, implications, and suggestions for future research, emphasizing the potential of the ESG Social Index to improve stakeholder-oriented ESG assessment systems.

2. Literature Review

2.1 ESG, SDGs, and the Theoretical Foundations of Sustainable Management

ESG (Environmental, Social, and Governance) is a framework that promotes long-term, ethical corporate behavior, originally advanced through the United Nations Global Compact [34]. ESG emphasizes responsibility in three key domains: environmental sustainability (e.g., emissions reduction, resource recycling), social impact (e.g., labor rights, diversity), and governance (e.g., transparency, anti-corruption) [11]. These principles have gained traction amid global crises like the COVID-19 pandemic and climate change, highlighting the growing relevance of non-financial corporate performance [6].

Closely linked to ESG is the United Nations' Sustainable Development Goals (SDGs), which offer a comprehensive blueprint to address poverty, inequality, and climate action through 17 goals and 169 targets. ESG dimensions directly correspond to SDG targets—environmental efforts align with goals on climate and ecosystems, social dimensions with goals on equity and education, and governance with goals on justice and institutional integrity [22, 29]. Together, ESG and SDGs present a unified vision of corporate responsibility aligned with global sustainability priorities.

The growing importance of ESG also reflects a shift from shareholder-centric models to shareholder-oriented value creation [37]. Regulatory mandates, increased ESG-based investment strategies, and demand from consumers and employees have accelerated ESG adoption. Corporations now face rising pressure to disclose ESG performance transparently and align with

international standards [7, 12, 36]. Despite increasing standardization efforts by organizations like SASB, ISO, and TCFD, inconsistency remains a major barrier [19], underscoring the need for unified yet adaptable ESG metrics that reflect stakeholder values.

2.2 Machine Learning and AI in ESG Evaluation

Recent research increasingly applies artificial intelligence and machine learning to ESG evaluation. Techniques such as Support Vector Regression, Random Forest, and deep learning models like LSTM and BERT are used to predict financial metrics based on ESG activities [9, 13]. News-based models like ESG2RISK and semi-supervised text mining approaches have demonstrated rapid responsiveness to ESG-related events [8, 17]. Additionally, social media-based models provide a dynamic layer of stakeholder sentiment analysis, offering new pathways for ESG monitoring and forecasting [32, 33]. Furthermore, sentiment analysis using local-language financial news—such as the KR-FinBERT-based portfolio optimization model applied to the South Korean stock market—highlights the potential of regionally customized NLP tools in enhancing financial decision-making [27]. These innovations highlight the potential of unstructured data in developing real-time, adaptive ESG metrics.

2.3 Research Gaps and Implications for ESG Social Index Development

Despite advances, ESG evaluation remains largely investor-centric and insufficiently reflects public opinion—a key determinant of corporate reputation and legitimacy [7]. Traditional models rely heavily on corporate disclosures and structured data, failing to account for stakeholder sentiment, particularly from the general public. Moreover, while many studies focus on the financial implications of ESG, they often overlook social value creation and long-term non-financial outcomes [15].

Methodologies that incorporate real-time public opinion into ESG evaluation are still in early stages. Technologies like natural language processing (NLP) can bridge this gap, as demonstrated by Lee and Kim [24], who developed a Social Index using public opinion data linked to the SDGs. In addition, ESG and public perception are both dynamic and time-sensitive, yet existing models often lack the ability to adapt over time [14].

In light of these limitations, this study proposes the ESG Social Index—a model that captures public sentiment using deep learning-based text analysis of social media. By doing so, it seeks to enhance the inclusiveness, responsiveness, and reliability of ESG evaluations, thereby offering a more holistic view of corporate sustainability performance.

3. Research Methodology

3.1 Overview of Analytical Framework and Theoretical Basis

This study proposes a real-time ESG evaluation methodology that incorporates public sentiment derived from social media data, aiming to address the limitations of traditional ESG assessments that rely heavily on structured data such as corporate disclosures or surveys. To achieve this, the GloVe (Global Vectors for Word Representation) embedding model was applied to quantify the semantic

similarity between social media texts and keywords related to the Sustainable Development Goals (SDGs). This approach allows for the development of a dynamic, socially attuned Social Index, which is distinct from conventional evaluation metrics.

The theoretical foundation of this approach lies in the structural alignment between ESG (Environmental, Social, and Governance) and the SDGs. ESG focuses on corporate-level sustainability performance, while the SDGs outline global sustainability objectives. Recent research [2, 4] has shown that ESG practices can meaningfully contribute to achieving SDG targets. Each component of ESG aligns closely with relevant SDGs: the Environmental (E) dimension is related to climate action and resource management (SDGs 12-15); the Social (S) dimension includes labor conditions and community engagement (SDGs 1-11); and the Governance (G) dimension addresses transparency and responsible management (SDGs 16-17).

To operationalize this connection, this study adopts a keyword-based text analysis methodology. Specifically, it employs the keyword set developed by Lee and Kim [2], which is based on the "Compiled Keywords for SDG Mapping" by SDSN Australia/Pacific and ACTS, along with additional terms from the SDG Compass. This comprehensive set categorizes keywords by SDG goal and thematic relevance and distinguishes between positive and negative connotations, allowing for sentiment-weighted analysis.

The Global Reporting Initiative [16] has similarly proposed a framework that integrates SDG-related keywords into ESG reports for assessing corporate sustainability activities. Other studies, such as those by Rosati and Faria [31], S&P Global [38], and Van Zanten and Van Tulder [40], have applied SDG keywords to sustainability reports to assess ESG performance. Building on these precedents, this study applies such keywords to public social media data to enable the real-time, quantitative analysis of stakeholder sentiment toward ESG efforts.

This section has outlined the conceptual rationale and methodological foundation of the study. The following sections provide a step-by-step explanation of the data collection, preprocessing, embedding, and similarity analysis processes that operationalize this framework.

3.2 Data Collection

Two main types of data were required for the analysis in this study. The first was the sustainability reports officially issued by companies, and the second was social data, such as comments on social media posts. These datasets play a crucial role in evaluating corporate ESG (Environmental, Social, and Governance) performance and analyzing public perceptions.

To select companies for analysis, the following criteria were established. First, it was considered whether the company officially published a sustainability report. Second, it was verified whether there was active public engagement on social media platforms and a sufficient volume of social data. Additionally, companies that received high ratings from major ESG evaluation agencies, such as MSCI ESG Research, S&P Global Ratings, and Sustainalytics, were prioritized. Based on these criteria, Company A in the food and beverage industry and Company B in the distribution industry were chosen as the final analysis subjects. Both Company A and Company B have been con-

sistently publishing reports and actively engaging with the public. Company A has maintained an AAA rating in MSCI evaluations from 2021 to 2023 and received scores of 87 and 94 in the DJSI evaluations for 2022 and 2023, respectively. Furthermore, Company A consistently held an A- rating in the CDP Climate Change assessment from 2021 to 2023. Company B maintained A and AA ratings in MSCI evaluations from 2021 to 2023 and scored 72 and 71 in the DJSI evaluations for 2022 and 2023, respectively. In the CDP Climate Change assessment, Company B received A-, B, and A- ratings from 2021 to 2023, showing excellent performance.

Social media serves as an essential channel for communication between companies and the public, allowing companies to gauge public perceptions and responses through comments left on posts. For this purpose, all comments on posts from each selected company's official social media accounts were collected, from the first post in 2022 to the most recent post as of August 30, 2024. The platforms targeted for comment collection were Instagram, X (formerly Twitter), and Facebook. However, despite changing Facebook's default comment sorting setting from "Most Relevant" to "All Comments," only some comments were publicly accessible. Due to the limited number of comments on Facebook, comments were collected only from Instagram and X.

During the comment collection process, the open-source web browser automation tool Selenium was used in a Python environment to extract comments from each company. Selenium is useful for systematically collecting comment data that is difficult to access manually by automatically loading and interacting with HTML elements on specific web pages. In this study, Selenium was integrated with ChromeDriver to automatically scroll through target pages and repeatedly search for HTML elements containing comments, loading all comments. Each comment's text element was then specified and extracted using CSS selectors. Through this process, a total of 13,989 comments were collected for Company A, and 26,075 comments for Company B.

Meanwhile, for sustainability reports, a total of 100 pages of documents issued in 2022 and 2023 were collected for Company A, and 160 pages for Company B.

3.3 Data Pre-processing

3.3.1 Preprocessing of Social Data

Since the two companies selected for analysis are global corporations, the comments collected from their official social media accounts included texts written in various languages, not just English. However, as the natural language processing tools used in this study primarily target English texts, preprocessing of non-English comments was necessary to ensure high-quality analysis results.

In particular, Company B is a French company, and a significant portion of the collected comments were written in French. Simply removing all non-English comments would result in a substantial loss of data, so to avoid this, non-English comments were translated into English using the DeepL translation tool. DeepL is a cutting-edge AI-based translation service that provides highly accurate and natural translation results, aiding the analysis of non-English text.

The translated texts were then processed to remove unnecessary words. For this purpose, Python's Natural Language Toolkit (nltk) library

was used. Nltk offers a variety of tools for analyzing text data, including a stop words set, which is useful for eliminating meaningless words during analysis. For example, words such as pronouns, particles, conjunctions, and auxiliary verbs, which do not significantly affect the text's meaning, were removed to highlight the core information in the text. Additionally, due to the nature of social media comments, users often mention others or emphasize specific topics using hashtags. For instance, mentions are written in the format

"@+targetID," and hashtags appear as "#+topic." Since these elements do not contribute meaningful information to the text's semantic analysis, unnecessary words starting with "@" or "#" were removed. Moreover, numbers, repeated characters, emojis, and special characters not found on a standard keyboard were also deleted.

The final step in the text preprocessing process was stemming. Stemming is the process of identifying the root or base form of a word, simplifying word forms to unify words with the same meaning. For instance, both "running" and "ran" can be converted to "run" for processing. This was achieved using the SnowballStemmer tool, which is one of the stemming algorithms included in nltk and is an improvement on the Porter stemming algorithm developed by Martin Porter [30].

After completing this preprocessing process, the collected comments were ultimately saved in the form of 13,947 rows for Company A and 26,034 rows for Company B.

3.3.2 Sustainability Report

Since the officially issued sustainability reports were all written in English, the preprocessing process could be carried out without any additional translation. This process was conducted similarly to the preprocessing of social data, as described earlier.

First, unnecessary characters that did not appear in the keyword list were removed from the reports. This involved cleaning the text by deleting special characters, extraneous symbols, and similar elements. Following this, nltk's stop word list was utilized to eliminate words deemed irrelevant for analysis, such as pronouns, conjunctions, and particles.

Additionally, as with the social data, numbers in the reports were removed as they did not provide meaningful insights for the analysis. SnowballStemmer was also used to perform stemming on the words. After completing this preprocessing process, Company A's sustainability report was saved as data in the form of 1,585 rows, and Company B's report was saved in 3,415 rows.

<Table 1> and <Table 2> below include portions of the preprocessed report and comment data for Companies A and B, with company names

〈Table 1〉 Pre-processed Data - Company A

Company A	Comments	Pre-processed
0	You have reached us at The Company w…	you have reach us at the compani wor...
1	Hi, I recently bought one of your coke product …	hi, i recent bought one of your coke product ...
2	Hi. Thanks for reaching out. We'd like to lo...	hi. thank for reach out. we'd like to look ...
…		
1,583	□□□	
1,584	Stop the millions of pounds of plastic garbage...	stop the million of pound of plastic garbag yo...

〈Table 2〉 Pre-processed Data - Company B

Company B	Comments	Pre-processed
0	groceries delivered : half of the st...	groceri deliv : half of the stuff i ...
1	this company supports state. therefore it sup...	this compani support state. therefor it suppo...
2	is your plan to continue being complicit with ...	is your plan to continu be complicit with huma...
…		
3,413	Happy New Year to the whole team. beautiful vi...	happi new year to the whole team. beauti vide...
3,414	happy new year to you too	happi new year to you too

omitted to ensure anonymity by leaving blank spaces.

3.4 Keyword Selection

This study utilized the keyword set presented by Lee and Kim [24]. This set is based on the “Keyword Set for SDG (Sustainable Development Goals) Mapping” provided by SDSN Australia/Pacific and ACTS, with additional reference to the SDG Compass, incorporating positive and negative keywords related to the 17 Sustainable Development Goals (SDGs).

The keywords selected by Lee and Kim [24] are organized around specific fields and key topics for each SDG goal. For example, in SDG 1, which emphasizes poverty alleviation and economic opportunity, keywords such as poverty, inequality, and income redistribution are included. For SDG 2, which addresses ending hunger and ensuring nutrition, food security, agriculture, and nutrition enhancement are used as core keywords. These keywords play a crucial role in setting concrete actions and strategies to achieve these goals in both the public and private sectors. The keywords from SDSN and ACTS provide specific and clear themes for goals such as education, gender equality, climate action, and the protection of aquatic and terrestrial ecosystems, facilitating a more practical approach toward the goals’ intended directions.

The additional keywords from the SDG Compass focus on strengthening the connection between SDGs and a business perspective. By adding detailed keywords that enable companies to create strategic synergies with sustainable development goals, the SDG Compass allows for a systematic evaluation of corporate ESG performance. For instance, in SDG 8 (Decent Work and Economic Growth), fair wages and employee engagement are added as major keywords, specifying ways in which companies can contribute to employment and economic growth. The keywords from the SDG Compass aim to assess practical sustainability performance by reflecting the comprehensive social goals of the SDGs while being tailored to align with corporate activities and their impacts through specialized keywords.

In this study, the process of developing a Social Index for SDG achievement using keywords from SDSN, ACTS, and the SDG Compass distinguishes itself from traditional evaluation indicators based on economic and social statistics. Traditional evaluation methods often rely on national-level statistics, which makes it challenging to reflect actual public opinion or sentiment. However, this study proposes a new evaluation metric based on social media data that reflects the genuine emotions of the public, providing a method for assessing how positively a company

〈Table 3〉 Keyword Set

SDG01	SDG02	SDG03	…	SDG15	SDG16	SDG17
-0.7 africa	0.7 agricultural orientation index	1.0 access to clean water and sanitation	…	0.8 afforestation	-1.0 abuse	0.9 capacity building
0.6 basic services	0.8 agricultural productivity	0.8 affordable medicines	…	0.6 agriculture	0.9 accountability	0.9 civil society partnerships
0.5 class	0.7 agriculture	-0.9 aids	…	0.7 animals	1.0 accountable institutions	0.9 communication technologies
…	…	…	…	…	…	…
0.9 health care services	0.6 sustainable sourcing	0.8 health care services	…	0.6 mountain ecosystems	0.7 physical displacement	0.6 weighted tariff average
0.7 electricity availability	0.7 genetic diversity of animals	0.9 air quality	…	-0.6 natural habit degradation	0.8 economic displacement	0.7 women entrepreneurs
0.9 non-discrimination	-0.6 labor practices in the supply chain	0.9 water quality	…	0.9 freshwater ecosystems	0.6 inclusive decision making	0.7 world trade organization

is perceived in terms of its progress toward achieving the SDGs. Specifically, the keywords from SDSN and ACTS offer a suitable framework for collecting public opinion and sentiment and connecting it to SDG evaluation. By adding business oriented keywords from the SDG Compass, this approach further strengthens the business relevance between ESG and SDG performance.

Keywords were processed by assigning different weights according to their positive or negative connotations. Positive keywords were given weights ranging from +0.5 to +1, while negative keywords were assigned weights between -0.5 and -1. These weights were calculated by averaging independent evaluations from text mining experts, reflecting the significance of each keyword in text analysis. <Table 3> below presents a sample of the keyword set proposed by Lee and Kim [24] used in this study.

3.5 Similarity Analysis

3.5.1 Embedding

In similarity analysis, the first step is embedding each text. Embedding text is a core technique in natural language processing, transforming the meaning of words or sentences into vector form in a multidimensional space. This process enables machine learning models to interpret text data, providing a foundation for understanding the relationships between words or sentences in the text. This study considered two different embedding techniques: GloVe (Global Vectors for Word Representation) and BERT (Bidirectional Encoder Representations from Transformers). The primary differences between these two embedding methods lie in how embeddings are generated and how contextual information is processed.

GloVe (Global Vectors for Word Representa-

tion) is a model that performs embeddings at the word level. GloVe generates fixed embedding vectors based on the frequency with which words co-occur with other words in text data. Generating fixed embedding vectors means that each word is assigned a unique vector without considering context. For instance, the word "밤" (which can mean "night" or "chestnut" in Korean) would have the same embedding vector whether it refers to evening or the fruit. While GloVe can capture basic meanings and similarities, it does not account for differences in contextual meanings.

On the other hand, BERT (Bidirectional Encoder Representations from Transformers) is a sentence-level embedding model that can capture the meaning of words in context. This model dynamically generates word vectors within a sentence. Unlike fixed embedding vectors, dynamic embedding vectors in BERT adjust each word's embedding vector according to the structure and context of the sentence by considering the interrelationships of all words within the sentence. This means that the word "밤" (which can mean "night" or "chestnut" in Korean) can be distinguished depending on whether it refers to evening or the fruit, based on context. This characteristic makes BERT more sensitive to semantic alignment than GloVe, allowing it to measure the similarity between two texts with different expressions in a detailed and flexible manner.

However, when using BERT, the computational cost increases with sentence length, requiring significantly higher computational resources compared to fixed embeddings. Additionally, when analyzing social media data and report data, as in this study, where the topics or contexts of the two texts are similar, there is a high likelihood of similarity beyond differences in expression. In contrast, GloVe provides a 100-dimensional embedding vector and vector representations for approximately 400,000 words, trained on a large corpus such as Wikipedia and Gigaword. GloVe, which is trained on a large-scale corpus, offers high computational efficiency and effectively reflects word similarity, allowing for a precise analysis of relationships between text data and keywords. Considering these points, this study determined that GloVe is a more suitable embedding method, while similarity analysis results using distilBERT, a lightweight version of BERT, are provided separately in Appendix 2.

3.5.2 Similarity Score Calculation

The similarity score calculation was conducted by calculating the cosine similarity between vectors using the GloVe model. Cosine similarity is a method for measuring word similarity based on the angle between two vectors, with scores closer to 1 indicating greater similarity. Specifically, the GloVe model extracts vectors for each word in a vector space, and cosine similarity is then calculated between these vectors. The calculated similarity score is used to evaluate the similarity between words extracted from the text and a pre-selected set of keywords.

In this study, a threshold of 0.6 was set for similarity. Only when the similarity between two words is 0.6 or higher is it considered that there is similarity between the keyword and the text, and this similarity is reflected in the score. Conversely, if the similarity is below 0.6, it is regarded that there is no similarity between the words, and the similarity is not recorded.

For words deemed similar with a similarity score exceeding 0.6, the derived similarity score was multiplied by the keyword's weight. This calcu-

lation was performed in the same manner for all keyword weights, and the sum of these values was recorded as the similarity score. This process was applied to all words in each row, and the total similarity score of each row was recorded as the final score for that row. This score serves

〈Table 4〉 Calculation Results: Company A's Comments

	Text	SDG01	SDG02	SDG03	…	SDG15	SDG16	SDG17
0	you have…	71.92	78.24	287.46	…	28.5	80.96	173.23
1	hi, i recent…	42.03	63.06	203.8	…	18.47	47.73	95.86
2	hi. thank…	44.94	52.76	227.27	…	29.51	68.78	125.09
…	…	…	…	…	…	…	…	…
13,986	dont act…	1.28	1.14	22.17	…	4.09	10.29	5.96
13,987	me and m…	3.38	3.47	7.98	…	0.12	4.24	2.3
13,947	stop the m…	13.55	18.47	61.54	…	6.03	15.5	27.86

〈Table 5〉 Calculation Results: Company A's Reports

	Text	SDG01	SDG02	SDG03	…	SDG15	SDG16	SDG17
0	refresh the…	4.99	2.28	3.56	…	-0.60	2.5	2.46
1	make a dif…	3.29	6.79	1.73	…	-0.57	6.26	5.27
2	busi & sust…	3.61	4.58	5.38	…	1.33	2.87	3.44
…	…	…	…	…	…	…	…	…
1,583	click here…	13.44	12.88	43.42	…	4.22	20.87	32.78
1,584	content wa…	19.51	35.07	59.92	…	11.73	14.29	27.97
1,584	assur state…	41.28	50.03	127.06	…	26.52	48.81	75.63

〈Table 6〉 Calculation Results: Company B's Comments

	Text	SDG01	SDG02	SDG03	…	SDG15	SDG16	SDG17
0	groceri deli…	106.83	134.71	489.96	…	43.26	132.49	231.12
1	this compa…	26.14	18.42	73.7	…	8.11	22.66	51.22
2	is your pla…	28.17	33.14	125.33	…	14.61	18.31	66.15
…	…	…	…	…	…	…	…	…
26071	happi new…	15.35	15.03	43.85	…	5.06	14.42	28.97
26072	happi new…	46.45	45.76	179.69	…	23.16	54.27	100.55
26034	happi new…	8.83	9.79	28.79	…	2.68	11.03	16.86

〈Table 7〉 Calculation Results: Company B's Reports

	Text	SDG01	SDG02	SDG03	…	SDG15	SDG16	SDG17
0	univers reg…	254.04	265.74	1028.15	…	107.21	209.65	592.61
1	in line with…	90.57	91.25	321.12	…	35.15	89.11	208.06
2	lastly, it rev…	49.74	59.5	179.35	…	17.88	47.51	100.53
…	…	…	…	…	…	…	…	…
3379	they ensur…	70.9	78.83	306.99	…	28.11	91.84	167.83
3380	lastly, they…	20.17	15.87	96.74	…	10.73	35.25	56.44
3414	not engag…	235.21	271.96	1000.87	…	90.97	281.78	559.05

as an indicator of how closely related each SDG category (SDG01~SDG17) is to the text. <Table 4>~<Table 6>, and <Table 7> below contain the similarity score results for Company A and Company B's comment data and report data.

3.5.3 Scaling

The Environmental (E) category is associated with four SDG goals, specifically SDGs 12 through 15. The Social (S) category is linked to eleven SDG goals, from SDGs 1 through 11, while the Governance (G) category is associated with two SDG goals, SDGs 16 and 17. Based on this, to evaluate ESG performance, the similarity scores of all comments for each SDG category were averaged and compared.

〈Table 8〉 Results of Similarity Calculations Statistics for Company A and Company B

A's Report	SDG01	SDG02	SDG03	…	SDG15	SDG16	SDG17
Max	1376.63	1505.49	5472.93	…	549.66	1356.86	329.71
75%	63.12	69	231.81	…	26.7	64.44	145.94
50%	42.48	45.01	154.9	…	17.89	44.1	97.98
25%	25.98	28.43	93.39	…	11.13	27.85	59.14
Min	0	-0.39	0	…	-0.92	0	0
Mean	50.94	55.17	185.73	…	21.92	51.97	117.89
Std.	53.2	59.1	207	…	24.14	52.83	128.82

A's Comment	SDG01	SDG02	SDG03	…	SDG15	SDG16	SDG17
Max	223.71	278.56	1143.13	…	125.59	333.41	518.06
75%	18.36	22.51	80.58	…	9.27	25.43	40.46
50%	5.45	7.05	21.67	…	2.45	8.94	10.23
25%	0.39	0	0	…	0	0.43	0
Min	-1.35	-1.82	-2.65	…	-1.78	-2.4	0
Mean	13.57	16.59	60.66	…	8.62	18.75	30.86
Std.	19.96	24.25	92.7	…	10.43	26.12	47.68

B's Report	SDG01	SDG02	SDG03	…	SDG15	SDG16	SDG17
Max	762.09	827.48	2986.51	…	286.7	655.05	1706.02
75%	60.61	68.66	228.35	…	24.7	63.14	142.24
50%	38.26	42.74	143.57	…	15.5	40.19	88.93
25%	18.66	20.84	73.22	…	7.73	20.92	44
Min	0	-0.89	-0.92	…	-1.78	-1.61	0
Mean	46.72	52.56	176.17	…	19.14	47.96	108.94
Std.	48.82	56.24	187.46	…	20.57	46.84	114.34

B'sComment	SDG01	SDG02	SDG03	…	SDG15	SDG16	SDG17
Max	483.82	613.92	2260.83	…	260.38	635.7	1128.12
75%	22.81	27.94	101.63	…	11.12	32.88	50.09
50%	7.11	8.44	27	…	2.83	11.63	13.15
25%	0.41	0.02	0.59	…	0	1.16	0
Min	-1.56	-3.94	-1.65	…	-1.87	-5.94	0
Mean	19.61	24.15	89.75	…	9.77	28.54	44.12
Std.	34.14	42.92	162.75	…	18.13	47.61	80.74

However, as shown in <Table 8>, which presents the statistical values of the similarity calculation results, the average values are significantly lower than the maximum values, and the standard deviation and quartiles reveal a skewed distribution. This skewness makes it challenging to compare the averages across different distributions.

To address this issue, log scaling was applied. Log scaling is useful for compressing the range of data values and reducing abnormally high skewness, resulting in a more balanced data distribution. This approach enhances the distinguishability of data clustered at lower values, enabling more effective analysis.

Furthermore, to effectively compare the average similarity scores for each SDG category between social data and sustainability reports, it was necessary to align the scales of the two distributions. For this purpose, min-max scaling was applied after log scaling.

Min-max scaling is a technique that transforms data values into a specific range, typically between 0 and 1 or between 0 and 10. Min-max scaling normalizes each data value based on the minimum and maximum values, converting it to a consistent scale. This scaling facilitates direct comparison between data sets with different ranges and aids in achieving more consistent results in analysis.

In this study, both distributions were scaled to a range between 0 and 10, enabling a comparison of similarity scores between social data and sustainability reports on the same basis. This allowed for a clearer understanding of the relative differences and relationships between the two data sets, making the comparison and interpretation of final analysis results easier.

4. Data Analysis and Findings

4.1 Introduction

This chapter presents a detailed analysis of public perception of ESG (Environmental, Social, and Governance) based on collected social media data and discusses the implications of the resulting Social Index. This chapter is the core of the empirical analysis, providing specific answers to the research questions and empirically examining how the proposed deep learning-based text analysis methodology effectively reflects public perception of ESG.

In this study, the GloVe embedding model was used to analyze the similarity between public social media opinions and SDG (Sustainable Development Goals) keywords. Through this analysis, a new Social Index was derived, reflecting public perception of each SDG. Unlike traditional structured data-based ESG evaluations, this index distinguishes itself by reflecting real-time public opinion through unstructured social media text.

Additionally, this chapter explores differences in public perception of ESG. This provides valuable insights into how public interest varies across ESG issues. To verify the performance and validity of the proposed methodology, statistical verification was conducted, ensuring the reliability and consistency of the analysis results.

The ultimate purpose of this chapter is to evaluate the effectiveness of the Social Index, which reflects public perception of ESG. The results provide a significant foundation for exploring the correlation between corporate ESG performance and public perception, demonstrating that this new index can make a practical contribution to ESG management evaluation.

4.2 ANOVA Analysis

ANOVA (Analysis of Variance) is a statistical method used to test the mean difference between two or more groups. ANOVA evaluates whether the differences between groups are due to random chance by comparing between-group variance with within-group variance.

Two-way ANOVA is a technique that analyzes the effect of two independent variables and their interaction on a dependent variable. This method allows for simultaneous evaluation of the effect of each independent variable and their interaction on the dependent variable. In this study, the first independent variable was considered to be the type of data—comment data and report data—while the second independent variable was the SDGs category. The mean similarity score differences for each group were analyzed based on these factors. The analysis results for the two companies are presented in <Table 9> and <Table 10> below.

A significant difference was found between comment data and report data for both companies. To identify which specific categories showed meaningful differences, a post hoc test was conducted, using the Tukey HSD (Honest Significant Difference) test in this study. Tukey HSD is a post hoc test method that evaluates significance by comparing the mean differences between all possible pairs of groups.

The Tukey HSD test results showed that, except for SDG05 in Company A, there were significant

〈Table 9〉 Two-way ANOVA Results of A Company's Report and Comments on ESG Category Evaluation

Factor	Degrees of Freedom	F-value	p-value
E: SDG12~15	3.0	605.71	0.0
Data (Social Data / Report)	1.0	76.61	2.13e-18

Factor	Degrees of Freedom	F-value	p-value
S: SDG01~11	10.0	1747.44	0.0
Data (Social Data / Report)	1.0	1765.34	0.0

Factor	Degrees of Freedom	F-value	p-value
G: SDG16, 17	1.0	5945.04	0.0
Data (Social Data / Report)	1.0	208.38	4.50e-47

〈Table 10〉 Two-way ANOVA Results of B Company's Report and Comments on ESG Category Evaluation

Factor	Degrees of Freedom	F-value	p-value
E of ESG (SDG12~15)	3.0	3070.28	0.0
Data (Social Data / Report)	1.0	3166.81	0.0

Factor	Degrees of Freedom	F-value	p-value
S of ESG (SDG1~11)	3.0	1908.82	0.0
Data (Social Data / Report)	1.0	9017.08	0.0

Factor	Degrees of Freedom	F-value	p-value
G of ESG (SDG16, 17)	3.0	5355.13	0.0
Data (Social Data / Report)	1.0	6341.46	0.0

differences between the SDG categories of all comments and the SDG categories of reports for both companies. The test results are included in Appendix 3. However, even when statistical significance is observed, it is necessary to determine whether the difference is practically meaningful. The data used in this study have a large sample size, which means that even minor differences can be detected as statistically significant. Therefore, researchers should visually analyze the patterns between comments and reports for each SDG category to make a direct judgment on whether the differences are meaningful.

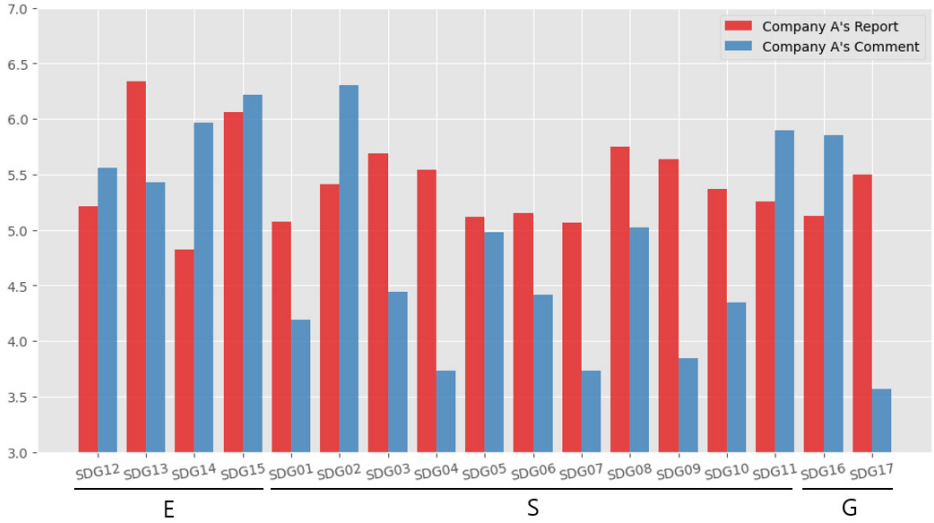
4.3 Evaluation and Comparison of Company A and Company B

4.3.1 Evaluation of Company A by Reports and Comments

The SDGs within the Environmental (E) category include SDG12 (Responsible Consumption and Production), SDG13 (Climate Action), SDG14 (Life Below Water), and SDG15 (Life on Land).

In Company A's report, the similarity scores for SDG13 and SDG15 recorded the highest average scores among all SDG categories, including the E category. In contrast, SDG14 received the lowest score among all SDGs. In the comments, the SDGs in the E category received relatively higher scores compared to other categories, particularly SDG14 and SDG15, which received high evaluations. However, while SDG13 scored the highest in the report, it received the lowest score in the E category in the comments. Additionally, SDG14, which received the lowest evaluation in the report, recorded relatively high scores in the comments. These patterns suggest that in the E category, the report and the comments provide contrasting evaluations.

The Social (S) category includes SDG01 (No Poverty), SDG02 (Zero Hunger and Sustainable Agriculture), SDG03 (Good Health and Well-being), SDG04 (Quality Education), SDG05 (Gender Equality), SDG06 (Clean Water and Sanitation),



〈Figure 1〉 Average Similarity Scores for Company A by ESG Category

SDG07 (Affordable and Clean Energy), SDG08 (Decent Work and Economic Growth), SDG09 (Industry, Innovation, and Infrastructure), SDG10 (Reduced Inequality), and SDG11 (Sustainable Cities and Communities).

In Company A's report, similar scores were assigned to SDG01, SDG05, SDG06, and SDG07, which were among the lower-rated categories within the Social (S) category. In contrast, SDG02, SDG03, SDG04, SDG08, and SDG09 received relatively favorable scores. In the comments, SDG02, SDG05, SDG08, and SDG11 received relatively higher scores. In addition to the differences in category-specific evaluations between the report and comments, there was a significant score discrepancy between the report and comments in most S category items, excluding SDG05 and SDG11. Specifically, SDG05 showed similar scores in both the report and comments, and ANOVA results indicated that it was the only ESG category without a significant difference between report and comment scores. For SDG11, the comments gave a relatively more positive evaluation than the report. Overall, the S category generally received higher scores in the report than in the comments. The large gap between report and comment scores suggests that while the report evaluated Company A's social performance relatively favorably, public perception did not align with this assessment.

The Governance (G) category includes SDG16 (Peace, Justice, and Strong Institutions) and SDG17 (Partnerships for the Goals).

For SDG16, the comments were relatively more positive than the report. Although the report's evaluation of SDG16 was neither particularly high nor low compared to other categories, the comment evaluation ranked relatively high across all SDGs, indicating a socially positive perception of Company A's performance related to SDG16. In contrast, SDG17 showed a large difference in scores between the report and comments, with SDG17 receiving the lowest score among all categories in the comments. Below is a table summarizing the comparison of Company A's report and comment scores by SDG category for each area.

Overall, in most SDG categories, Company A's report provided a more positive evaluation compared to the comments, maintaining a generally consistent assessment. In contrast, the comments gave relatively high ratings in the Environmental (E) category but assigned mostly lower scores in the Social (S) and Governance (G) categories, with a few exceptions. Notably, the categories that received lower scores in the comments showed a pattern of greater variability, unlike the

〈Table 11〉 Comparison of A Company's Report and Comments on SDG Category Evaluation

SDG Category	Specific SDGs	Average Report Score	Average Comment Score	Key Features
Environment (E)	SDG12 ~ SDG15	SDG13, SDG15 high, SDG14 low	SDG14, SDG15 high, SDG13 low	Contrasting scores between the report and comments.
Social (S)	SDG01 ~ SDG11	SDG01, SDG05, SDG06, SDG07 low; SDG02, SDG03, SDG04, SDG08, SDG09 high	SDG02, SDG05, SDG08, SDG11 high	Report scores were generally much higher than comments.
Governance (G)	SDG16, SDG17	SDG16 middle, SDG17 high	SDG16 high, SDG17 low	Relatively higher comment score for SDG16, large difference with SDG17, where the comment score was lowest.

balanced evaluation in the report.

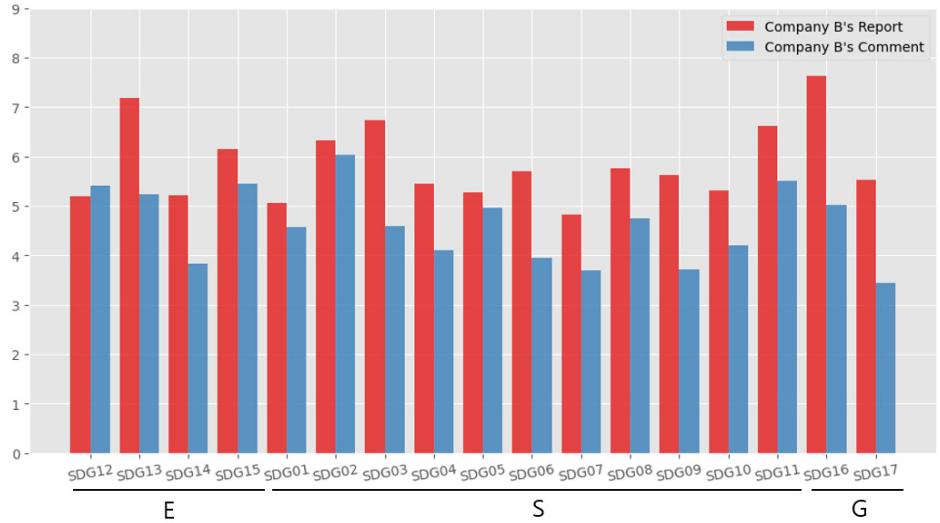
4.3.2 Evaluation of Company B by Reports and Comments

In the Environmental (E) category, Company B's report received similar scores for SDG12 and SDG14, which can be considered average when compared to other SDG categories. In contrast, SDG13 and SDG15 recorded relatively high scores, ranking among the highest compared to other categories.

In the comments, all three categories (SDG13, SDG14, SDG15) except for SDG12 received lower scores than in the report. SDG12 recorded a slightly higher score than in the report, though at almost the same level. Notably, SDG12, SDG13, and SDG15 received similar scores in the comments, and these scores were relatively high compared to other SDGs in the Social (S) and Governance (G) categories. However, SDG14 received a lower score than these three categories and was rated relatively low compared to other SDGs.

In the Social (S) category, Company B's report recorded relatively high scores for SDG02, SDG03, and SDG11 when considering all SDGs. The other SDGs in the S category also received generally average evaluations without particularly low scores.

In contrast, the pattern of evaluation in the comments was somewhat different. Overall, the scores in the comments were lower than those in the report, and the range of score variation in the comments was much greater than in the report. Some categories showed similar scores to the report, while others exhibited significant differences. Specifically, SDG01, SDG02, and SDG05 had lower scores in the comments than in the report, though they received relatively similar evaluations. Notably, SDG02 received the highest score in the comments, just as in the report. In contrast, SDG03, SDG06, and SDG09 showed considerable differences in scores between the report and comments. SDG11 also displayed a large difference, though it was rated highly in the comments relative to



〈Figure 2〉 Average Similarity Scores for Company B by ESG Category

〈Table 12〉 Comparison of B Company's Report and Comments on SDG Category Evaluation

SDG Category	Specific SDGs	Average Report Score	Average Comment Score	Key Features
Environment (E)	SDG12 ~ SDG15	SDG13, SDG15 high	SDG13, SDG14, SDG15 low	Report scores higher than comment scores across the board.
Social (S)	SDG01 ~ SDG11	SDG02, SDG03, SDG11 high	SDG02, SDG11 high, SDG03, SDG06, SDG09 low	Comments generally have lower scores compared to reports, with a larger variance in scores.
Governance (G)	SDG16, SDG17	SDG16 high, SDG17 average	SDG17 low	The largest score discrepancy was seen for SDG16, with comments rating it lower than the report.

other categories. This indicates that SDG11 was also positively perceived socially, but the report rated it even higher.

In the Governance (G) category, SDG16 received the highest score among all SDGs in the report, while in the comments, it received an average evaluation compared to other categories. However, SDG16 showed the largest score difference between the report and comments. SDG17 received the lowest score among all SDGs in the comments, while in the report, it was evaluated at a relatively average level. Below is a table summarizing the comparison of SDG category scores in Company B's report and comments for each area.

Overall, in the comments, the evaluation dynamically changes across categories, while the report maintains a relatively balanced assessment, with some items receiving higher scores than others at certain points.

4.4 Addressing Gaps in ESG Evaluation: A Dynamic Public Opinion Approach

4.4.1 Limitations of Existing Research

Conventional ESG evaluation methodologies have primarily relied on structured data provided by corporate reports or reputable rating agencies. This approach has limitations, as it fails to adequately reflect public perception and offers static evaluation results. For example, survey data reflects the opinions of a limited number of respondents at specific points in time, which may not encompass the diverse voices of stakeholders in ESG evaluation [5, 41]. Additionally, existing studies dependent on structured data struggle to effectively capture how corporate ESG activities are perceived by the public in real-time, diminishing the effectiveness of these evaluations.

Some prior studies in the literature have attempted ESG evaluation methods utilizing social media data; however, most of these studies employed surface-level keyword frequency analysis or structured sentiment analysis models for analyzing text data [17]. This approach has limitations, as it does not sufficiently account for the context within the text and struggles to capture the semantic relationships between words, which is crucial when handling complex concepts like ESG.

4.4.2 Study Contribution: Developing ESG Metrics with Real-Time Public Opinion

This study aimed to overcome the limitations

of previous research by analyzing social media data in real-time using deep learning-based natural language processing (NLP) techniques. By applying the GloVe (Global Vectors for Word Representation) embedding model, the study enabled a deeper analysis of the semantic relationships between words within the text. This approach goes beyond traditional keyword frequency analysis, allowing for a more accurate understanding of the connections between ESG and SDG (Sustainable Development Goals). The new ESG evaluation metric proposed in this study enhances the timeliness of ESG evaluations through a dynamic evaluation methodology that reflects real time public opinion.

First, the importance of real-time data analysis was confirmed. Previous research generally relied on data from specific points in time, while this study used unstructured social media data to reflect real-time public opinion. This allowed for a dynamic analysis of how corporate ESG activities are perceived by the public in real-time, which contributed significantly to enhancing the reliability of ESG evaluations.

Second, a refined analysis of text data was performed. While previous studies relied on simple sentiment or frequency analysis, this study used the GloVe model to capture word similarity and semantic relationships precisely. This enabled a more in-depth analysis of public perception related to ESG and the relationships between various SDG categories, contributing to the evolution of ESG evaluation from a static report-based approach to one that dynamically reflects public voices in real-time.

Third, the study strengthened the linkage between ESG and SDGs. By analyzing the connections between SDG goals and ESG activities, this study quantitatively evaluated how public perception of ESG aligns with sustainability goals. This approach provided a methodology to analyze the real-time impact of ESG activities on specific SDG goals, offering a way to capture real-time public opinions that previous ESG evaluations may have overlooked.

This study academically addresses the limitations of traditional ESG evaluation methodologies, which fail to incorporate real-time public opinion, and enables a deeper understanding of the relationship between ESG and SDGs through unstructured data analysis. By introducing deep learning-based text analysis techniques, the study was able to capture the dynamic changes in public perception, which were not discussed in previous research, thereby highlighting the potential for ESG evaluation to develop into a more comprehensive and reliable approach.

Furthermore, this study offers important practical implications. Companies can perform more transparent and dynamic ESG evaluations through an ESG Social Index that reflects real-time public opinion, providing a foundation for establishing tailored ESG strategies. Especially by considering the possibility of differing public perceptions, companies can differentiate their ESG strategies and potentially implement a real-time response system to achieve more accurate evaluation outcomes.

5. Conclusions

This study proposed a novel approach to ESG (Environmental, Social, and Governance) performance evaluation by developing a Social Index that integrates real-time public opinion using social media data and natural language processing techniques. Traditional ESG assessments, based

largely on structured and periodically updated corporate disclosures, often fail to reflect evolving public sentiment. In contrast, the proposed model, employing the GloVe embedding algorithm, analyzes the semantic similarity between public comments and SDG-related concepts to quantify alignment between corporate ESG activities and stakeholder perceptions.

Empirical findings from Company A and B revealed notable discrepancies between corporate sustainability reports and public sentiment, particularly in the Social (S) and Governance (G) domains. While Environmental (E) categories showed relatively high consistency, public opinion often diverged from corporate self-assessments in S and G areas, especially around SDG16 (Peace, Justice, and Strong Institutions). ANOVA tests confirmed that these differences were statistically significant, highlighting the need for companies to improve transparency and communication in underperforming domains. Temporal analysis also demonstrated the responsiveness of public opinion to external events, suggesting that real-time sentiment tracking can provide valuable insights for adjusting ESG strategies dynamically.

The ESG Social Index enhances the responsiveness and reliability of ESG evaluations by capturing real-time, unstructured data and offering tailored insights across industries and regions. For instance, companies operating in environmentally sensitive areas or socially scrutinized sectors can use the Index to develop targeted ESG strategies aligned with local stakeholder expectations. Ultimately, this index provides a bridge between internal ESG initiatives and external stakeholder perceptions, improving the alignment between corporate activities and societal demands.

Meanwhile, in January 2024, the International Sustainability Standards Board (ISSB) officially implemented its sustainability disclosure standards namely, IFRS S1 (General Requirements for Disclosure of Sustainability-related Financial Information) and IFRS S2 (Climate-related Disclosures). This marks a shift in regulatory frameworks toward unifying global ESG disclosure standards and requiring more quantitative and comparable sustainability information. Such regulatory developments indicate that companies will be expected to go beyond reporting internal data alone, and instead provide more multidimensional disclosures that incorporate the perceptions and responses of external stakeholders.

In this context, the ESG Social Index proposed in this study can function as a real-time, unstructured, perception-based indicator that complements the existing structured, internally focused disclosure metrics. Specifically, it has the potential to expand the scope of “financially material sustainability information” emphasized in IFRS S1, and to serve as a real-time monitoring tool aligned with IFRS S2’s emphasis on identifying and responding to climate-related risks. Whether this index will remain a supplementary reference tool or evolve into a core indicator within ESG evaluation frameworks will depend on its practical adoption by companies, acceptance by regulatory bodies, and the accumulation of academic validation.

Accordingly, this study aims to make a substantive contribution to the advancement of ESG disclosure and evaluation systems by introducing a model that integrates real-time public perception into traditional disclosure-based ESG assessment approaches.

To advance the practical application and academic value of this model, several recom-

mendations are proposed:

- Diversify Data Sources: Incorporate additional sources such as news media, governmental publications, and NGO reports to enhance representativeness.
- Advance NLP Techniques: Utilize state-of-the-art models beyond GloVe for greater semantic nuance and context sensitivity.
- Conduct Longitudinal Analyses: Evaluate long-term shifts in public opinion to assess the sustainability of ESG initiatives over time.
- Pursue Comparative Research: Examine public ESG perception across different industries and regions to inform more localized and effective strategies.
- Refine Evaluation Models: Develop hybrid models that combine qualitative and quantitative metrics for a multidimensional assessment.
- Foster Industry-Academic Collaboration: Strengthen partnerships with companies to promote practical implementation and refine ESG strategies through real-time feedback.

In conclusion, this study demonstrates the potential of integrating public opinion into ESG evaluation frameworks, enabling companies to enhance transparency, accountability, and trust. The proposed Social Index offers a dynamic and inclusive tool for aligning corporate sustainability efforts with stakeholder expectations, contributing to more credible and adaptive ESG management practices.

While this study presented a new methodology for ESG performance evaluation, it has several limitations, and future research should address various areas to complement these limitations.

First, as this study relied on social media data, there may be inherent data biases. Future research should expand by incorporating diverse data sources such as news, government reports, and non-profit organization publications to minimize bias.

Second, the text analysis technique used in this study may not fully capture the complexity and context of ESG-related discourse. This limitation becomes more pronounced when dealing with multilingual data. In particular, for Company B, a substantial number of public comments were originally written in French and subsequently translated into English using the DeepL AI translation tool. Although DeepL is known for its high-quality, context-aware translations, subtle nuances and cultural expressions inherent to the original language may still be lost or distorted in the translation process. Therefore, future research should consider validating translated text against original-language meanings or applying multilingual NLP models that can directly analyze texts in their native languages to ensure more accurate semantic interpretation.

Third, as this study used data collected at a specific point in time, longitudinal research is needed to analyze how public perception of ESG performance changes over time. This would enable a more comprehensive evaluation of the long-term impacts of ESG initiatives.

Fourth, since this study used data limited to specific industries and regions, future research should conduct comparative studies across industries and regions to enable a broader ESG performance assessment. This would facilitate the development of tailored ESG strategies that reflect industry characteristics and regional contexts.

Fifth, while the Social Index proposed in this study reflects real-time public opinion, it may not

be temporally aligned with conventional ESG indicators or corporate performance metrics, which are typically reported on a quarterly or annual basis. This time discrepancy raises important questions about how to interpret the Social Index in relation to other ESG variables. Nevertheless, this temporal characteristic also suggests that the Social Index may serve as a leading indicator, capturing early signs of reputational risks or shifts in stakeholder perception before they are reflected in formal disclosures. Future research should explore this potential by designing longitudinal studies or time-lagged models to examine whether changes in the Social Index can predict subsequent shifts in internal ESG ratings or corporate behavior.

Finally, it is important to acknowledge that social media data may be susceptible to user-related biases influenced by corporate communication strategies. For instance, large corporations often have more resources to engage in promotional activities, including partnerships with influencers or coordinated campaigns, which may disproportionately shape public sentiment on social platforms. Such dynamics can distort the perceived ESG performance, especially in the short term. Future research should therefore explore methods to detect and control for such influence-driven biases, such as weighting mechanisms, bot filtering, or sentiment validation.

References

[1] Abughniem, M.S., Al Aishat, M.H., and Hamdan, A.M.(2019), “Corporate sustainability as an antecedent to the financial performance: An empirical study,” Polish Journal of Management Studies, 20(2), 35-44.

[2] Barbosa, A., and Basilio, M.C.(2023), “Integration of Environmental, Social, and Governance (ESG) criteria: their impacts on corporate sustainability performance,” Humanities and Social Sciences Communications.

[3] Bassen, A., and Kovacs, A.M.(2008), “Environmental, social and governance key performance indicators from a capital market perspective,” Zeitschrift Für Wirtschafts- Und Unternehmensethik, 9(2), 182-193.

[4] Bebbington, J., and Unerman, J.(2020), “Achieving the United Nations Sustainable Development Goals: An enabling role for accounting research,” Accounting, Auditing .(Accountability Journal, 33(7), 1657-1670.

[5] Berg, F., Kölbel, J.F., and Rigobon, R.(2020), “Aggregate confusion: The divergence of ESG ratings. MIT Sloan School of Management.

[6] Bhandari, K.R., Ranta, M., and Salo, J.(2022), “The resource-based view, stakeholder capitalism, ESG, and sustainable competitive advantage: The firm’s embeddedness into ecology, society, and governance,” Business Strategy and the Environment, 31(4), 1525-1537.

[7] Bhattacharya, C.B., Sen, S., and Korschun, D.(2008), “Using corporate social responsibility to win the war for talent,” MIT Sloan Management Review, 49(2), 37-44.

[8] Borms, S., Boudt, K., Van Holle, F., and Willems, J.(2021), “Semi-supervised text mining for monitoring the news about the ESG performance of companies,” In Data Science for Economics and Finance(pp. 217-239). Springer, Cham.

[9] Chang, R., Chu, L., Tu, J., Zhang, B., and Zhou, G.(2020), “ESG and the market return,” Social Science Research Network.

[10] Cho, C., and Lee, H.(2023), “A comparison of ESG evaluation methods: Based on the K-ESG guidelines,” Intelligent Information Research, 29(1), 1-25.

[11] Chouaibi, Y., and Zouari, G.(2022), “The mediating role of real earnings management in the relationship between CSR practices and cost of equity: Evidence from European ESG data,” EuroMed Journal of Business. Advance online publication.

[12] Cornell, B., and Shapiro, A.C.(2021), “Corporate stakeholders, corporate valuation and ESG,” European Financial Management, 27(2), 196-207.

[13] De Lucia, C., Pazienza, P., and Bartlett, M.(2020), “Does good ESG lead to better financial performance by firms? Machine learning and logistic regression models of public enterprises in Europe,” Sustainability, 12(13), 5317.

[14] Eccles, R.G., Ioannou, I., and Serafeim, G.(2014), “The impact of corporate sustainability on organizational processes and performance,” Management Science, 60(11), 2835-2857.

[15] Galbreath, J.(2013), “ESG in focus: The Australian evidence,” Journal of Business Ethics, 118(3), 529-541.

[16] GRI(Global Reporting Initiative)(2019), “Business reporting on the SDGs: Integrating the SDGs into Corporate Reporting: A Practical Guide. https://www.globalreporting.org/public-policy/sustainable-development/integrating-sdgs-into-sustainability-reporting/.

[17] Guo, T., Jamet, N., Betrix, V., Piquet, L.-A., and Hauptmann, E.(2020), “ESG2Risk: A deep learning framework from ESG news to stock volatility prediction.

[18] Jin, W.(2022), “Nike’s relation with non-government-organizations. In 2021 International Conference on Public Art and Human Development(ICPAHD 2021), Atlantis Press, 1132-1135.

[19] Jung, S.(2024), “A study on the development of ESG safety evaluation indicators from a disaster management perspective,” Master’s thesis, Hanyang University.

[20] Kang, K., Park, J., and Na, H.(2023), “The impact of corporate ESG ratings on accrual earnings management and real earnings management,” Korean Journal of Accounting Research, 107, 139-163.

[21] Kim, R., Kim, T., and Heo, J.(2022), “Structural relationship analysis between ESG management, job satisfaction, and financial performance,” Ethical Management Review, 22(2), 27-47.

[22] Lee, H.(2021), “A study on the utilization of environmental impact assessment in ESG management evaluation indicators.(Master’s thesis, Hanse University).

[23] Lee, K.(2022), “A study on the calculation of input criteria for ESG performance indicators by country,” Korean Journal of Port Economics, 38(2), 31-47.

[24] Lee, R., and Kim, J.(2021), “Developing a social index for measuring the public opinion regarding the attainment of sustainable development goals,” Sustainability Research Journal, 12(3), 123-135.

[25] Park, H.(2020), “Recent trends and major issues in global ESG investment,” Capital Market Focus, 11(2).

[26] Park, S., and Shin, E.(2021), “Analysis of the correlation between ESG performance and

financial value for large, medium, and small companies," Environmental Policy, 29(4), 151199.

[27] Park, J.M., Jung, M.S., Kim, H.S., and Kim, S.M.(2024), "Adjusted portfolio selection model reflecting news sentiment analysis: Focusing on South Korean stock market," Journal of the Korean Operations Research and Management Science Society, 49(4), 57-72.

[28] Park, S.R., and Jang, J.Y.(2021), "The impact of ESG management on investment decision: Institutional investors' perceptions of country-specific ESG criteria," International Journal of Financial Studies, 9(3), 48.

[29] Park, Y., and Han, S.(2021), "The impact of corporate ESG activities on corporate image, perceived price fairness, and consumer response," Journal of Management Studies, 50(3), 643-664.

[30] Porter, M.F.(1980), "An algorithm for suffix stripping," Program: Electronic library and information systems, 14(3), 130-137.

[31] Rosati, F., and Faria, L.G.D.(2019), "Addressing the SDGs in sustainability reports: The relationship with institutional factors," Journal of Cleaner Production, 215, 1312-1326.

[32] Serafeim, G., and Yoon, A.(2022), "Which corporate ESG news does the market react to?," Financial Analysts Journal, 78(1), 59-78.

[33] Sokolov, A., Mostovoy, J., Ding, J., and Seco, L.(2020), "Building machine learning systems to automate ESG index construction.

[34] Sokolova, N.A., and Teymurov, E.S.(2022), "Correlation of sustainable development goals and ESG principles," Courier of Kutafin Moscow State Law University(MSAL),(12), 171-183.

[35] Song, G.(2022), "Directions for the formation of an ESG labor ecosystem," Korea Labor Society Research Institute Issue Paper, 2022(9), 1-21.

[36] Stead, J.G., and Stead, W.E.(2014), "Building spiritual capabilities to sustain sustainabilitybased competitive advantages," Journal of Management, Spirituality & Religion, 11(2), 143-158.

[37] Stout, L.(2012), "Shareholder value myth," Berrett-Koehler Press.

[38] S&P Global.(2023), "Key sustainability trends that will drive decision-making in 2023," S&P Global.

[39] Yoo, H., Lee, S., and Nam, J.(2021), "The impact of consumers' ESG management demands on ESG management support," Entrepreneurship & ESG Research, 1(2), 1-24.

[40] Van Zanten, J.A., and Van Tulder, R.(2018), "Multinational enterprises and the Sustainable Development Goals: An institutional approach to corporate engagement," Journal of International Business Policy, 1(3-4), 208-233

[41] Whelan, T., Atz, U., Van Holt, T., and Clark, C.(2021), "ESG and financial performance," Uncovering the Relationship by Aggregating Evidence from, 1, 2015-2020.