applied sciences

MDPI

Article

Green Finance in the Digital Public Sphere: A Multi-Method NLP Analysis

Mehmet Kayakuş [1], Mustafa Terzioğlu [2], Dilşad Erdoğan [3], Serdar Paçaci [4,*], Georgiana Moiceanu [5] and Razvan Dobrescu [5,6,*]

[1] Department of Biomedical Engineering, Faculty of Engineering, Akdeniz University, Antalya 07058, Türkiye; mehmetkayakus@akdeniz.edu.tr
[2] Accounting and Tax Department, Korkuteli Vocational School, Akdeniz University, Antalya 07800, Türkiye; mterzioglu@akdeniz.edu.tr
[3] Department of Finance, Korkuteli Vocational School, Akdeniz University, Antalya 07800, Türkiye; dilsaderdogan@akdeniz.edu.tr
[4] Department of Computer Engineering, Faculties of Technology, Isparta University of Applied Sciences, Isparta 32200, Türkiye
[5] Department of Entrepreneurship and Management, National University of Science and Technology Politehnica Bucharest, 60042 Bucharest, Romania; georgiana.moiceanu@upb.ro
[6] Department of Economics, National University of Science and Technology Politehnica Bucharest, 60042 Bucharest, Romania
* Correspondence: serdarpacaci@isparta.edu.tr (S.P.); razvan.dobrescu@upb.ro (R.D.)

Abstract

Green finance has become a key component of sustainability transitions by supporting the alignment of financial systems with environmental and climate-related objectives. As discussions on sustainable investments and climate finance increasingly take place in digital environments, analyzing large-scale user-generated content has become important for understanding online discussions and emerging discourse patterns. This study investigates green finance discourse in the digital public sphere using a multi-method Natural Language Processing (NLP) framework. A dataset of 23,452 posts collected from the platform X was analyzed using TF-IDF-based word frequency analysis, FinBERT-based sentiment analysis, Latent Dirichlet Allocation (LDA) topic modelling, and keyword co-occurrence network analysis. The results indicate that green finance discourse is dominated by positive (44.61%) and neutral (45.85%) sentiment, suggesting a generally favorable and institutionalized public perception. Topic modelling identified three dominant themes: climate finance and sustainability transition, digital finance and speculative investment narratives, energy investments and financial infrastructure. Network analysis revealed that sustainability, climate, energy, investment, and finance constitute the core conceptual structure of discourse. These findings demonstrate the effectiveness of NLP-based approaches for analyzing large-scale digital discussions and provide insights for policymakers, financial institutions, and organizations seeking to better understand online discussions surrounding green finance and support more effective sustainability communication strategies.

Keywords: green finance; natural language processing; FinBERT; sentiment analysis; Latent Dirichlet Allocation; text mining; social media analytics; sustainable finance

Check for updates

Academic Editors: Vincenza Carchiolo and Michele Malgeri

Received: 23 June 2026
Revised: 13 July 2026
Accepted: 20 July 2026
Published: 21 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The accelerating urgency of climate change and sustainability transitions has fundamentally reshaped the strategic landscape in which businesses, financial institutions, and

policymakers operate. In this context, green finance has emerged as a critical mechanism for aligning capital allocation with environmental objectives, enabling the transition toward low-carbon and sustainable economic systems. Beyond its role as a financial instrument, green finance increasingly functions as a strategic tool, through which organizations signal commitment to sustainability, manage risk, and position themselves within evolving regulatory and market environments. As such, it occupies a central role in contemporary debates on sustainable development and business strategy.

While green finance has been widely examined from financial and regulatory perspectives, growing volumes of user-generated content on digital platforms provide new opportunities for understanding how sustainability-related financial concepts are discussed and interpreted within online discussions on the platform X. Social media platforms have become important sources of real-time information, reflecting public reactions, emerging themes, and sentiment dynamics related to sustainability and finance. Consequently, analyzing digital discourse can provide valuable insights into the evolving public perception of green finance.

Despite the growing prominence of green finance in both academic and policy agendas, existing research has largely adopted a finance-centric perspective, focusing on instruments such as green bonds, ESG investments, and climate finance mechanisms. While these studies provide important insights into institutional design and market performance, they tend to overlook the strategic and discursive dimensions of green finance, particularly how it is framed, perceived, and contested within digital communication environments. At the same time, research utilizing social media data often isolates analytical dimensions—such as sentiment or topic modelling—without capturing the multidimensional nature of discourse, where meaning is simultaneously shaped by thematic structures, emotional orientations, and relational patterns.

Addressing these limitations, this study investigates green finance discourse in the digital public sphere using a multi-method Natural Language Processing (NLP) framework. By integrating TF-IDF-based text mining, FinBERT sentiment analysis, LDA topic modelling, and keyword co-occurrence network analysis, the study provides a comprehensive examination of thematic structures, sentiment patterns, and conceptual relationships within large-scale social media data. This integrated analytical framework enables a multidimensional understanding of public discussions surrounding green finance.

The study makes three key contributions. First, it applies an integrated NLP framework to examine large-scale social media discussions on green finance, combining complementary analytical techniques within a single workflow. Second, rather than introducing a new analytical method, the study demonstrates how TF-IDF analysis, FinBERT sentiment analysis, LDA topic modelling, and keyword co-occurrence network analysis can be integrated to provide a multidimensional interpretation of digital green finance discourse. Third, by jointly examining thematic structures, sentiment patterns, and conceptual relationships, the study offers practical insights for policymakers, financial institutions, and sustainability practitioners while illustrating the value of combining established NLP techniques for sustainability communication research.

Conceptually, the proposed framework assumes that digital discourse emerges through multiple complementary analytical dimensions. Lexical patterns constitute the linguistic basis of discourse, thematic structures organize these lexical elements into coherent narratives, sentiment analysis reveals their affective orientation, and keyword co-occurrence networks capture the structural relationships through which narratives become reinforced. Integrating these complementary dimensions provides a conceptual interpretative framework for understanding how green finance discourse is constructed and communicated in the digital public sphere.

https://doi.org/10.3390/app16147315

It should be noted that the objective of this study is not to develop or benchmark new NLP algorithms. Rather, the proposed framework integrates well-established and complementary analytical techniques to provide a multidimensional interpretation of green finance discourse. Accordingly, the contribution of this study lies in methodological integration and practical application rather than algorithmic performance comparison.

2. Literature

The concept of green finance has been widely examined in recent years, primarily within the domains of financial economics, sustainability policy, and institutional governance. Existing studies have largely focused on financial instruments such as green bonds, ESG investments, and climate finance mechanisms, as well as their roles in supporting low-carbon transitions and sustainable development goals [1,2]. While this body of research provides valuable insights into the structural and regulatory dimensions of green finance, it tends to conceptualize the phenomenon as a technical and market-oriented construct, often overlooking its social and discursive dimensions.

Recent studies have increasingly emphasized the importance of examining public discourse surrounding sustainability-related financial concepts. As green finance intersects with environmental policy, investment decisions, and technological innovation, public discussions on digital platforms provide valuable insights into how these concepts are perceived, interpreted, and associated with broader sustainability objectives. Consequently, analyzing large-scale social media data can contribute to a better understanding of emerging themes, public sentiment, and communication patterns related to green finance.

The transformation of communication technologies has further amplified these dynamics. Social media platforms function as a contemporary digital public sphere in which individuals and institutions actively participate in shaping public discourse. Drawing on the notion of the public sphere, digital environments enable the circulation of multiple and often conflicting narratives, thereby influencing how concepts such as sustainability and green finance are perceived and legitimized. Studies examining climate change discourse on social media have demonstrated that such platforms operate as "unsolicited public opinion spaces", where sentiment, framing, and thematic clustering reflect broader societal attitudes [3,4]. These insights are directly transferable to green finance, as the concept intersects with climate policy, economic transformation, and sustainability transitions.

Methodologically, the increasing availability of digital trace data has led to the growing use of text mining and natural language processing techniques in analyzing public discourse [5]. Approaches such as TF-IDF-based term weighting, topic modelling, and sentiment analysis have been widely applied to short-text environments such as social media [6,7]. TF-IDF enables the identification of distinctive terms within large text corpora by balancing term frequency and document-level distribution [8,9]. Topic modelling techniques, especially Latent Dirichlet Allocation (LDA), provide a probabilistic framework for uncovering latent thematic structures within text collections [10]. At the same time, advances in deep learning have led to the adoption of transformer-based models such as BERT, attention mechanisms, and graph-based deep learning architectures, which have demonstrated strong performance in extracting complex semantic and structural patterns across various application domains [11–13]. These developments further highlight the growing capability of modern artificial intelligence techniques for analyzing complex textual and relational data.

In addition to thematic and sentiment-based analyses, the relational dimension of discourse has gained increasing attention. Keyword co-occurrence networks allow researchers to examine how concepts are interconnected and clustered within a discourse, revealing the structural relationships that underpin meaning construction [14–16]. This re-

lational perspective is particularly important for understanding complex phenomena such as green finance, where multiple domains—environmental, financial, technological, and policy-related—intersect.

Despite these methodological advancements, a key limitation in the literature is the tendency to analyze these dimensions in isolation. Many studies focus either on thematic structures, sentiment patterns, or conceptual networks, without integrating these perspectives into a unified analytical framework. This fragmented approach restricts the ability to fully capture the multidimensional nature of discourse, particularly in the context of short-text digital data where meaning is often context-dependent and relational.

Addressing this gap, the present study adopts an integrated, multi-method approach that combines TF-IDF-based word frequency analysis, LDA topic modelling, FinBERT-based sentiment analysis, and keyword co-occurrence networks. By doing so, it enables a comprehensive examination of green finance discourse across three interrelated dimensions: thematic structure, emotional orientation, and conceptual relationships. This approach contributes to the literature by offering a more holistic understanding of how green finance is constructed, interpreted, and negotiated within the digital public sphere.

3. Materials and Methods

This study analyzed 27,376 tweets collected from the platform X between 1 and 31 January 2026, to investigate the digital discourse on green finance. During the data cleaning and text preprocessing stage, duplicate content, URL links, semantically weak terms, and contextually irrelevant expressions were removed, resulting in a final dataset of 23,452 tweets for analysis.

The study employed a multi-method text mining framework, including TF-IDF-based word frequency analysis, FinBERT-based sentiment analysis, and LDA topic modelling. In addition, the findings were supported by visualization techniques such as density visualizations, and keyword co-occurrence networks.

Figure 1 presents the overall NLP workflow adopted in this study.



Figure 1. NLP workflow.

Overview of the NLP-based analytical framework. The workflow includes data collection, preprocessing, TF-IDF word frequency analysis, FinBERT sentiment analysis, LDA topic modelling, visualization, and result interpretation for examining green finance discourse on the platform X.

This multi-layered methodological approach enables a comprehensive assessment of the thematic structures, emotional orientations, and conceptual relationships underlying green finance discourse in the digital public sphere.

Within this framework, each analytical technique serves a distinct but complementary interpretative role. TF-IDF identifies salient lexical signals, LDA organizes these signals into thematic narratives, FinBERT captures the emotional framing of these narratives, and keyword co-occurrence network analysis reveals the conceptual relationships through which discourse is structurally organized. Rather than interpreting these analytical outputs independently, the study integrates them into a single conceptual interpretative framework for analyzing green finance discourse.

Each analytical component was selected because it captures a different dimension of digital discourse. TF-IDF identifies salient terms, FinBERT captures domain-specific

sentiment, LDA uncovers latent thematic structures, and keyword co-occurrence networks reveal conceptual relationships among frequently occurring terms. These methods are therefore complementary rather than competing analytical approaches. Consequently, this study does not aim to evaluate the predictive superiority of individual algorithms but to demonstrate their combined analytical value for interpreting green finance discourse.

3.1. Data Set

This study utilized the platform X as a data source to examine the digital public discourse surrounding the concept of green finance. As a widely used social media platform, X functions as a contemporary digital public sphere in which diverse actors—including policymakers, financial institutions, investors, and individuals—engage in real-time discussions on global issues such as climate change, sustainability, and financial markets [17]. Its open and interactive structure makes it particularly suitable for capturing public interpretations and discursive dynamics related to emerging socio-economic concepts.

Data were collected from the platform X using the X API v2 by searching the keyword "green finance", which was selected due to its comprehensive nature as an umbrella term encompassing related subtopics such as sustainable finance, climate finance, and green investment [1,2]. This approach enabled the study to capture the broad discursive landscape of green finance without restricting the analysis to specific financial instruments or subdomains. Although the keyword "green finance" was selected as a broad umbrella term, we acknowledge that related discussions using alternative expressions such as "sustainable finance", "ESG investing", "climate finance", "green bonds", or "transition finance" may not have been fully captured. Consequently, the findings should be interpreted as representing online discourse explicitly associated with the term "green finance" rather than the entirety of sustainability-related financial discussions in the digital public sphere. Future studies may adopt broader keyword strategies to capture a more comprehensive range of sustainability finance discourse.

The data collection period spanned from 1 January to 31 January 2026. This timeframe was selected to provide a clearly defined and temporally consistent snapshot of public discourse on green finance. Rather than aiming to represent long-term or stable public opinion, the selected period enables the examination of discourse within a specific temporal context. Because public discussions may be influenced by policy announcements, market developments, or sustainability-related events occurring during the collection period, the findings should be interpreted as reflecting discourse during January 2026. Future studies employing longitudinal datasets would enable the investigation of temporal changes and the assessment of discourse stability over time [5].

A total of 27,376 tweets were initially collected, and after preprocessing, duplicate entries, irrelevant content, and textual noise were removed, resulting in a final dataset of 23,452 tweets. This dataset size provides a sufficiently large corpus for exploring thematic patterns and sentiment distributions in large-scale social media analyses [18,19]. Only publicly accessible English-language posts were retained to ensure compatibility with FinBERT. Duplicate posts and repeated content were removed to reduce redundancy, while retweets were retained only when they contained unique textual content; exact duplicate retweets were excluded. Replies and quote posts were included when they contained original text relevant to the search query. No automated bot- or spam-detection algorithm was applied, and this is acknowledged as a limitation of the study.

3.2. Text Mining

Text mining represents an interdisciplinary approach that combines natural language processing (NLP) and statistical techniques to identify patterns, relationships, and trends

within unstructured textual data [20]. In large-scale and unstructured environments such as social media, text mining provides a systematic framework for analyzing conceptual densities and discursive structures, enabling researchers to uncover the underlying organization of digital discourse [21]. In this study, text mining techniques were employed to examine both the conceptual and emotional dimensions of green finance discourse. Specifically, TF-IDF-based word frequency analysis and FinBERT-based sentiment analysis were utilized within a multi-method analytical framework.

Prior to analysis, the dataset underwent a series of preprocessing steps to enhance data quality and analytical reliability. All text was converted to lowercase, while URL links, special characters, and punctuation marks were removed. Additionally, stop words with limited semantic value were excluded, and duplicate content was eliminated. These preprocessing procedures reduced noise within the dataset and improved the consistency and robustness of subsequent analyses [22].

All analyses were conducted using Python 3.14.6. Data handling and preprocessing were performed using pandas and NumPy 2.5.0, while regular expressions were used to remove URLs, mentions, hashtags, special characters, and redundant spaces. Missing tweet texts were removed before analysis. For word frequency and visualization analyses, NLTK stopword lists were used, and only alphabetic tokens longer than two or three characters were retained depending on the analysis. Sentiment analysis was performed using the FinBERT model through the Hugging Face Transformers library. The model classified tweets into positive, negative, and neutral sentiment categories. Topic modelling was implemented using scikit-learn's CountVectorizer and LatentDirichletAllocation. The CountVectorizer was configured with English stop words, max_df = 0.95, and min_df = 2. The LDA model was estimated with three topics, random_state = 42, and batch learning. Keyword co-occurrence networks were generated using NetworkX 3.6.1 and Matplotlib 3.10.0, with a co-occurrence window size of 2, a minimum co-occurrence threshold of 2, and the 60 most frequent keywords included as nodes. A grid-based sensitivity analysis was performed by varying the minimum co-occurrence threshold (2–10), the number of retained high-frequency keywords (40–100), and the co-occurrence window size (2–5). The selected configuration (threshold = 2, top words = 60, window = 2) was retained because it preserved the complete conceptual structure while maintaining comparable network properties across all analyses.

The word frequency analysis was conducted using the TF-IDF (Term Frequency–Inverse Document Frequency) method, which identifies distinctive terms by balancing their frequency within individual documents against their distribution across the corpus [23]. This approach enabled the identification of salient concepts and thematic markers that structure green finance discourse.

From the perspective of information retrieval theory, TF-IDF is a lexical weighting method designed to estimate the discriminative importance of terms within a document collection rather than to capture semantic meaning [24,25]. Consequently, TF-IDF is effective for identifying salient lexical features but cannot resolve contextual ambiguity or polysemous expressions commonly found in short social media texts and financial discourse. In the present study, this limitation is intentionally addressed through a complementary analytical framework. FinBERT provides contextual semantic representations based on transformer architectures, LDA identifies latent thematic structures from document-level word distributions, and keyword co-occurrence network analysis captures conceptual relationships among frequently associated terms. Therefore, semantic interpretation in this study emerges from the integration of lexical, contextual, probabilistic, and relational analyses rather than from TF-IDF in isolation.

3.3. Sentiment Analysis

Sentiment analysis is a natural language processing technique used to determine the emotional orientation of textual expressions, typically categorized as positive, negative, or neutral [26]. It serves as a critical analytical tool for examining social attitudes and public perceptions reflected in user-generated content, particularly within social media environments [27]. In this study, sentiment analysis was employed to capture the emotional dimension of digital discourse surrounding the concept of green finance.

The sentiment analysis was conducted using FinBERT, a transformer-based language model specifically developed for financial text analysis [28]. Built upon the Bidirectional Encoder Representations from Transformers (BERT) architecture, FinBERT has been pre-trained and fine-tuned on large-scale financial corpora, enabling effective sentiment classification in finance-related textual data [29]. Given that green finance discourse frequently contains investment, market, capital, and sustainability-related terminology, FinBERT was considered particularly suitable for capturing contextual sentiment patterns within social media discussions. Unlike traditional lexicon-based approaches that rely on predefined sentiment scores, FinBERT evaluates the contextual meaning of words and sentences, allowing for more accurate identification of sentiment in short and domain-specific texts [30]. This capability is especially valuable in social media environments, where meaning is often shaped by contextual nuances and financial terminology [31]. The model was employed to classify tweets into positive, negative, and neutral sentiment categories, providing a robust representation of public attitudes toward green finance. Although FinBERT has demonstrated strong performance in financial sentiment analysis, no additional manual annotation or model-specific validation was conducted within the scope of this study. Consequently, the sentiment classifications rely on the pre-trained capabilities of the model. Furthermore, social media texts may contain sarcasm, promotional language, spam, or context-dependent expressions that remain challenging for transformer-based sentiment models. Therefore, the sentiment analysis results should be interpreted as an approximation of overall public sentiment rather than an error-free classification of individual posts. Future studies could improve analytical reliability by incorporating manually annotated validation datasets, reporting performance metrics, and comparing FinBERT with alternative transformer-based sentiment models.

3.4. Topic Modelling

Topic modelling is a probabilistic text analysis technique designed to uncover latent thematic structures within large text corpora by identifying patterns of word co-occurrence [32]. Rather than assigning texts to predefined categories, this approach enables the discovery of underlying themes that structure discourse, making it particularly suitable for analyzing complex and unstructured data such as social media content [7,33]. In this study, topic modelling was employed to reveal the thematic organization of green finance discourse within the digital public sphere.

The analysis utilized the Latent Dirichlet Allocation (LDA) model, a widely used probabilistic generative approach that assumes each document (i.e., each tweet) is composed of a mixture of topics, and each topic is represented by a distribution of words [34]. By examining patterns of word co-occurrence across the dataset, LDA identifies clusters of terms that correspond to latent thematic structures [35]. In this framework, documents are represented through topic distributions, while topics are defined by probabilistic word associations, enabling the systematic identification of dominant themes and their representative terms [36].

LDA was selected due to its interpretability and its extensive use in the literature for analyzing thematic structures in large-scale textual datasets, particularly in short-text environments such as social media [37]. Furthermore, the model allows for the quantitative

assessment of thematic prevalence by estimating both the distribution of topics across documents and the relative importance of each topic within the corpus [38]. This facilitates a structured examination of the thematic axes that underpin green finance discourse.

To determine the optimal number of topics, multiple LDA models ranging from two to ten topics were evaluated using coherence scores, which measure the semantic consistency and interpretability of generated topics. The resulting coherence scores were 0.630 for two topics, 0.680 for three topics, 0.670 for four topics, 0.640 for five topics, 0.610 for six topics, 0.590 for seven topics, 0.560 for eight topics, 0.580 for nine topics, and 0.600 for ten topics. Among the evaluated models, the three-topic model achieved the highest coherence score (0.680), indicating the best balance between semantic coherence and topic interpretability. Although the four-topic model produced a comparable coherence score (0.670), qualitative examination indicated that it generated more fragmented and overlapping themes. Therefore, the three-topic solution was selected for the final analysis.

As this study adopts an exploratory NLP framework, conventional statistical significance tests were not directly applicable to the analyses performed. Instead, methodological robustness was enhanced through several complementary procedures. The topic modelling results were evaluated using coherence scores to ensure semantic consistency. Sentiment analysis was conducted using the domain-specific FinBERT model, and the findings were interpreted by comparing results across TF-IDF analysis, topic modelling, sentiment analysis, and keyword co-occurrence network analysis. The consistency observed across these complementary analytical approaches provides additional support for the reliability of the findings. Although LDA has recognized limitations when applied to short texts such as social media posts, it remains one of the most widely adopted and interpretable topic modelling techniques in the literature. To improve topic quality, standard preprocessing procedures were applied to reduce textual noise, and the optimal number of topics was selected based on coherence score evaluation. Nevertheless, no formal human interpretability assessment or external validation of the extracted topics was conducted. Therefore, the identified topics should be interpreted as probabilistic thematic structures rather than definitive semantic categories. Future research may compare LDA with topic modelling approaches specifically developed for short texts, such as BERTopic or the Biterm Topic Model (BTM), and incorporate expert-based topic validation.

4. Results

The results of text mining, TF-IDF-based word frequency analysis, LDA topic modelling, and FinBERT-based sentiment analysis applied to a dataset of 23,452 tweets collected from the platform X between 1 and 31 January 2026 are presented within a comprehensive analytical framework following the preprocessing stage. The findings provide insights into how green finance is discussed within the digital public sphere, the concepts with which it is associated, and the emotional orientations shaping public discourse.

In addition, visualization techniques, including keyword co-occurrence networks and density visualizations, were employed to explore the thematic densities and conceptual clusters underlying the discourse. This multi-layered analytical approach enables a systematic evaluation of the social perception of green finance across both thematic and emotional dimensions.

The distribution of sentiment across the dataset is presented in Table 1.

Table 1. Sentiment analysis results.

Positive	Negative	Neutral
10.460	2.236	10.756

Based on the analysis of 23,452 tweets, 44.61% of the content was classified as positive, 9.54% as negative, and 45.85% as neutral. The findings indicate that positive and neutral discourse dominate at nearly equal levels, while negative sentiment remains comparatively limited.

This distribution suggests that green finance is predominantly discussed in an informative, policy-oriented, and supportive context within the digital public sphere. The high proportion of positive content reflects an optimistic perception of green finance, particularly in relation to climate finance, ESG practices, renewable energy investments, and sustainable economic transformation.

In contrast, negative discourse is primarily associated with concerns regarding regulatory complexity, transition costs, financial risks, and debates surrounding issues such as greenwashing. These findings indicate that while green finance is generally perceived positively, critical perspectives persist, particularly in relation to its economic and regulatory implications.

Overall, the sentiment distribution suggests that green finance is predominantly discussed in a supportive and institutional context within the analyzed dataset.

4.1. Word Frequency Analysis

To identify the conceptual axes underlying green finance discourse, a word frequency analysis was conducted using the TF-IDF method, with results categorized by sentiment. This approach enables the identification of distinctive terms within each sentiment category, thereby revealing how thematic emphases vary across emotional contexts. The results are presented in Table 2.

Table 2. Word frequency results.

Positive	Negative	Neutral
Growth	Risk	Global
Future	Market	Capital
Innovation	Stock	Financial
Clean	Billion	Summit
Strong	Fund	Innovation
Token	Investment	Development
Crypto	Transition	Market
Tech	Carbon	Bank
Project	Development	Transition
Good	Capital	Sustainability

The distribution of keywords across sentiment categories reveals that green finance discourse is structured around distinct thematic orientations. In the positive category, terms such as growth, future, innovation, cleanliness, and technology indicate that green finance is associated with progress, transformation, and forward-looking economic development. The presence of terms such as crypto and token further suggests a linkage between green finance and digital financial innovation, reflecting the integration of sustainability narratives with emerging financial technologies.

In contrast, the negative sentiment category is dominated by terms such as risk, market, investment, capital, and transition, indicating that critical discourse is primarily framed in terms of financial uncertainty, regulatory pressures, and structural transformation costs. The prominence of the term carbon suggests that environmental objectives are often discussed in relation to economic trade-offs and policy constraints. This indicates that negative sentiment is not directed toward sustainability itself but toward the complexities associated with its implementation.

The neutral category is characterized by terms such as global, financial, capital, bank, and sustainability, reflecting a predominantly institutional and macroeconomic framing of green finance. The presence of terms such as summit and development suggests that the discourse is frequently embedded in policy communication, corporate reporting, and international governance contexts.

Overall, the findings indicate that green finance discourse varies systematically across sentiment categories, with positive narratives emphasizing innovation and growth, negative narratives focusing on risk and transition challenges, and neutral discourse reflecting institutional and informational communication. This differentiation highlights the multidimensional nature of green finance as both a policy concept and a financial practice within the digital public sphere.

4.2. Topic Modelling Analysis

The thematic structure of green finance discourse in the digital public sphere was examined using the Latent Dirichlet Allocation (LDA) model. The final model was selected based on coherence score optimization, with the three-topic solution achieving the highest coherence score (0.68). This approach identifies latent topics by analyzing patterns of word co-occurrence across the dataset, enabling the detection of underlying thematic clusters within large-scale textual data [39].

For each topic, the mean topic probability and prevalence were calculated to assess both the relative importance and distribution of themes across the corpus. The results are presented in Table 3, which summarizes the dominant topics and their associated keyword structures.

Table 3. Topic modelling results.

Topic N	Mean Topic Probability	Prevalence	Topmost Co-Occurring Words
2	0.4404	45.44%	Green, finance, climate, sustainability, energy, capital, transition, investment, development, carbon
3	0.3400	35.05%	Green, finance, crypto, market, trading, stocks, bitcoin, tech, money, bank
1	0.2196	19.51%	Green, finance, energy, solar, power, investment, market, global, growth, data

4.2.1. Topic 2: Climate Finance and Sustainability Transition

Topic 2 was the most dominant topic (45.44%). It is structured around keywords such as climate, sustainability, energy, transition, investment, and carbon, indicating that green finance discourse is primarily framed within the context of environmental sustainability and the transition to a low-carbon economy. The co-occurrence of terms such as development, capital, and innovation further suggests that financial mechanisms are perceived as integral to enabling long-term structural transformation.

From a theoretical perspective, this topic reflects the alignment of financial systems with environmental governance, where green finance functions as a policy-oriented instrument supporting sustainability transitions. The dominance of this theme indicates that, within the digital public sphere, green finance is largely interpreted through a sustainability-driven narrative that integrates economic growth with environmental responsibility.

4.2.2. Topic 3: Digital Finance and Speculative Investment Narratives

Topic 3 was the second most prevalent topic (35.05%). It is characterized by keywords such as crypto, bitcoin, trading, market, stocks, and technology, revealing a strong association between green finance and digital financial ecosystems. This thematic cluster

indicates that green finance discourse is partially embedded within speculative investment environments and emerging financial technologies.

The findings indicate a strong association between green finance discourse and investment-oriented themes, including digital assets, trading, and market activities. The integration of green finance into cryptocurrency and digital trading narratives suggests that its meaning is being expanded beyond environmental concerns, indicating that green finance is increasingly discussed alongside digital finance and investment-related activities.

4.2.3. Topic 1: Energy Investments and Financial Infrastructure

Topic 1 accounted for 19.51% of the corpus, making it the least prevalent of the three identified topics. It is characterized by keywords such as energy, solar, power, investment, global, and growth, indicating a focus on large-scale energy projects and infrastructure financing. This topic reflects a more sector-specific discourse in which green finance is framed as a capital-intensive mechanism supporting renewable energy development and global investment flows.

From a theoretical standpoint, this theme highlights the role of green finance as a bridge between sustainability objectives and capital allocation processes. It emphasizes the material and infrastructural dimension of green finance, demonstrating how environmental goals are operationalized through large-scale financial investments in energy systems.

4.3. Data Visualization

Density visualization was employed to identify conceptual concentrations and co-occurrence patterns within green finance discourse. This method represents the spatial distribution of keywords based on their frequency and relational proximity, where areas of higher density indicate stronger thematic clustering and conceptual centrality [14,40].

The density map provides an overview of how key concepts are organized within the discourse and highlights the relative importance of thematic clusters. By revealing areas of high conceptual intensity, this approach complements word frequency analysis and co-occurrence network results, enabling a more comprehensive understanding of the structural organization of green finance discourse.

The distribution of conceptual densities in the dataset is presented in Figure 2.

Figure 2 displays the conceptual density and clustering of keywords within the dataset. The concentration of terms around green and finance indicates that these concepts function as the central anchors of discourse. High-density clusters formed by terms such as climate, energy, sustainability, and investment reveal the core thematic focus on environmental transition. Peripheral clusters suggest the presence of secondary themes related to digital finance and policy discourse. This structure demonstrates the multi-layered organization of green finance discourse.

A keyword co-occurrence network was constructed to examine the relational structure of concepts within green finance discourse. This method identifies conceptual clusters by analyzing the co-occurrence patterns of keywords within the same textual context [16]. In the network structure, nodes represent keywords, while edges indicate the frequency of co-occurrence between terms [15]. The size of nodes and the density of connections reflect the relative importance and centrality of concepts within the discourse.

This approach enables the identification of how key concepts are interconnected and how thematic clusters are formed within the dataset. By mapping these relationships, co-occurrence network analysis provides insight into the structural organization of discourse and the interaction between different conceptual domains.

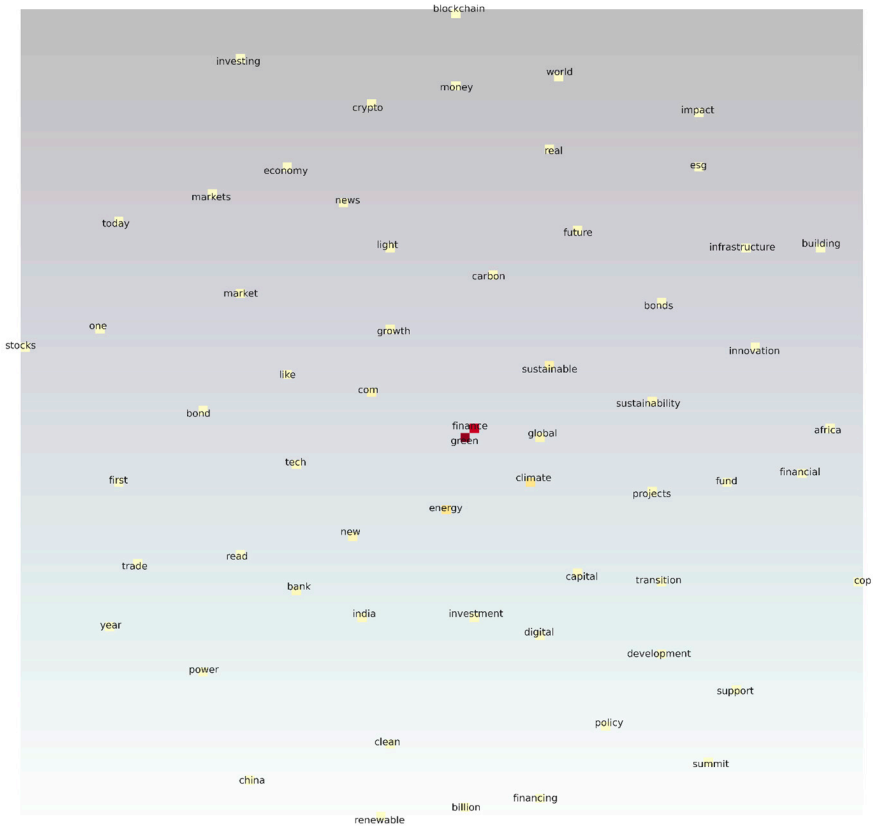
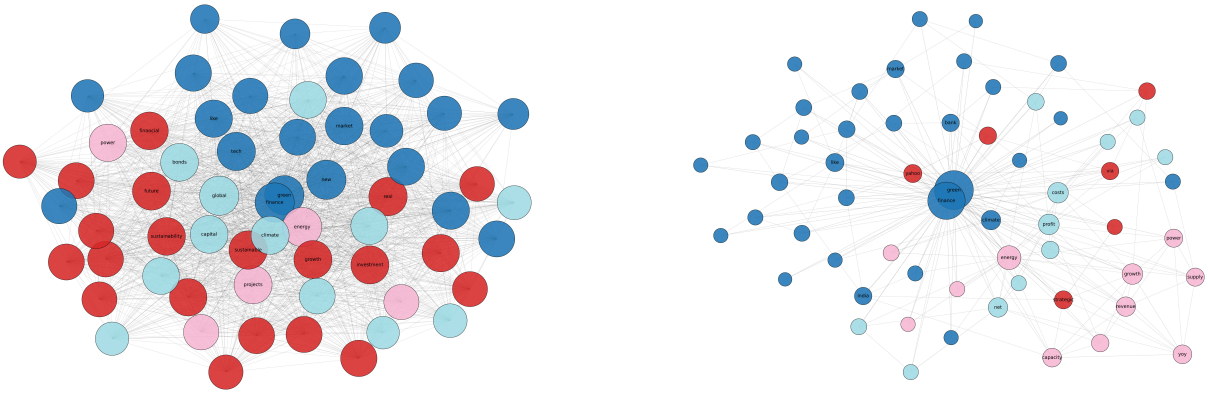


Figure 2. Density visualization.

Figure 3 presents separate keyword co-occurrence networks for positive and negative tweets, enabling a comparison of their conceptual structures. The corresponding quantitative network characteristics are summarized in Table 4.



(**a**) Positive keyword co-occurrence network (**b**) Negative keyword co-occurrence network

Figure 3. Sentiment-specific keyword co-occurrence networks for (**a**) positive and (**b**) negative tweets. Node size represents degree centrality, and node colors indicate communities identified using modularity-based community detection.

As shown in Table 4, the overall keyword co-occurrence network was highly interconnected (density = 0.831), indicating extensive conceptual relationships among the most frequently occurring terms. Compared with the overall network, the positive discourse network remained relatively dense (density = 0.545; modularity = 0.078), whereas the negative discourse network was considerably sparser (density = 0.128) and exhibited the highest modularity (0.155), suggesting a stronger separation into distinct conceptual communities.

Table 4. Quantitative characteristics of sentiment-specific keyword co-occurrence networks.

Network	Tweets	Density	Average Degree	Modularity	Average Path Length
Overall	5802	0.831	49.00	0.046	1.169
Positive	1339	0.545	32.17	0.078	1.455
Negative	240	0.128	7.53	0.155	1.886

Node-level centrality analysis further revealed structural differences between positive and negative discourse. In the positive network, green, finance, climate, sustainability, and energy exhibited the highest degree centrality values, indicating that supportive discussions were broadly organized around sustainability and climate-transition concepts. In contrast, the negative network was dominated by green and finance, which also displayed substantially higher betweenness centrality values, while economically oriented concepts such as profit, costs, and revenue became relatively more central. These findings indicate that negative discourse is concentrated around a limited number of economic concerns, whereas positive discourse exhibits a broader conceptual organization.

The keyword co-occurrence analysis was extended through standard graph-theoretical measures, including degree centrality, betweenness centrality, closeness centrality, eigenvector centrality, modularity, and community detection. In addition, separate positive and negative networks were constructed to compare sentiment-specific conceptual structures. Parameter sensitivity analyses were also conducted to evaluate the robustness of the selected co-occurrence settings.

5. Discussion

This study examines how green finance is represented and discussed within the digital public sphere using a multi-method NLP framework. By integrating sentiment analysis, topic modelling, and network analysis, the findings provide a comprehensive overview of thematic structures, sentiment patterns, and conceptual relationships within large-scale social media discourse. By integrating thematic, sentiment, and relational analyses, the findings extend existing literature, which has largely examined green finance through institutional, regulatory, and performance-based lenses [1,41]. In doing so, the study contributes to the growing literature on sustainable finance by demonstrating how NLP-based techniques can be used to analyze large-scale digital discussions and uncover thematic, emotional, and relational patterns within green finance discourse.

The results indicate that green finance discourse is positioned at the intersection of sustainability, governance, and investment-related themes. Discussions frequently combine environmental objectives with economic and financial considerations, highlighting the multidimensional nature of green finance in digital communication environments. Consistent with previous studies, our findings confirm that sustainability, climate action, and low-carbon transition remain the dominant themes in green finance discourse [1,2]. Like earlier research, environmental and financial concepts constitute the core structure of public discussions. However, our findings also reveal a stronger association between green finance and digital finance topics, including cryptocurrency, trading, and financial technology. This differs from much of the existing literature, which has primarily focused on institutional, regulatory, and market perspectives of green finance. The prominence of digital finance themes in our dataset suggests that contemporary online discussions increasingly integrate sustainability narratives with emerging financial technologies, thereby extending previous research on green finance discourse.

From a sentiment perspective, the predominance of positive and neutral discourse suggests a generally favorable pattern of sentiment expressed by X users regarding green

finance within the digital public sphere. This finding supports earlier studies indicating that sustainability-related financial practices are becoming normalized through regulatory integration and market adoption [41]. However, the persistence of negative sentiment associated with regulatory complexity, transition costs, and market uncertainty indicates that concerns regarding the implementation and economic implications of green finance continue to be present within public discourse. Rather than reflecting opposition to sustainability goals, these concerns point to tensions associated with the implementation and economic implications of green finance, reinforcing the notion that sustainability transitions are inherently complex and uneven processes.

Importantly, the relational patterns revealed through keyword co-occurrence networks and density visualization demonstrate that green finance discourse is highly interconnected and structurally complex. Core concepts such as climate, energy, and investment form a dense central cluster, while peripheral clusters indicate the presence of specialized and evolving thematic domains. The network structure indicates strong conceptual relationships among environmental, financial, and technological themes, demonstrating the interconnected nature of green finance discussions within the digital public sphere [14–16].

From a business strategy perspective, these findings provide several practical implications for financial institutions and sustainability practitioners. First, financial institutions should enhance the transparency of their green finance products and sustainability initiatives to strengthen public trust and address concerns related to greenwashing. Second, because digital platforms play a central role in shaping public discourse, organizations should incorporate digital communication strategies into their sustainability efforts and actively engage with stakeholders through online channels. Third, sustainability practitioners can use NLP-based text analytics to monitor perceptions expressed by X users, identify emerging concerns, and evaluate the effectiveness of sustainability communication. These practices can support more informed decision-making and improve stakeholder engagement.

From a policy perspective, the findings suggest several practical recommendations. Policymakers should consider public discourse when designing and evaluating green finance policies, as social media discussions provide valuable insights into public expectations and emerging concerns. In addition, clearer reporting standards and transparency requirements may help reduce uncertainty and mitigate concerns related to greenwashing. Incorporating insights derived from large-scale NLP analyses into policy evaluation processes may also contribute to more evidence-based and socially responsive green finance strategies [5].

The findings indicate that green finance discourse should be interpreted as a multidimensional communication process rather than as a collection of isolated textual features. Within the proposed conceptual interpretative framework, lexical prominence identified by TF-IDF forms the vocabulary of public discussions, thematic organization derived from LDA provides narrative coherence, sentiment orientation identified by FinBERT reflects evaluative framing, and conceptual connectivity revealed through keyword co-occurrence networks explains how narratives become structurally reinforced. Together, these complementary analytical dimensions provide a coherent interpretation of how green finance discourse is constructed, communicated, and maintained within the digital public sphere.

This study contributes to the literature by demonstrating the practical value of integrating multiple established NLP techniques within a single analytical framework for examining green finance discourse. Rather than introducing a new methodological approach, the study illustrates how complementary methods can be combined to capture thematic, emotional, and relational dimensions of sustainability-related online discussions. This integrated perspective provides a richer interpretation of digital discourse and may serve as a useful analytical framework for future studies examining sustainability communication across different contexts and platforms. Accordingly, the proposed analytical

framework follows an information retrieval perspective in which lexical weighting, contextual language modelling, probabilistic topic discovery, and relational network analysis jointly contribute to a more comprehensive interpretation of digital discourse than any individual method alone.

Because this study does not propose a novel NLP architecture or prediction model, benchmark comparisons with alternative algorithms and module ablation experiments were considered beyond its intended scope. Instead, the proposed framework emphasizes analytical complementarity by combining established methods that address different dimensions of digital discourse. Future research may systematically compare alternative sentiment analysis models (e.g., VADER and general-purpose BERT), short-text topic modelling approaches (e.g., BERTopic and Biterm Topic Model), and alternative network analysis techniques while conducting module ablation experiments to quantify the individual contribution of each analytical component.

6. Conclusions

Green finance has become an important component of sustainability transitions, attracting increasing attention across digital communication environments. This study investigated green finance discourse in the digital public sphere using a multi-method NLP framework integrating text mining, FinBERT sentiment analysis, LDA topic modelling, and keyword co-occurrence network analysis. The findings demonstrate that green finance discussions are characterized by distinct thematic structures, sentiment patterns, and conceptual relationships that reflect the interaction of sustainability, finance, and technology-related topics.

The results indicate that green finance discourse encompasses both sustainability-oriented and investment-oriented themes. While climate action, sustainability transition, and renewable energy remain dominant topics, discussions related to digital finance, investment activities, and emerging financial technologies also occupy a substantial share of the discourse. This demonstrates the multidimensional nature of green finance discussions within social media environments.

From a managerial perspective, the findings emphasize the importance of monitoring online discussions and sentiment dynamics related to green finance on the platform X. Organizations can benefit from social media analytics and NLP-based monitoring systems to better understand stakeholder expectations, identify emerging concerns, and support more effective sustainability communication strategies.

From a policy perspective, the findings demonstrate the value of utilizing large-scale social media analytics to support the design and evaluation of green finance policies and sustainability-related financial frameworks. Policymakers can benefit from insights into discussions expressed on the platform X, enabling the development of more adaptive and evidence-informed sustainability policies. This is particularly important in ensuring that green finance initiatives are not only economically viable but also socially legitimate and broadly supported.

This study contributes to the literature by applying an integrated NLP framework to the analysis of green finance discourse in the digital public sphere. The findings highlight the importance of understanding online discussions, thematic structures, and sentiment dynamics related to green finance on the platform X. Furthermore, the study demonstrates the usefulness of FinBERT sentiment analysis, LDA topic modelling, and keyword network analysis for examining large-scale sustainability discussions on social media platforms. Future research may extend this approach through cross-platform, longitudinal, and cross-cultural analyses to provide a broader understanding of evolving sustainability-related discourse.

https://doi.org/10.3390/app16147315

This study has several limitations that should be acknowledged. First, the analysis was based on English-language posts collected from the platform X during a single month (January 2026), providing a temporal snapshot rather than a representation of long-term online discourse. Because discussions on social media may change in response to policy developments, market events, and sustainability-related agendas, the findings should be interpreted within the selected period and platform rather than as representative of broader public opinion. Second, the dataset was retrieved using only the keyword "green finance". Although this term represents the central concept examined in the study, discussions employing related expressions such as "sustainable finance", "ESG investing", "climate finance", "green bonds", or "transition finance" may not have been captured. Furthermore, the dataset may include posts from individual users, organizations, institutional accounts, and promotional content, while no distinction was made between these account types and geographic information was not considered. Finally, although LDA provided an interpretable thematic representation of the corpus and the keyword co-occurrence network offered useful exploratory insights, no formal human validation of topic interpretability was performed. While quantitative network analysis was conducted, future research could incorporate additional graph-theoretical metrics and expert-based validation to further strengthen the robustness and interpretability of the findings.

Author Contributions: Conceptualization, M.K. and M.T.; Methodology, M.K., D.E. and S.P.; Software, M.K.; Validation, M.T. and S.P.; Formal analysis, M.K. and S.P.; Investigation, M.T., D.E., G.M. and R.D.; Resources, D.E. and S.P.; Data curation, M.K.; Writing—original draft, M.K., D.E., S.P., G.M. and R.D.; Writing—review & editing, M.K., M.T., G.M. and R.D.; Project administration, M.T. and S.P.; Funding acquisition, G.M. and R.D. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by the National University of Science and Technology Politehnica Bucharest through the program PubArt.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study were collected from publicly available posts on the platform X. Due to the platform's terms of service and privacy considerations, the complete dataset cannot be publicly shared. Processed data and the analysis code are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ozili, P.K. Green finance research around the world: A review of literature. *Int. J. Green Econ.* **2022**, *16*, 56–75. [CrossRef]
2. Akomea-Frimpong, I.; Kukah, A.S.; Jin, X.; Osei-Kyei, R.; Pariafsai, F. Green finance for green buildings: A systematic review and conceptual foundation. *J. Clean. Prod.* **2022**, *356*, 131869. [CrossRef]
3. Cody, E.M.; Reagan, A.J.; Mitchell, L.; Dodds, P.S.; Danforth, C.M. Climate change sentiment on Twitter: An unsolicited public opinion poll. *PLoS ONE* **2015**, *10*, e0136092. [CrossRef] [PubMed]
4. Pearce, W.; Holmberg, K.; Hellsten, I.; Nerlich, B. Climate change on Twitter: Topics, communities and conversations about the 2013 IPCC Working Group 1 report. *PLoS ONE* **2014**, *9*, e94785. [CrossRef] [PubMed]
5. Ohme, J.; Araujo, T.; Boeschoten, L.; Freelon, D.; Ram, N.; Reeves, B.B.; Robinson, T.N. Digital trace data collection for social media effects research: APIs, data donation, and (screen) tracking. *Commun. Methods Meas.* **2024**, *18*, 124–141.
6. Murshed, B.A.H.; Mallappa, S.; Abawajy, J.; Saif, M.A.N.; Al-Ariki, H.D.E.; Abdulwahab, H.M. Short text topic modelling approaches in the context of big data: Taxonomy, survey, and analysis. *Artif. Intell. Rev.* **2023**, *56*, 5133–5260. [PubMed]
7. Isoaho, K.; Gritsenko, D.; Mäkelä, E. Topic modeling and text analysis for qualitative policy research. *Policy Stud. J.* **2021**, *49*, 300–324.
8. Salton, G.; Buckley, C. Term-weighting approaches in automatic text retrieval. *Inf. Process. Manag.* **1988**, *24*, 513–523. [CrossRef]
9. Jiang, Z.; Gao, B.; He, Y.; Han, Y.; Doyle, P.; Zhu, Q. Text Classification Using Novel Term Weighting Scheme-Based Improved TF-IDF for Internet Media Reports. *Math. Probl. Eng.* **2021**, *2021*, 6619088. [CrossRef]

10. Blei, D.; Ng, A.; Jordan, M. Latent dirichlet allocation. *Adv. Neural Inf. Process. Syst.* **2001**, *14*.. [CrossRef]
11. Liu, J.; Li, K.; Zhu, A.; Hong, B.; Zhao, P.; Dai, S.; Wei, C.; Huang, W.; Su, H. Application of deep learning-based natural language processing in multilingual sentiment analysis. *Mediterr. J. Basic Appl. Sci.* **2024**, *8*, 243–260. [CrossRef]
12. Xing, Z.; Li, L.; Yang, Y.; Zhou, W.; Ma, G.; Chen, S.; Lyu, L.; Li, X. A lightweight model for indoor object detection in unstructured scenes based on joint attention and prior knowledge in the context of home rehabilitation. *J. King Saud. Univ. Comput. Inf. Sci.* **2026**, *38*, 186. [CrossRef]
13. Xing, Z.; Meng, Z.; Zheng, G.; Yang, L.; Guo, X.; Tan, L.; Jiang, Y. Human-computer interactive rehabilitation: A 3D graph deep learning method for non-contact gesture recognition in post-epidemic and aging societies. *Measurement* **2025**, *257*, 118794.
14. Kim, T.; Lee, D.; Lim, H.; Lee, U.-K.; Cho, H.; Cho, K. Exploring research trends and network characteristics in construction automation and robotics based on keyword network analysis. *J. Asian Archit. Build. Eng.* **2021**, *20*, 442–457.
15. Yuan, C.; Li, G.; Kamarthi, S.; Jin, X.; Moghaddam, M. Trends in intelligent manufacturing research: A keyword co-occurrence network based review. *J. Intell. Manuf.* **2022**, *33*, 425–439. [CrossRef]
16. Zhao, Y.; Yin, J.; Zhang, J.; Wu, L. Identifying the driving factors of word co-occurrence: A perspective of semantic relations. *Scientometrics* **2023**, *128*, 6471–6494. [CrossRef]
17. Anbarini, R.; Abdullah, A.; Sjafirah, N.A. Democracy on Social Media: The Analysis of the New Criminal Code Ratification Polemic on Twitter. *J. Kaji. Jurnalisme* **2023**, *7*, 40–56. [CrossRef]
18. Ghosh, A.; Parisian, P.; Yan, Y.; Lin, W. The Intersection of Social Media and Public Health: A Review of Sentiment Analysis, Social Network Analysis, Trend Analysis, and Topic Modeling. In *Intelligent Data Analytics in Media Communication to Improve Public Health Care*; Apple Academic Press: Palm Bay, FL, USA, 2025; pp. 195–242.
19. Xu, Q.A.; Chang, V.; Jayne, C. A systematic review of social media-based sentiment analysis: Emerging trends and challenges. *Decis. Anal. J.* **2022**, *3*, 100073. [CrossRef]
20. Mohammadi, E.; Karami, A. Exploring research trends in big data across disciplines: A text mining analysis. *J. Inf. Sci.* **2022**, *48*, 44–56.
21. Alhazmi, H.N. Text mining in online social networks: A systematic review. *Int. J. Comput. Sci. Netw. Secur.* **2022**, *22*, 396–404.
22. Naveen, V.; Sountharrajan, S. A Lightweight Multimodal Framework With Dilated CNN Audio Modeling and SE-Enhanced MobileNetV2 for Real-Time Sentiment Analysis of Customer Product Reviews. *IEEE Access* **2025**, *13*, 217825–217845. [CrossRef]
23. Santamato, V.; Marengo, A.; Aversano, L. Unveiling the Multifaceted Landscape of Biohacking: Insights From a Machine Learning-Driven Systematic Review. *IEEE Access* **2026**, *14*, 51391–51430. [CrossRef]
24. Rathi, R.N.; Mustafi, A. The importance of Term Weighting in semantic understanding of text: A review of techniques. *Multimed. Tools Appl.* **2023**, *82*, 9761–9783. [PubMed]
25. Guo, J.; Cai, Y.; Fan, Y.; Sun, F.; Zhang, R.; Cheng, X. Semantic models for the first-stage retrieval: A comprehensive review. *ACM Trans. Inf. Syst.* **2022**, *40*, 1–42. [CrossRef]
26. Hung, L.P.; Alias, S. Beyond sentiment analysis: A review of recent trends in text based sentiment analysis and emotion detection. *J. Adv. Comput. Intell. Intell. Inform.* **2023**, *27*, 84–95. [CrossRef]
27. Santos, M.L.B.d. The "so-called" UGC: An updated definition of user-generated content in the age of social media. *Online Inf. Rev.* **2022**, *46*, 95–113.
28. Qi, R. Financial news sentiment analysis and market sentiment prediction based on large language models. *J. Comput. Signal Syst. Res.* **2025**, *2*, 22–31.
29. Kong, Y.; Nie, Y.; Dong, X.; Mulvey, J.M.; Poor, H.V.; Wen, Q.; Zohren, S. Large language models for financial and investment management: Applications and benchmarks. *J. Portf. Manag.* **2024**, *51*, 162–210. [CrossRef]
30. Rostam, Z.R.K.; Kertész, G. Advances in pre-trained language models for domain-specific text classification: A systematic review. *ACM Trans. Intell. Syst. Technol.* **2025**, *16*, 1–41. [CrossRef]
31. Choi, P.M.S.; Huang, S.H.; Wang, Q. *Large Language Models in Finance: An Overview*; Finance and Large Language Models; Springer: Singapore, 2025; pp. 1–26.
32. Tong, X.; Willcock, I.; Sun, Y. Unraveling player's insights: A comparative analysis of topic modeling techniques on game reviews and video game developers' perspectives. *IEEE Trans. Games* **2024**, *17*, 167–180.
33. Chang, I.-C.; Yu, T.-K.; Chang, Y.-J.; Yu, T.-Y. Applying text mining, clustering analysis, and latent dirichlet allocation techniques for topic classification of environmental education journals. *Sustainability* **2021**, *13*, 10856. [CrossRef]
34. Farkhod, A.; Abdusalomov, A.; Makhmudov, F.; Cho, Y.I. LDA-based topic modeling sentiment analysis using topic/document/sentence (TDS) model. *Appl. Sci.* **2021**, *11*, 11091. [CrossRef]
35. Du, B.-X.; Liu, G.-Y. Topic analysis in lda based on keywords selection. *J. Comput.* **2021**, *32*, 1–12. [CrossRef]
36. Negara, E.S.; Triadi, D. Topic modeling using latent dirichlet allocation (LDA) on twitter data with Indonesia keyword. *Bull. Soc. Inform. Theory Appl.* **2021**, *5*, 124–132.
37. Topaloğlu, D.; Polat, O. Trends and methods in intensive care unit (ICU) research using machine learning: Latent dirichlet allocation (LDA)-based thematic literature review. *BMC Med. Inform. Decis. Mak.* **2025**, *25*, 282. [CrossRef] [PubMed]

38. Zheng, X.; Pan, L.; Lu, B.; Jin, Y. Theme mining and quantitative evaluation of library policies: Based on LDA model. *J. Comput. Methods Sci. Eng.* **2022**, *22*, 2227–2241. [CrossRef]
39. Yin, B.; Yuan, C.-H. Detecting latent topics and trends in blended learning using LDA topic modeling. *Educ. Inf. Technol.* **2022**, *27*, 12689–12712. [CrossRef]
40. Kamaruzzaman, K.; Salleh, S.A.; Pardi, F.; Abdullah, M.F.; Foronda, V.; Bergonio, E.L.; Rahmawaty, R. Review of environmental monitoring in freshwater lakes using geospatial techniques (remote sensing and GIS). *Geocarto Int.* **2025**, *40*, 2448978. [CrossRef]
41. Ahlström, H.; Monciardini, D. The regulatory dynamics of sustainable finance: Paradoxical success and limitations of EU reforms. *J. Bus. Ethics* **2022**, *177*, 193–212. [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.