

Article

The Structural Evolution of Corporate Sustainability Disclosure: Evidence from China's "Dual Carbon" Goal

Yuqi Xie, Eunyoung Kim *

School of Knowledge Science, Japan Advanced Institute of Science and Technology,
1 Chome-1 Asahidai, Nomi, Ishikawa 923-1211, Japan; s2320003@jaist.ac.jp (YX)

* Correspondence: Eunyoung Kim, Email: kim@jaist.ac.jp.

ABSTRACT

International efforts to address climate change have emphasized the necessity for achieving net-zero emissions by 2050. In alignment with these efforts, China introduced the "Dual Carbon" goal as a key climate and energy policy. Since companies are the primary actors in implementing this policy, this study examines the structural evolution of sustainability reporting of Chinese-listed companies under the "Dual Carbon" goal, with a focus on changes in keyword emphasis, thematic structure, and keyword centrality. Drawing on a sample of 7860 sustainability disclosure (CSR, ESG, SD, and ENV) from 2019 to 2023, this study adopted word frequency analysis, LDA topic modelling, and co-occurrence network analysis. The main results are a) "Dual Carbon" related keywords saw a statistically significant increase after 2020; b) the thematic lineage Climate & Emission Strategy became the fastest growing and dominant disclosure theme; and c) in 2023 co-occurrence network "Dual Carbon" exhibited high semantic centrality and formed a super hub; "Digitalization" and "Governance" functioned as connector to link separate narrative clusters. These results suggest a substantial shift of narrative priority and logic toward climate and emission reduction strategies in Chinese-listed companies' sustainability disclosure within the institutional context of the "Dual Carbon" goal.

Open Access

Received: 24 Dec 2025

Accepted: 06 Feb 2026

Published: 09 Feb 2026

Copyright © 2026 by the author. Licensee Hapres, London, United Kingdom. This is an open access article distributed under the terms and conditions of Creative Commons Attribution 4.0 International License.

KEYWORDS: sustainability disclosure; textual analysis; LDA; co-occurrence Network; Chinese-listed companies; institutional pressure

ABBREVIATIONS

CSR, Corporate Social Responsibility; ESG, Environmental Social and Governance; SD, Sustainable Development; ENV, Environmental Report.

INTRODUCTION

Background

Climate change, particularly global warming resulting from carbon emissions (CO₂), has been identified as an interdisciplinary issue affecting

Economics, Sociology, Politics, Ecology, Engineering, and other fields [1]. International efforts on carbon emission reduction have been made, so that global annual emissions are expected to decrease to 25–30 Gt CO₂ by 2030 and to net-zero by 2050 [2]. More than 132 economies have committed to achieving net-zero emissions by 2050 and many have introduced local climate policies to mitigate carbon emissions, such as the carbon market [3,4]. In the real world, global warming has drawn significant attention from the public and the administration among high-income and emerging economies. For high-income economies such as the EU and Japan, governments and sector associations have set emissions reduction as an essential indicator to evaluate domestic or even multinational firms' performance on dealing with the climate change issue; customers are encouraged to show a preference for decarbonized or low-carbonized products. For emerging economies like China and India, the administration regarded green development as a new direction to improve the quality of economic growth; domestic companies made actions to align with green development both in publicity and operation, hoping to enhance their products and services' competitiveness in international markets.

However, the Glasgow Agreement highlights that emerging economies, such as China and India [5], may experience significant hardship in satisfying these goals. For instance, China's rapid economic growth is accompanied by huge energy consumption, which has thereby substantially increased CO₂ and other greenhouse gas emissions. Specifically, China is the largest energy consumer with the highest proportion of 23.6% in global total consumption, and its energy consumption structure is relatively unreasonable [6]. As the world's largest developing economy and a major energy consumer and carbon emitter, China is expected to take the responsibility of reducing carbon emissions, even at a cost of weakening its robust economic growth. In response, China officially proposed the "Dual Carbon" Goal in 2020 at the 75th session of the United Nations General Assembly on 22 September 2020. In a more specific clarification, the Chinese leadership has added to achieve a CO₂ emission peak by 2030 and to reach carbon neutrality (net zero carbon emissions) by 2060. To date, the "Dual-Carbon" Goal has become one of China's most important national strategies, especially in terms of Climate and energy policies.

Considering that companies and individuals are the primary actors in reacting to and practicing these goals [7], it is worthy to explore their response at an academic level. In this study, Chinese-listed companies were chosen as the research focus. More specifically, since the "Dual-Carbon" goal emphasizes more on climate and energy policy, we perceived that sustainability disclosure, such as corporate sustainable report (CSR) and Environmental-Social-Governance report (ESG) were suitable raw data source. To explain, these reports are typically assumed as professional disclosures that reflect their commitment, preference, and

response to address environmental, energy, and social issues [7,8]. They often contain explicit information on how listed companies describe the alignment of their business visions with environmental and energy policies, such as the “Dual-Carbon” Goal.

Literature Review

Institutional Theory and Corporate Disclosure

Institutional theory is widely utilized across disciplines like Sociology, Economics, Business, and Management to interpret why micro units such as organizations and companies, despite their differences, will eventually behave in remarkably similar ways over time. By definition, institutional theory states that corporate behaviors and disclosure are not only aiming for optimal efficacy or maximum profit, but for recognized legitimacy in a given social and regulatory context [9]. Under such conditions, corporates tend to align themselves with powerful regulations, policy agendas, and prevailing norms adopted by peer organizations. Consequently, this identical mechanism clarifies a causal relationship between seeking legitimacy and exogenous pressures for every participating firm. Specifically, under institutional pressure, firms are encouraged to adopt similar reporting practices to maintain legitimacy.

From a theoretical perspective, institutional pressure originates from three sources: regulatory requirements, professional standards, and peer recognition [9]. Regulatory pressure arises from formal rules and policy frameworks, such as governmental plans, national strategies, and public regulations. Under such pressure, firms are required to comply with official requirements in order to avoid regulatory risks or potential legal consequences. Normative pressure refers to expectations derived from professional standards, industry norms, and widely accepted practices. Firms that fail to follow these norms may be perceived as unprofessional or irresponsible, which can weaken trust and legitimacy among stakeholders and peer organizations [10]. Cognitive pressure reflects the tendency of firms to imitate practices that are perceived as successful, legitimate, or widely adopted by others. Through such imitation, companies seek to enhance their competitiveness and align themselves with prevailing organizational models. Ultimately, corporate disclosure serves as a strategic response to multiple forms of institutional pressure [11], through which firms adjust their reporting practices to maintain legitimacy within their institutional environment.

Institutional theory further emphasizes that institutions are not static, but evolve through policy shifts, regulatory reforms, and changes in widely accepted norms [11,12]. When institutional environments change, firms are expected to adjust their practices to maintain legitimacy. Besides, such an adjustment does not necessarily take the form of abandoning existing practices. Prior studies suggest that organizational responses to institutional change often involve reorganization rather than substitution

[13]. On one hand, new rules and expectations may be layered onto existing practices. On the other hand, earlier routines continue to function but are repositioned within a revised institutional logic. In the context of corporate disclosure, this implies that companies may reorganize the structure, emphasis, and internal coherence of disclosed information instead of merely increasing disclosure volume [14]. Accordingly, disclosure change under institutional transition should be interpreted as a process of reconfiguration, in which certain issues gain prominence, others become peripheral, and the overall narrative logic is reshaped to align with emerging institutional priorities. This perspective highlights the importance of examining whether companies disclose new information and how disclosure structures are reorganized in response to institutional change.

Corporate Sustainability Disclosure & Narrative Evolution

Sustainable disclosure, including ESG, CSR, SD, AND EV reports, is commonly perceived as a neutral technical document intended to objectively present firms' performance in social and environmental responsibilities. However, from an institutional perspective, corporate sustainability reporting also functions as an organizational practice through which firms signal their compliance with existing regulations, norms, and dominant expectations [10,15]. Therefore, these disclosures are assumed to construct narratives that reflect firms' priorities and strategic interests [16,17].

Based on this perspective, corporate sustainability disclosure is expected to be both a static reporting outcome and a dynamic narrative. As external institutional environments change, except for introducing new terms into disclosures, companies also adjust the relative emphasis, internal structure, and semantic connections among disclosure elements. Under such conditions, sustainability reporting evolves through changes in narrative priorities, thematic organization, and the way key concepts are linked to each other. Existing studies suggest that firms tend to reorganize disclosure structures to reflect shifting institutional expectations, especially under major policy or regulatory changes [18,19]. Such reorganization may involve redefining core themes, elevating certain issues to central positions, or integrating previously separate topics into a more coherent narrative framework. Therefore, examining the evolution of disclosure narratives provides a valuable way to understand how companies interpret and respond to institutional change over time.

Research Gaps and Questions

Existing studies on China's "Dual-Carbon" goal have largely focused on policy design, regulatory frameworks, and macro-level governance mechanisms. Numerous studies apply textual analysis techniques to policy documents, such as government reports, official policy texts, and

leadership speech drafts, to examine policy logic and functional mechanisms [20–23]. However, comparatively less attention has been paid to how firms, Chinese listed companies in particular, reflect and express their engagement with the “Dual Carbon” goal through their own sustainable disclosures. As corporate reports constitute an important interface between policy signals and organizational practices, the limited focus on applying textual analysis techniques to firm-level disclosure restricts our understanding of how the “Dual Carbon” goal is associated with changes in corporate sustainable narratives.

At the same time, a growing number of studies have begun to evaluate the micro-level impacts of the “Dual-Carbon” goal using firm-level data from Chinese listed companies. These studies primarily adopt economic or policy evaluation frameworks to analyze changes in financial performance, investment behavior, technological innovation, or emission efficiency [23–25]. Corporate responses are often interpreted as efficiency-driven or incentive-based outcomes. By contrast, relatively little attention has been paid to corporate sustainability disclosure as an institutional response shaped by legitimacy-seeking behavior. The application of institutional theory to explain how the “Dual Carbon” goal is reflected in disclosure priorities and narrative structures remains limited. Therefore, existing research provides insufficient insight into why firms articulate and organize their sustainability narratives except for purely economic interpretations.

To address the above research gaps, this study aims to answer the main research question: How has the sustainability reporting of Chinese listed companies evolved under the “Dual Carbon” goal? Specifically, we decompose this into three sub-questions:

- SRQ1: Did carbon-related keywords display a significant rise following the introduction of the “Dual Carbon” goal?
- SRQ2: How did the main themes of corporate sustainability disclosure evolve?
- SRQ3: How did the central positions of keywords in corporate sustainability disclosure change?

To address these research questions within a coherent theoretical and analytical logic, this study develops a conceptual framework grounded in institutional theory, as illustrated in Figure 1. The framework organized in the logic of Context-Process-Outcome illustrates how institutional pressure associated with the “Dual-Carbon” goal is conceptually linked to changes in corporate sustainability disclosure and how this process is investigated through text-based analytical methods.

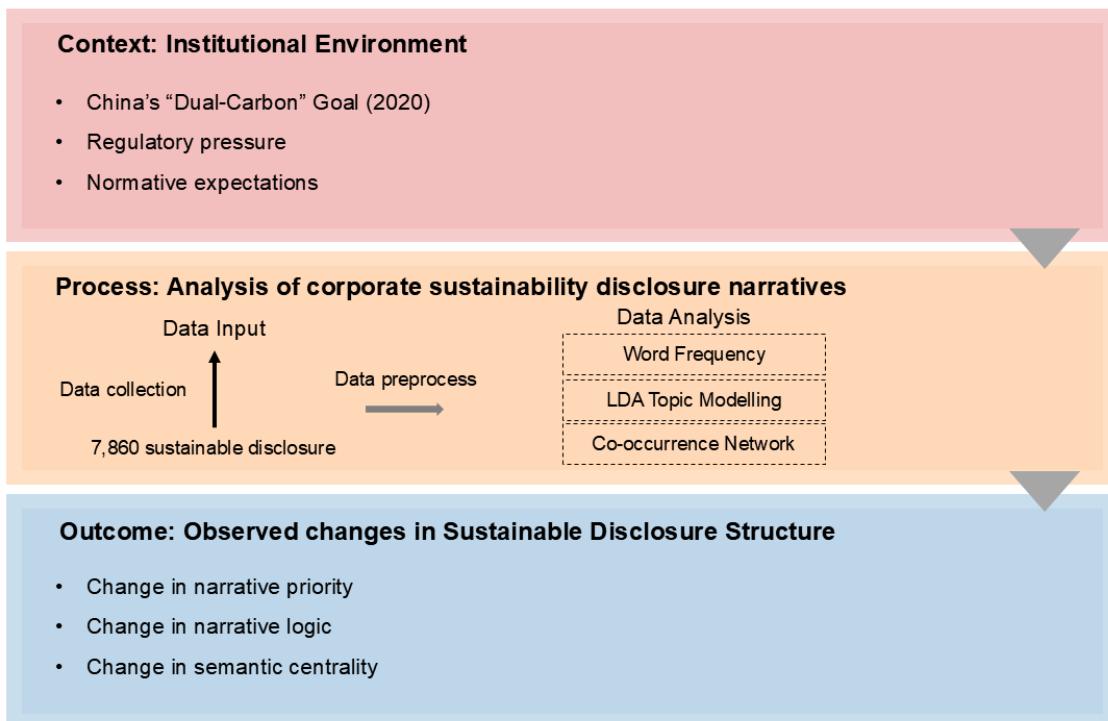


Figure 1. Conceptual Framework.

Contributions

This study is expected to make several modest contributions to existing studies on corporate sustainability disclosure and the "Dual Carbon" goal in China's context.

First, by focusing on sustainable disclosure of Chinese-listed companies, it extends existing research on the "Dual-Carbon" goal from policy texts and macro-level analysis to firm-level disclosure narratives. Second, by adopting an institutional theory vision, this study complements predominantly economics-based explanations and highlights legitimacy-seeking considerations in corporate sustainability disclosure. Third, methodologically, the combined use of word frequency analysis, LDA topic modeling, and co-occurrence network analysis provides a structural view of how corporate sustainability narratives evolve within an institutional context.

MATERIALS AND METHODS

Data Collection

Sample

In recent studies, crawler technology has been widely used to scrape textual data from websites, online forums, and social platforms for research and business analysis. The website, uc.cninfo.com.cn, established in 1994, is qualified by the China Securities Regulatory Commission as the earliest professional website in China to comprehensively disclose announcements and market data for companies listed on the Shenzhen

and Shanghai stock exchanges. This study applied the Python 3.12.4 programming tool to download CSR, ESG, SD and ENV reports of Chinese-listed Companies on uc.cninfo.com.cn. Specifically, targeted reports were identified by matching keywords in the dataset of uc.cninfo.com.cn, which were social responsibility, ESG, environmental report, sustainable development, governance, and management. Targeted reports were collected and aggregated as a raw dataset to cover the latest and most complete information.

This study gathered 1033, 1154, 1512, 1882 and 2279 sustainability-related disclosures of Chinese-listed companies in the period of 2019 to 2023. It should be noted that sustainability-related disclosures are typically published in the year following the reporting period in China. To further explore ownership-based heterogeneity, involved companies were categorized into state-owned enterprises (SOEs) and non-state-owned enterprises (non-SOEs) based on publicly available ownership information disclosed in corporate filings and their profiles. The final sample includes 3839 SOE disclosures and 3850 non-SOE disclosures, while 170 ones could not be matched to an ownership category due to missing or ambiguous ownership information.

This five-year window allows for a robust, panel-data comparison between the pre-policy environment (2019–2020) and the post-policy implementation stage (2021–2023), following the announcement of the Dual Carbon goal in late 2020. A total of 7860 targeted reports represent a significant expansion and more comprehensive market coverage compared to previous studies.

Data Preprocessing

To ensure the high quality and consistency of the text analysis, a stringent preprocessing procedure was adopted, establishing the initial framework of our prior work:

The first step was to convert all the PDF reports into plain text (.txt) format. Redundant elements such as image captions, tables, and standard boilerplate text were removed to prevent methodological noise. Subsequently, an identical step was conducted to address the mixed usage of simplified and traditional Chinese characters in the reports: the *zhconv* library was employed to uniformly convert all text to simplified Chinese. This step was crucial for preventing the repeated counting of core keywords due to character variations and ensuring consistency with the subsequent segmentation tool. Ultimately, the *Jieba* library was employed to process the texts into accurate word segmentation. A custom user dictionary, comprising the 21 core Dual Carbon and sustainability keywords, was added into *Jieba* to ensure these multi-character terms were treated as singular entities (e.g., “碳中和”) during the analysis. Finally, a comprehensive stop-word list was applied to the corpus to remove common function words before proceeding with the LDA and Co-occurrence analyses.

Data Analysis

Word Frequency Analysis & Significance Test

To quantify the change in Chinese-listed companies' focus on the lexical level, we first selected a matrix of 21 policy-relevant keywords derived from the State Council's "Action Plan for Carbon Dioxide Peaking Before 2030" and the "1+N" policy framework, covering key dimensions such as policy directives (e.g., Dual Carbon, Carbon Neutrality), action indicators (e.g., Carbon Emissions, Emission Reduction), and generalized sustainability terms (e.g., Green Development). The full of 21 keywords are: "Dual Carbon (双碳)", "Carbon Neutrality (碳中和)", "Carbon Peaking (碳达峰)", "Carbon Tariff(碳关税)", "CBAM", "Carbon Border Adjustment Mechanism(碳边境调节机制)", "Emission(排放)", "Carbon Emissions(碳排放)", "Emission Volume(排放量)", "Emission Reduction (减排)", "Environmental Policy(环境政策)", "Policy (政策)", "National Policy(国家政策)," "National Strategy(国家战略)", "Environmental Governance(环境治理)", "Environmental Protection (环境保护)", "Environmentally Friendly (环保)", "Green (绿色)", "Green Development (绿色发展)", "Sustainable Development(可持续发展)", and "Developmental Goal(展目标)".

Our primary metric, Document Coverage, is defined as the percentage of total reports in a given year that mention a specific keyword at least once. This metric is chosen over raw word count to effectively normalize for annual variations in sample size and report length. To formally verify the policy's structural impact, a Two-sample Z-test for Proportions was conducted. This test compares the Document Coverage of each keyword in the baseline period (2019) against the post-policy period (2023) to determine if the observed shifts are statistically significant.

To further examine ownership-based heterogeneity, the same document coverage measures were calculated separately for SOEs and non-SOEs. This subsample analysis applies an identical keyword dictionary and measurement procedure, allowing for a descriptive comparison of disclosure patterns across ownership types.

Latent Dirichlet Allocation (LDA) Topic Modeling

To further understand the embedded focus beyond single keywords and capture the overall thematic structure of the corporate sustainability disclosure, we employed the LDA topic modeling technique. In a topic model, an unsupervised algorithm is applied to cluster the latent semantic structure of the set of documents and then sort out hidden topics [24]. LDA is a typical topic model that calculates probability distribution and then assigns it to the topics of each document [25]. Therefore, generally, LDA is employed to identify abstract themes or topics from a collection of documents.

In this study, the LDA analysis was performed on the entire five-year corpus to achieve two goals: (1) discover the latent topics comprising corporate sustainable discourse, and (2) track the annual topic prevalence,

which is the proportion of words in the corpus assigned to each topic, to visualize the evolution of the thematic structure. The optimal number of topics (K) was determined using the coherence score metric, ensuring a balance between interpretability and statistical fit.

Topic Prevalence and Trend Analysis Beyond identifying the content of each theme, we calculated the Topic Prevalence for each lineage across the five years. Topic prevalence is defined as the expected proportion of a document (or the entire annual corpus) dedicated to a specific topic. By aggregating the prevalence of annual topics belonging to the same semantic lineage, we quantified the shifting "narrative gravity" of different sustainability dimensions. This allows us to systematically examine changes in the relative prominence of different sustainability narrative dimensions in corporate disclosure.

Co-occurrence Network Analysis

As the last step of analysis, we utilized co-occurrence network analysis [26] to investigate changes in the relational structure of the corporate narrative from the perspective of sustainability reporting. This method examines the relational structure between the 19 core keywords concluded from Word Frequency and LDA analyses, where a link between two keywords is established if they appear within the same sentence or paragraph.

The network analysis was performed separately for the pre-policy baseline year (2019) and the final year (2023). To quantifiably describe the structural features of the co-occurrence networks, three indicators were applied. Network Density indicates the proportion of realized co-occurrence links among all possible links within the selected keyword set [27]. Degree centrality presents the number of times a given keyword co-occurs with other selected keywords, which illustrates the breadth of its semantic connections [28]. Weighted degree measures the total frequency of co-occurrence between a keyword and all its connected keywords, which identifies the intensity of its involvement in the overall narrative. Because the number of sample disclosures differs between 2019 and 2023 (1033 vs. 2279), the weighted degree was normalized to ensure comparability across years. Specifically, the raw weighted degree of each keyword was divided by the total number of reports in the corresponding year and multiplied by 100. Finally, the topological changes in semantic associations were visualized to facilitate comparison in the two co-occurrence networks.

(AI-assisted tools were used only for language editing and proofreading and did not contribute to the research design, data analysis, or interpretation of results).

RESULTS

Changes in Keyword Coverage

Comprehensive-Level Changes in Keyword Coverage

To describe changes in keyword coverage over time, we analyze the document coverage of 21 targeted keywords from 2019 to 2023. The statistical significance of these changes is assessed using a two-sample Z-test for Proportions (see Table 1), comparing the pre-policy baseline (2019) with the latest post-policy period (2023).

Table 1. Keyword Document Coverage by year.

Keywords	2019	2020	2021	2022	2023	z-test for 2023 to 2019
Dual Carbon	0.00%	1.39%	44.11%	60.15%	65.69%	$p < 0.001$
Carbon Neutrality	0.68%	24.26%	52.18%	57.86%	59.59%	$p < 0.001$
Carbon Peaking	0.00%	16.20%	46.10%	50.69%	51.38%	$p < 0.001$
Carbon Tariff	0.00%	0.00%	0.13%	0.69%	2.06%	$p < 0.001$
CBAM/Carbon Border Adjustment Mechanism	0.00%	0.00%	0.00%	0.37%	0.88%	$p < 0.002$
Emission	69.60%	72.18%	76.39%	83.00%	86.53%	$p < 0.001$
Carbon Emissions	16.65%	23.57%	43.52%	54.20%	65.99%	$p < 0.001$
Emission Volume	32.43%	34.40%	44.25%	53.29%	63.45%	$p < 0.001$
Emission Reduction	72.22%	74.52%	78.77%	83.21%	85.56%	$p < 0.001$
Environmental Policy	3.00%	2.51%	3.17%	4.25%	4.43%	$p = 0.051$
Policy	87.51%	88.82%	89.75%	91.71%	93.07%	$p < 0.001$
National Policy	15.00%	18.54%	22.16%	24.87%	30.01%	$p < 0.001$
National Strategy	18.01%	20.10%	25.20%	29.01%	36.02%	$p < 0.001$
Environmental Governance	15.00%	16.46%	17.33%	20.78%	23.69%	$p < 0.001$
Environmental Protection	85.67%	86.14%	87.76%	89.85%	92.72%	$p < 0.001$
Environmentally Friendly	93.71%	92.63%	93.45%	96.12%	95.61%	$p = 0.020$
Green	85.58%	85.79%	89.42%	93.15%	93.86%	$p < 0.001$
Green Development	49.66%	53.64%	62.83%	69.98%	77.36%	$p < 0.001$
Sustainable Development	91.38%	91.59%	91.87%	94.74%	95.30%	$p < 0.001$
Developmental Goal	24.10%	27.90%	28.37%	30.77%	37.17%	$p < 0.001$

A noticeable growth is observed in keywords directly related to the national “Dual Carbon” strategy. The term “Dual carbon” exhibited a document coverage of 0.00% in 2019 but surged to 65.69% in 2023. Similarly, “Carbon Neutrality” and “Carbon Peaking” experienced rapid increases, reaching 59.59% and 51.38%, respectively. The Z-test results ($p < 0.001$) confirm that these changes are statistically significant, which indicates a substantial expansion in the presence of policy-related climate terms in corporate sustainability disclosures.

Apart from these policy-specific terms, the results also show prominence growth of other categories of keywords within corporate sustainability disclosures. Broad and long-established terms such as “Environmental Protection” remained consistently high, increasing

moderately from 85.67% to 92.72%. At the same time, action-oriented and measurement-focused terms expanded substantially. For instance, “Carbon Emissions” and “Emission Volume” rose from 16.65% and 32.43% in 2019 to 65.99% and 63.45% in 2023, respectively ($p < 0.001$).

In contrast, universal sustainability concepts such as “Sustainable Development” already exhibited a high baseline coverage in 2019 (91.38%) and continued to increase slightly to 95.30% in 2023 ($p < 0.001$), which indicated a sustained trend. Specifically, these changes in keyword coverage were continually followed by existing sustainability concepts, instead of being replaced by policy-specific terms.

To capture corporate attention toward emerging international climate trade mechanisms, the technical abbreviation “CBAM” and its formal term “Carbon Border Adjustment Mechanism” were jointly examined. While neither term appeared in the 2019 baseline, the combined indicator reached a document coverage of 0.88% in 2023 ($p = 0.002$). Together with the rise of “Carbon Tariff” (0.00% to 2.06%, $p < 0.001$). These results indicate the emergence of explicit references to external carbon border adjustment and carbon tariffs in corporate sustainability disclosures.

Ownership-Based Differences in Disclosure Patterns

To explore organizational heterogeneity in corporate sustainability disclosure under the “Dual-Carbon” goal, we further examine keyword document coverage patterns by ownership type, which distinguishes between SOEs and non-SOEs. Through the same 21 targeted keywords and word frequency analysis, annual keyword document coverage is calculated separately for the two subsamples over the period 2019–2023.

Consistent with the comprehensive trends, both SOEs and non-SOEs exhibited a rapid increase in keywords directly related to the national “Dual Carbon” strategy. However, SOEs displayed a sharper and more substantial increase among these keywords (see Figure 2). From 2019 to 2023, the document coverages of “Dual Carbon,” “Carbon Neutrality,” and “Carbon Peaking” in SOEs sustainable disclosures grew from 0.00%, 0.54%, and 0.00% to 75.12%, 67.19%, and 62.81%, respectively. Over the same period, the corresponding rates among non-SOEs rose to 57.01%, 52.62%, and 41.10%, remaining consistently below the SOEs level.

Focusing on the post-policy implementation stage in 2023, SOEs saw higher document coverage than the overall sample across nearly all targeted keywords, while non-SOEs generally fell below the comprehensive performance. The largest gap of document coverage was observed for “National Strategy”, which was 51.07% of SOEs’ sustainable disclosures, compared with 36.02% in the comprehensive performance and only 21.67% among non-SOEs. In contrast, the document coverages of “Environmental Policy” and “Sustainable Development” experienced higher figures among non-SOEs.

Besides, universal sustainability concepts such as “Environmental Protection”, “Environmentally Friendly”, “Green”, and “Sustainable

Development" were consistently seen with high document coverages throughout the whole study period across the comprehensive performance, SOEs and non-SOEs sustainable disclosures. For these broadly defined sustainability terms, differences in ownership-based disclosure patterns appeared relatively modest.

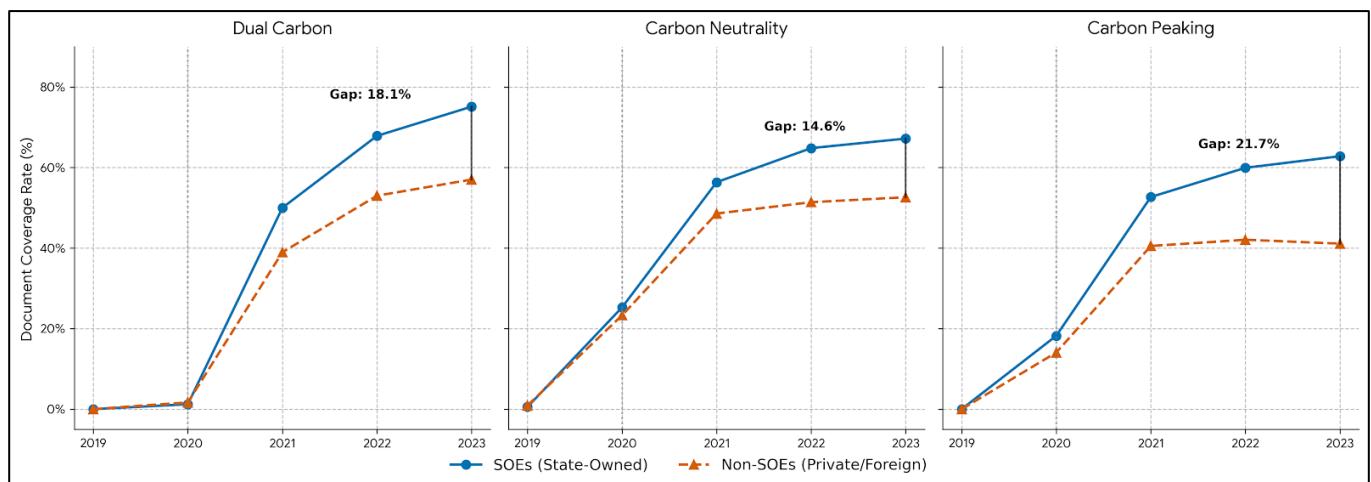


Figure 2. Document Coverage of Selected Keywords: SOEs vs. Non-SOEs.

Evolution of Corporate Sustainability Disclosure Topics

Thematic Identification and Definition

Before topic interpretation, the optimal number of topics for each year was determined using coherence score analysis across a range of K (2 to 11) values (see Figure 3). While lower values of K exhibited slightly higher cross-year stability, K = 7 was selected to balance semantic granularity and interpretability within each year's sustainable disclosures.

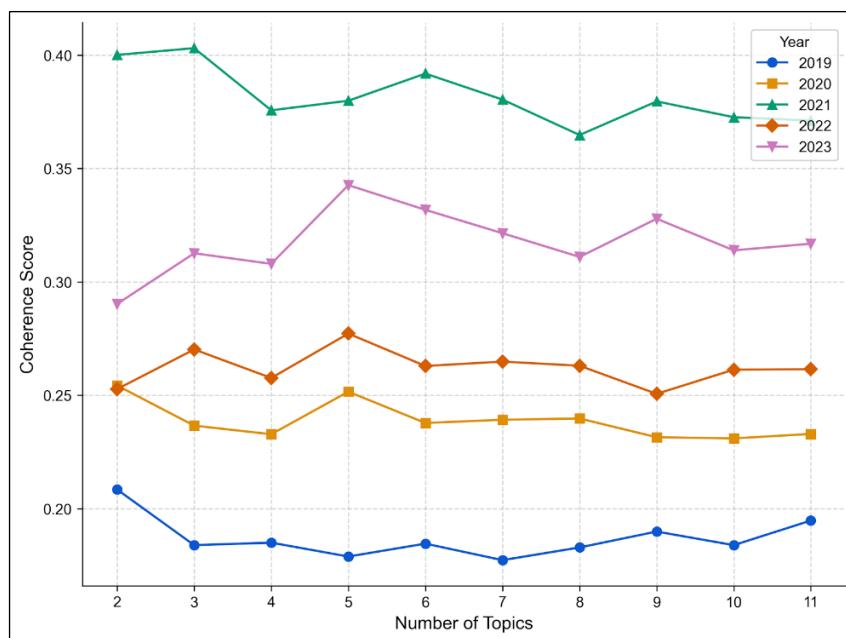


Figure 3. Coherence score from 2019 to 2023.

Adopting the optimal topic number $k = 7$ and individually modeling each year's sustainable disclosures from 2019 to 2023, a total of 35 latent topics (7 topics per year) was generated. To enhance interpretability and enable cross-year comparison, these topics were categorized based on the semantic similarity of their high probability keywords and dominant thematic content. Consequently, five recurrent thematic lineages were identified, which collectively represent recurring thematic patterns in corporate sustainability narratives in China. These five lineages are: Climate & Emission Strategy, Clean Energy & Technical Upgrades, Green Finance & Risk Control, Resource Utilization & Pollution Control, and Medical & Bio-safety.

Table 2 presents the mapping between the annual topic indices and the corresponding thematic lineages. Figure 4 shows the annual prevalence of these lineages in the form of a stacked area chart, visualizing corporate sustainability disclosure structural evolution at the macro-level. This visualization presents changes in the relative prevalence of thematic lineages over time, showing an expansion of climate-oriented themes together with a gradual rebalancing of traditional pollution-related narratives.

It should be noticed that a limited number of topics exhibiting high year-specific heterogeneity (e.g., niche sectoral events) were excluded from these core lineages to ensure analytical focus on the structural shifts associated with the "Dual-Carbon" goal. This classification framework provides the basis for the following detailed analysis of content evolution and prevalence trends.

Table 2. Mapping of Annual Topics to Semantic Lineages.

Lineage	2019	2020	2021	2022	2023
Climate & Emission Strategy	T1: Governance, Waste, GHG	T2: Governance, GHG, Climate Change	T5: Governance, Issues, Info Security	T0: Governance, GHG, Climate Change	T2: Governance, Chemicals, Climate
Clean Energy & Technical Upgrades	T6: Automotive, New Energy, Factory	T1: Power, Gas, Construction	T1: Power, Generation, Nuclear	T1: Wind Power, Generation, Equipment	T3: Wind Power, Energy Storage, Natural Gas
Green Finance & Risk Control	T0: Financial Services, Insurance, AML	T3: Financial Services, Insurance, SMEs	T3: Green Finance, AML, Insurance	T3: Green Finance, Cybersecurity, Data Security	T5: Green Finance, Cybersecurity, Data Security
Resource Utilization & Pollution Control	T4: Wastewater, Exhaust Gas, Mining Area	T5: Mining, Coal, Chemical Industry	T0: New Materials, Coal, Sewage	NA	T1: Mining, Steel, Metals
Medical & Biosafety	T3: Medical Science, Hospital, Medicine	T6: Biological, Medicine, Vaccine	T4: Biological, Drug, Vaccine	T2: Biological, Medicine, Drug	T6: Medicine, Drug, Patient

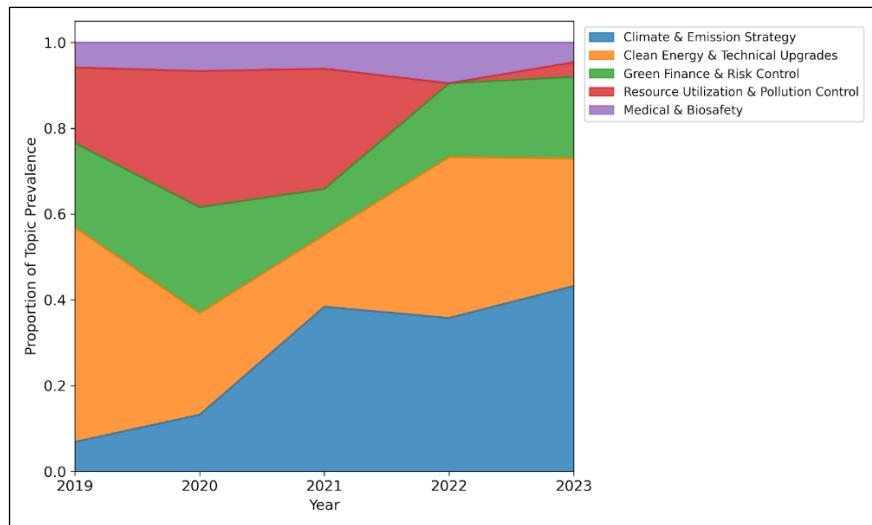


Figure 4. Lineages Prevalence from 2019 to 2023.

Lineage 1: Climate & Emission Strategy

From the perspective of Topic evolution, the “Climate & Emission Strategy” lineage shows a clear structural evolution in disclosure content across the sample period (see Figure 5). Specifically, terms such as “governance (管治)”, “greenhouse gases (温室气体)”, “emission volume (排放量)”, and “labor rights (劳工)” consistently appeared as high-probability keywords, forming a stable institutional backbone of this theme. Conversely, obvious shifts were seen in the composition of the rest of the topic terms. The term “waste (废弃物)”, which was prominent in 2019 and 2020, disappeared from this lineage after 2021. Beginning in 2020, climate-specific terminology such as “climate change (气候变化)” increased markedly in probability. After 2021, the topic further incorporated sector-related and technology-related terms, including “photovoltaics(光伏)”, “electric(电动)”, “automotive (汽车)”, “chips (芯片)”, and “chemicals (化学品)”. Additionally, the term “privacy (隐私)” was observed frequently in 2021 and 2022, indicating corporate concerns related to data became increasingly associated with sustainable disclosure during this period.

From the perspective of Annual prevalence evolution, Lineage 1 presents the largest increase among all identified themes. Its topic annual prevalence rose sharply from 0.0593 in 2019 to 0.3458 in 2023 (see Figure 6), suggesting a growth of nearly six times in narrative weight. A considerable expansion of lineage 1 prevalence was first observed between 2020 and 2021, which rose from 0.0682 to 0.1826, and it aligned with the post-announcement phase of China’s “Dual-Carbon” goal. This notable and rapid growth indicates issues related to climate and emissions switched from a marginal position to a central component of corporate sustainability reporting over the study period.

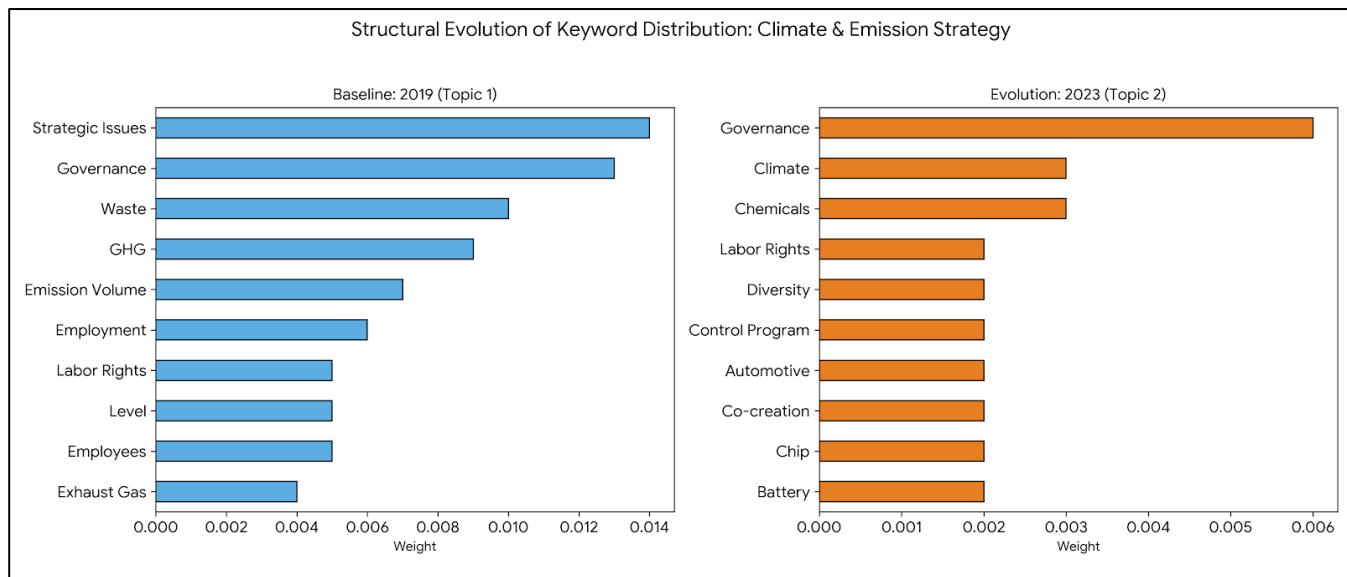


Figure 5. Structural Evolution of Keywords Distribution: Climate & Emission Strategy (2019 vs. 2023).

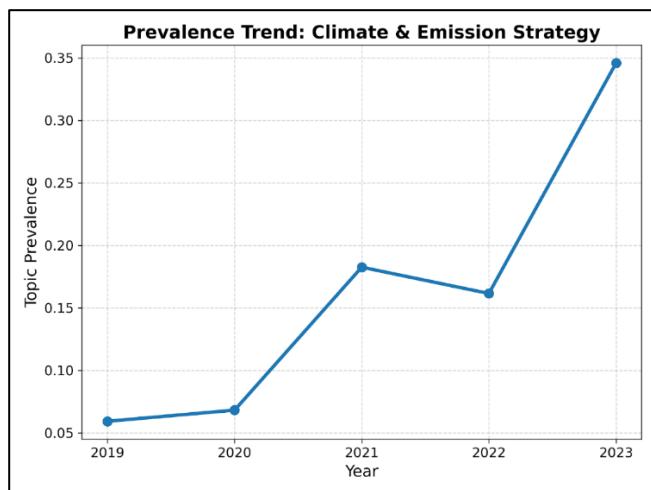


Figure 6. Prevalence Trend: Climate & Emission Strategy (from 2019 to 2023).

Lineage 2: Clean Energy & Technical Upgrades

From the perspective of topic evolution, lineage 2 Clean Energy & Technical Upgrades displays a moderately wide-distributed structure, mainly involving sectors of energy systems, equipment, and transportation (see Figure 7). Compared with other lineages, its high-probability keywords are distributed across multiple sectoral and technical domains rather than concentrated on a single category or concept. Apparent changes in the composition of high probability keywords can be seen over time. In 2019, the term “new energy (新能源)” first appeared as a representative broad concept within this lineage and was subsequently followed by the appearance of more detailed terms in the following period. From 2020 onward, terms related to power systems, such as “power (电力)” and “power generation (发电)”, became consistently prominent. After 2021, more specific energy technologies, including “wind

power (风电)”, “photovoltaics (光伏)”, “nuclear power (核电)”, and “power generation (发电)”, appeared with increased probabilities. Additionally, the term “governance (管治)” emerged as the highest probability keyword in 2021. This was followed in 2023 by the appearance of terms such as “youth (青年)”, “SASAC (国资委)”, “governance (管治)”, and “rectification (整治)”, reflecting that lineage 2 tended to cover terms or concepts related to organizational or administrative respects.

From the perspective of annual prevalence evolution, the “Clean Energy & Technical Upgrades” lineage experienced a significant decline followed by a partial recovery over the study period (see Figure 8). In detail, its topic prevalence decreased from 0.4307 in 2019 to a dramatically lower score of 0.0797 in 2021, but rebounded to 0.2376 in 2023. Overall, this trend reflects a substantial reduction in the relative narrative weight of this lineage compared with its initial prominence in 2019.

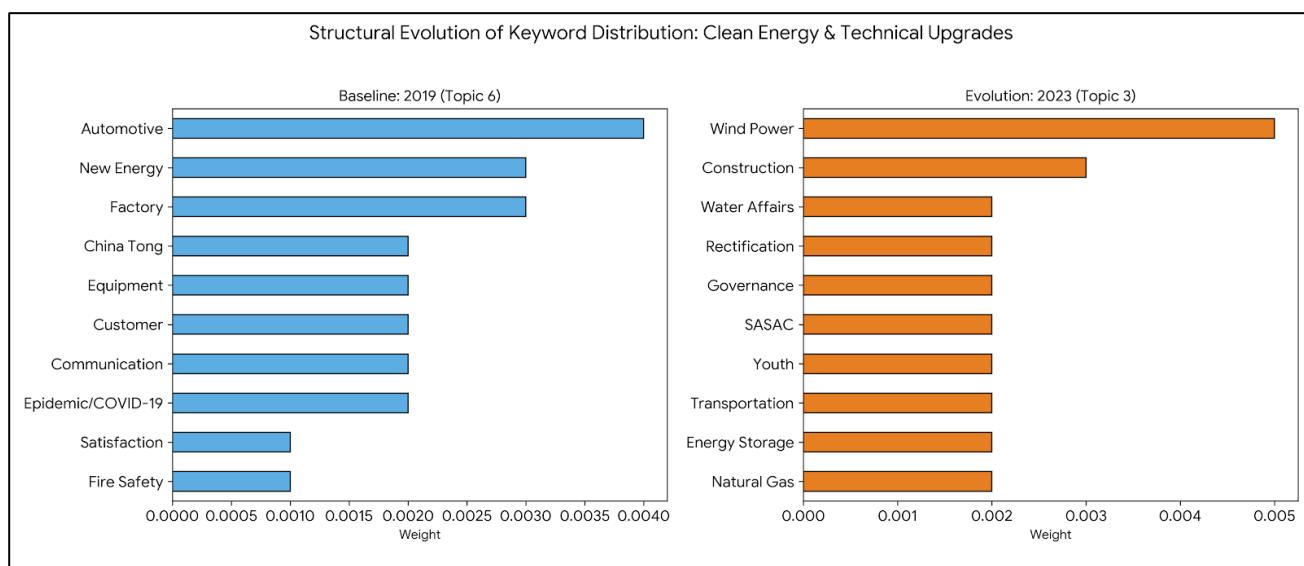


Figure 7. Structural Evolution of Keywords Distribution: Clean Energy & Technical Upgrades (2019 vs. 023).

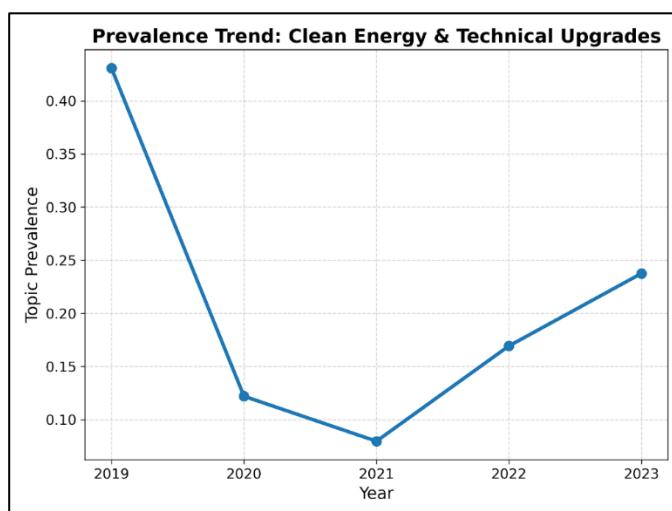


Figure 8. Prevalence Trend: Clean Energy & Technical Upgrades (from 2019 to 2023).

Lineage 3: Green Finance & Risk Control

From the perspective of topic evolution, the “Green Finance & Risk Control” lineage exhibits a high degree of semantic stability over the sample period (see Figure 9). Terms such as “financial services (金融服务),” “green finance (绿色金融),” and “insurance (保险)” consistently appeared as high-probability keywords from 2019 to 2023, forming the core vocabulary of this theme. Several recurrent features further characterize this topic. The term “anti-money laundering (反洗钱)” appeared occasionally in 2019, 2021, and 2023, indicating the financial regulatory dimension embedded in this lineage. In addition, socially oriented financial terms including “small and micro enterprises (小微),” “inclusive finance (普惠),” and “rural areas (乡村)” were consistently seen from 2019 to 2023, demonstrating the continued presence of vocabulary related to corporate social responsibility. This lineage experienced an obvious shift as well in the technological dimension. The generalized term “digitalization (数字化)” was first observed in 2020 and subsequently expanded into a more specific set of keywords related to digital technology, including “cybersecurity (网络安全),” “data security (数据安全),” “software (软件),” “models (模型),” and “information technology (信息技术)” in 2022 and 2023.

From the perspective of annual prevalence evolution, lineage 3 “Green Finance & Risk Control” possesses a fluctuating pattern with an overall recovery trend over the study period (see Figure 10). Its topic prevalence declined from 0.1689 in 2019 to its lowest level of 0.0508 in 2021 and then bound back to 0.1523 in 2023. This U-shaped trajectory indicates that while the relative narrative weight of this lineage temporarily decreased, it regained prominence toward the end of the sample period.

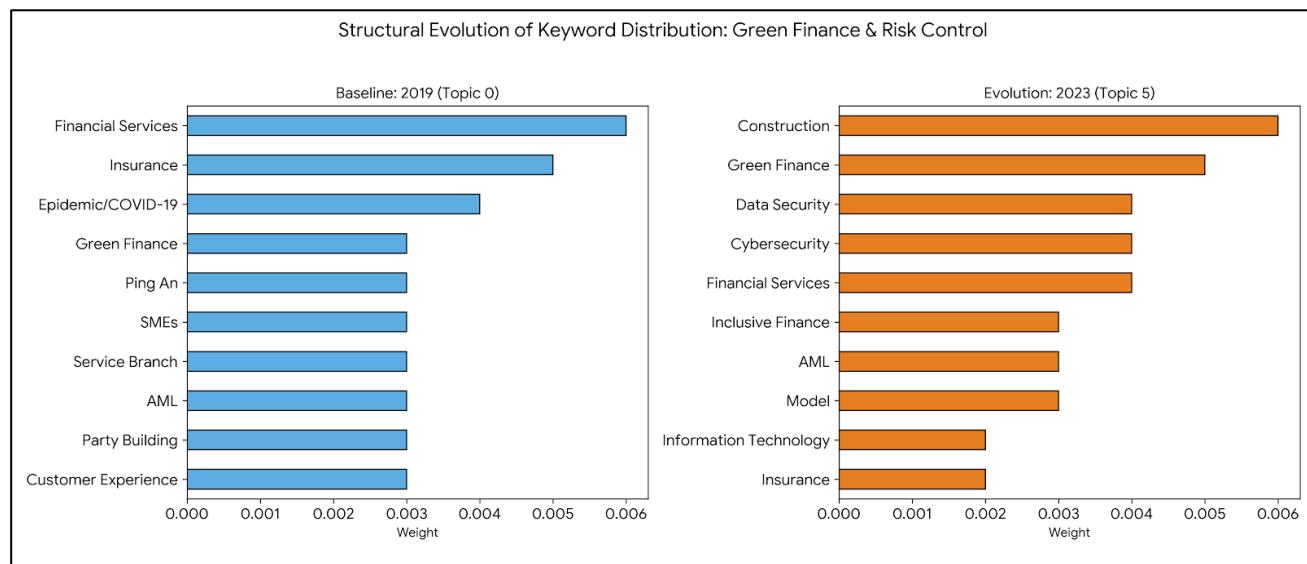


Figure 9. Structural Evolution of Keywords Distribution: Green Finance & Risk Control (2019 vs 2023).

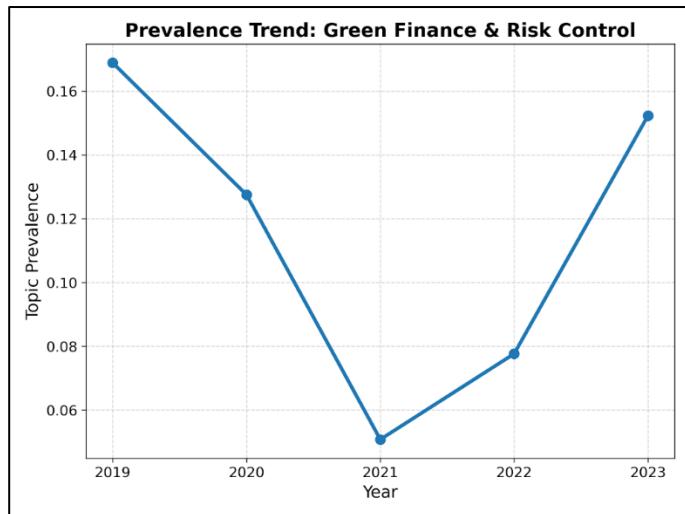


Figure 10. Prevalence Trend: Green Finance & Risk Control (from 2019 to 2023).

Lineage 4: Resource Utilization & Pollution Control

From the perspective of topic evolution, lineage 4 Resource Utilization & Pollution Control presents the highest degree of structural instability among all identified themes (see Figure 11). The topic of 2019 in this lineage was basically featured by terms reacting to pollution, which were, but not limited to, “wastewater (废水),” “exhaust gas (废气),” and “prevention and control (防治).” To further understand it, a high probability keywords distribution like this indicates a focus on end-of-pipe environmental management. Across the study period, keywords related to traditional resource-intensive and high-pollution industries, such as “mining areas (矿区),” “coal (煤炭),” “chemical industry (化工),” “steel (钢铁),” and “metals (金属),” emerged recurrently, which suggest a persistent association of lineage 4 with heavy industrial activities.

However, the internal composition of this lineage did not indicate a stable or cumulative evolution. In 2021, the topic temporarily incorporated terms such as “new materials (新材料),” “new energy (新能源),” “occupational diseases (职业病),” and “hazards (危险),” which reflected a broadened response to pollution-related risks. To be noticed, there were no identified topics possessing sufficient semantic coherence to be consistently mapped into this lineage in 2022. Therefore, it was observed that a dominant theme, resource utilization or pollution control, was missing in that year. By 2023, the remaining terms associated with this lineage mainly focused on industrial materials and components, such as “steel (钢铁),” “metals (金属),” and “components (零部件),” indicating a narrowed and residual form of discourse.

From the perspective of annual prevalence evolution, “Resource Utilization & Pollution Control” lineage shows a significant and continuous decline over the study period (see Figure 12). Its topic prevalence decreased sharply from 0.1515 in 2019 to 0.0272 in 2023, which was the most substantial reduction among all five lineages. Despite minor

fluctuations in ongoing years, the overall trend indicates a rapid marginalization of this theme within corporate sustainability disclosures.

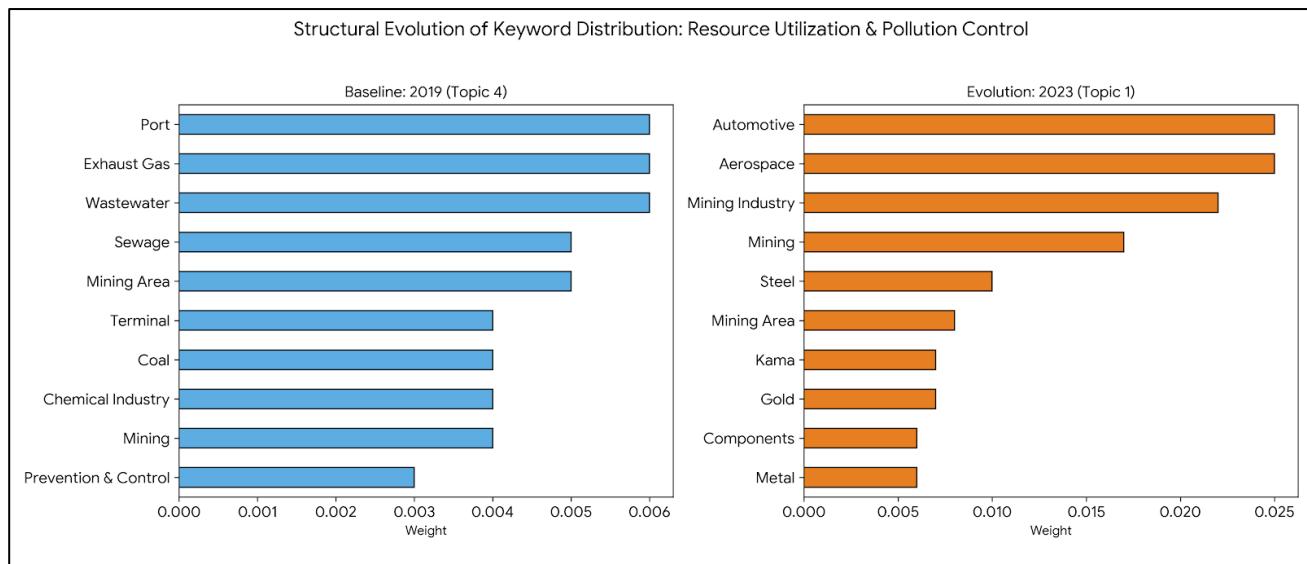


Figure 11. Structural Evolution of Keywords Distribution: Resource Utilization & Pollution Control (2019 vs. 2023).

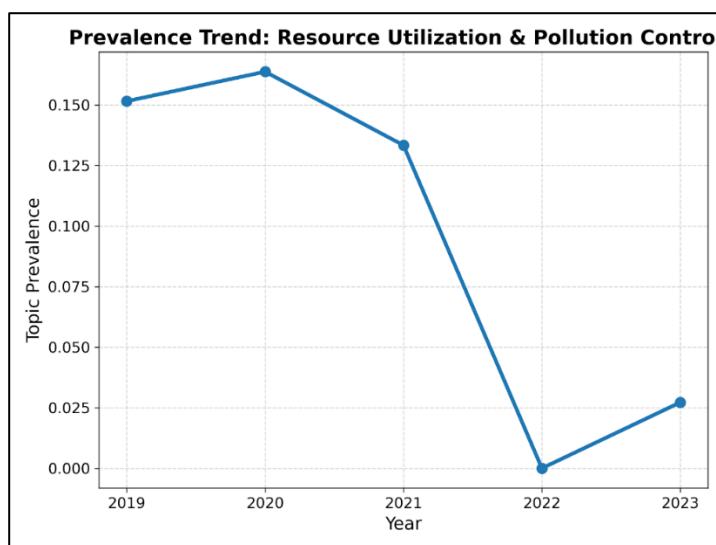


Figure 12. Prevalence Trend: Resource Utilization & Pollution Control (from 2019 to 2023).

Lineage 5: Medical & Bio-safety

From the perspective of topic evolution, the “Medical & Bio-safety” lineage demonstrates a high level of semantic stability throughout the sample period (see Figure 13). Across all five years, the topic was consistently taken up by sectoral-specific terms including “medicine (医学),” “pharmaceuticals (医药),” “hospitals (医院),” “drugs (药品),” and “patients (患者),” with no substantial changes in the structural distribution. In contrast to other lineages, this theme barely exhibited obvious internal diversification or lexical replacement. The main changes in this lineage were largely confined to pandemic-related terms within the healthcare

sector. The term “epidemic (疫情)” appeared with relatively high probability in 2019 and 2020, while “vaccine (疫苗)” remained a prominent term from 2019 to 2022. By 2023, although the core medical vocabulary remained stable, pandemic-specific terms no longer appeared among the high-probability keywords, indicating a convergence toward routine medical and healthcare-related disclosures.

From the perspective of annual prevalence evolution, the “Medical & Bio-safety” lineage sees a gradual declining trend over the study period (see Figure 14). Its topic prevalence decreased from 0.0501 in 2019 to 0.0368 in 2023. In spite of minor fluctuations across years, the overall decline suggests a relative reduction in the narrative weight of this theme within corporate sustainability disclosures.

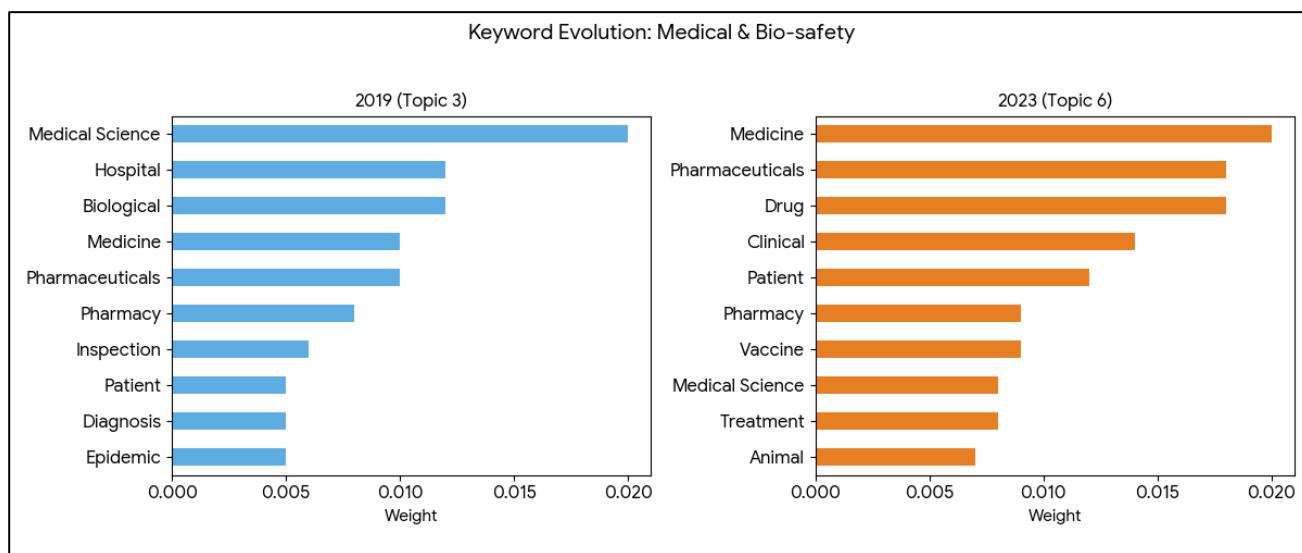


Figure 13. Structural Evolution of Keywords Distribution: Medical & Bio-safety (2019 vs. 2023).

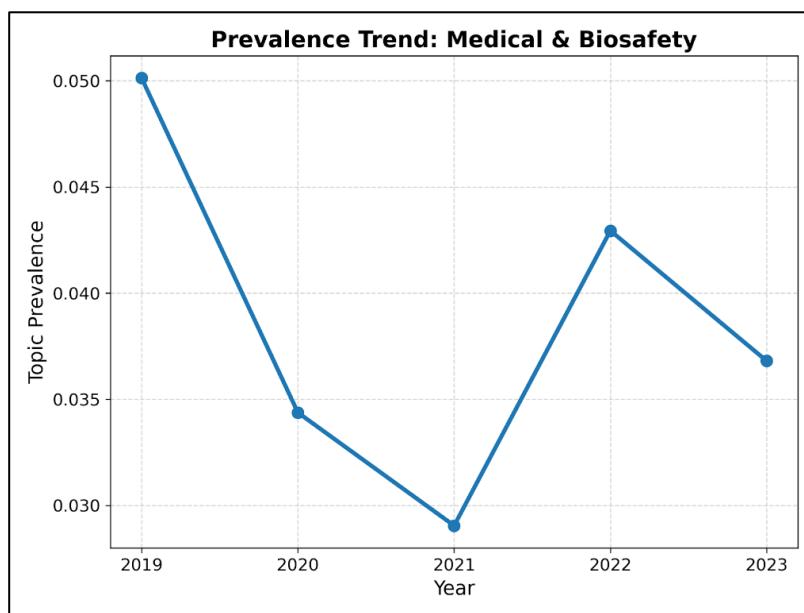


Figure 14. Prevalence Trend: Medical & Bio-safety (from 2019 to 2023).

Co-occurrence Network Structure of Corporate Sustainability Disclosure

A total of 19 keywords were selected as network nodes for the co-occurrence analysis. These keywords were chosen based on three dimensions. First, several keywords were directly related to China's "Dual Carbon" goal framework, including "Dual Carbon," "Carbon Neutrality," "Carbon Peaking," "Carbon Emissions," and "Emission Reduction," which had shown significant changes in document coverage in the previous word frequency analysis. Second, representative keywords were selected from the major thematic lineages identified in the LDA results, covering climate and emissions (e.g., "Climate Change," "Greenhouse Gases"), clean energy technologies (e.g., "New Energy," "Photovoltaics," "Wind Power," "Energy Storage"), governance and finance (e.g., "Governance," "Environmental Governance," "Green Finance," "Insurance"), and digital-related dimensions (e.g., "Digitalization," "Data Security," "Information Technology," "Cybersecurity"). Third, emerging keywords that played a connective role across different themes were secured to ensure consistency with the semantic structure outlined earlier.

By establishing co-occurrence networks for 2019 (baseline) and 2023 (post-policy), this analysis enables a comparative examination of how the relational structure among these core concepts evolved after the implementation of China's "Dual-Carbon" goal.

Comprehensive Structural Changes in the Co-occurrence Network (2019 vs. 2023)

By comparing the overall structural characteristics of the co-occurrence networks in 2019 and 2023, Figure 15 shows changes in the organization of corporate sustainability disclosure over time.

On the whole, a substantially denser and more interconnected network structure in 2023 was observed than in 2019. The network density increased remarkably from 0.3801 in 2019 to 0.5848 in 2023, which suggested a significant rise in the proportion of realized connections among all possible keyword pairs. Correspondingly, the number of valid co-occurrence relationships enlarged from 65 pairs in 2019 to 100 pairs in 2023, reflecting a clear increase in semantic linkage intensity within corporate disclosure narratives.

Besides the increase in density, the overall connectivity of the network also grew. In 2019, only 15 out of the 19 selected keywords appeared in co-occurrence pairs, while two keywords, "Carbon Peaking" and "Dual Carbon," were completely isolated, with their degree centrality equal to 0. In other words, the keywords' isolation meant they did not co-occur with other keywords. This structural feature aligns with the empirical context of the sample period, as these concepts had not yet been formally introduced into policy discourse at that time. In contrast, in 2023, all 19 keywords were fully integrated into the network, with no isolated nodes.

observed. This finding indicates a comprehensive expansion of semantic coverage in the 2023 co-occurrence network.

From a structural perspective, the 2019 co-occurrence network exhibited a relatively loose and uneven configuration because its co-occurrence relationships concentrated around a limited number of keywords and it possessed weak linkages across different semantic areas. Conversely, the 2023 network displayed a more focused and complex configuration, with a pattern of denser interconnections and a higher degree of overall cohesion. This transformation suggests that corporate sustainability disclosure evolved from a relatively fragmented structure toward a more integrated semantic system over the study period.

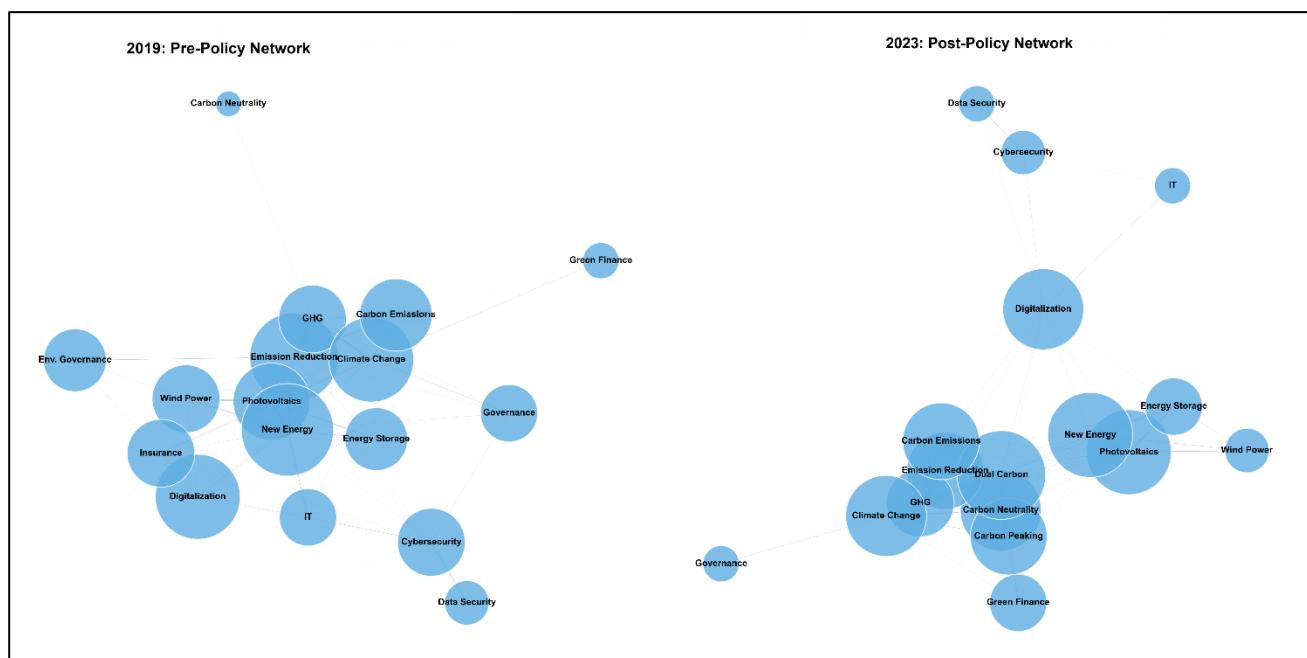


Figure 15. Co-occurrence Network (2019 vs. 2023).

Shifts in Core Keywords

In 2019, the co-occurrence network was characterized by a small number of core keywords. “New Energy” and “Emission Reduction” occupied the most central positions. “New Energy” showed the highest degree centrality (14) and a high normalized weighted degree (12.58), with extensive connections to technology-related keywords such as “Photovoltaics,” “Wind Power,” and “Energy Storage.” Meanwhile, “Emission Reduction” demonstrated the strongest co-occurrence intensity in the network, with a normalized weighted degree of 19.46. It was closely linked to environmental keywords like “Greenhouse Gas,” “Climate Change,” and “Carbon Emissions.” In contrast, policy-focused concepts such as “Carbon Neutrality” and “Dual Carbon” remained isolated during the baseline period.

In 2023, the distribution of core keywords changed significantly. “Dual Carbon” emerged as the most prominent core keyword, displaying both high degree centrality (15) and the highest normalized weighted degree

(94.25) among all selected keywords. It formed strong co-occurrence pairs with multiple keywords related to climate or policy, including “Carbon Neutrality,” “Carbon Peaking,” “Emission Reduction,” “Climate Change,” and “Carbon Emissions.” Meanwhile, “New Energy”, with a degree centrality of 17 and a normalized weighted degree of 88.64, continued to hold a central position. Similarly, “New Energy” was still surrounded by “Photovoltaics (98.95)” and “Energy Storage (55.42), but their roles were increasingly embedded within a broader carbon-focused semantic structure rather than functioning as independent cores.

Overall, the comparison between 2019 and 2023 reveals a clear reconfiguration of core keywords within the co-occurrence network, marked by the emergence of carbon-related policy concepts as dominant semantic anchors.

Emergence of Multiple Hubs in the Network

The co-occurrence network in 2023 also indicated a clear transition from a single-core structure to a multi-hub configuration. In 2019, the network was mainly organized around two major hubs, which were “New Energy” and “Emission Reduction.” However, most other keywords were positioned far from the center with relatively low degree centrality and weighted degree. The overall structure was therefore characterized by a limited number of central nodes and a weak differentiation among secondary keywords.

By contrast, the 2023 network displayed that multiple highly connected hubs were functioning simultaneously (see Figure 16). Aside from “Dual Carbon,” which held the most central position with a degree centrality of 15 and a normalized weighted degree of 94.25, several other keywords also demonstrated increased structural importance. “New Energy” remained a major hub with a degree centrality of 17 and a normalized weighted degree of 88.64 and maintained strong connections with technological keywords such as “Photovoltaics” and “Energy Storage.” At the same time, “Emission Reduction” continued to function as a key node within the carbon-related theme since it presented high connectivity with “Carbon Emissions,” “Greenhouse Gas,” and “Climate Change”.

Overall, the emergence of multiple hubs in the co-occurrence network of 2023 reflects a more complex semantic structure. In this new structure, carbon policy concepts, energy technologies, and governance-related keywords jointly contribute to the organization of corporate sustainability disclosure narratives.

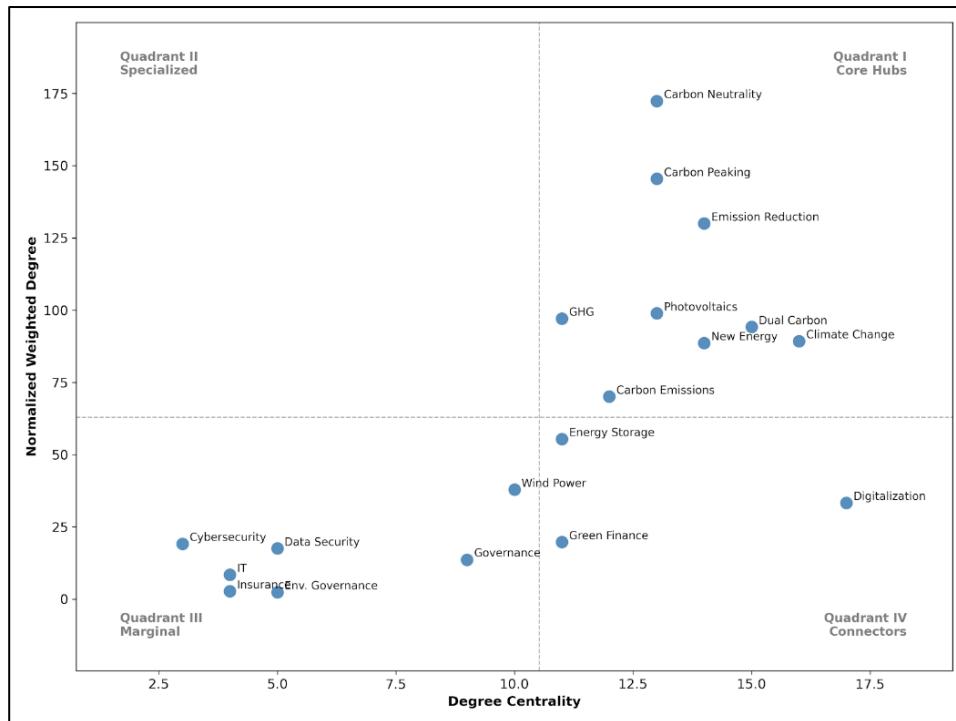


Figure 16. Strategic Matrix of Keyword Centrality and Intensity in 2023.

Connector Keywords across Semantic Clusters

The 2023 co-occurrence network also revealed that connector keywords bridging other distinct semantic clusters were becoming more important. In 2019, cross-cluster connections were relatively limited. Most keywords were embedded within narrowly defined thematic groups, such as energy technologies or environmental ones and only a small number of nodes played an intermediary role between these groups. For example, “Digitalization” showed a moderate degree centrality and was mainly connected with keywords related to technology and finance, but its links with carbon-related keywords remained weak.

In 2023, several keywords had evolved into strong structural connectors, linking carbon policy concepts, energy technologies, and governance-related themes. Among them, “Digitalization” exhibited a notable increase in both degree centrality and normalized weighted degree and formed direct co-occurrence pairs with keywords from multiple clusters, including “Dual Carbon,” “New Energy,” “Governance,” and “Data Security.” This pattern indicates that “Digitalization” no longer functions as an external technological concept but has become embedded across different dimensions of corporate sustainability disclosure.

Similarly, “Governance” appeared as another important connector keyword in 2023. This keyword was seen to expand its connections with “Dual Carbon,” “Climate Change,” “Carbon Emissions,” as well as with digital-related terms such as “Cybersecurity” and “Data Security.” These enlarged connections suggested a growing semantic overlap between carbon strategies and governance-related discourse. To compare,

“Governance” primarily co-occurred with compliance-oriented or environmental management keywords in 2019, while its role in 2023 became more integrative across multiple clusters.

From the perspective of major indicators, these connector keywords did not necessarily display the highest weighted degrees in the network as those in core hubs. However, their relatively high degree centrality across different semantic clusters highlights their structural function as bridges rather than anchors. Overall, the increased prominence of connector keywords in 2023 reflects a shift toward a more interconnected narrative structure. Under this trend, carbon policy, energy transition, digital infrastructure, and governance issues are increasingly discussed in combination rather than in isolation.

DISCUSSION

Disclosure Reorganization under Institutional Change

The results of this study suggest that China’s “Dual-Carbon” goal constituted a major institutional intervention associated with substantial changes in the structure and focus of corporate sustainability disclosure. Through word frequency analysis, LDA topic modeling, and co-occurrence network results, a distinct pattern emerges, which is climate-related concepts shifted from peripheral mentions to central organizing elements of corporate sustainability disclosure, after 2020.

From an institutional perspective, this transformation can be interpreted as a response to strengthened regulatory and normative pressure. Institutional theory argues that firms tend to adjust their organizational practices and disclosure strategies in response to dominant rules, expectations, and policy signals in their external environment [9]. The rapid rise in document coverage of policy-specific terms such as “Dual Carbon,” “Carbon Neutrality,” and “Carbon Peaking” after 2020 provides clear empirical evidence that corporate narratives became closely aligned with the national strategic agenda. Instead of keeping the same pace on general environmental commitment, Chinese-listed companies increasingly adopted the official policy vocabulary, indicating a convergence toward institutionally legitimized discourse.

However, the process of institutional realignment contained substantial heterogeneity across organizational types, particularly with respect to ownership. The ownership-based analysis indicated that SOEs and non-SOEs performed differently in the intensity and manner of their discursive alignment with the “Dual Carbon” goal. While both groups exhibited rising attention to keywords directly related to the “Dual Carbon” goal after 2020, SOEs consistently demonstrated stronger incorporation of policy-framed keywords, particularly those explicitly linked to national strategy. Applying the institutional theory to interpret this divergence, differences in organizational embeddedness can be observed within the state policy system. As units or organizations more affiliated with

governmental authority and political accountability structures, SOEs face stronger normative pressure [9]. Therefore, they indicated a preference for signaling alignment with national strategic goals in sustainability reporting. Conversely, non-SOEs appear to pay more balanced attention among generalized environmental and sustainability narratives, and the “Dual Carbon” goal. This tendency suggests that non-SOEs face alternative pathways of legitimacy construction oriented toward market and societal audiences [14,29].

To be clear, this institutional realignment was not just about keyword frequency or document coverage. The LDA results demonstrate that “Climate & Emission Strategy” became the most prominent thematic line, and the co-occurrence network analysis shows that carbon-related concepts became key semantic anchors linking multiple clusters. This indicates that Chinese-listed companies did not merely add new climate-related terms to their existing sustainable disclosures but reorganized their reporting structure around carbon governance as a central strategic focus. Similar shifts have been seen in previous studies examining how major environmental policies influence corporate reporting priorities by redefining what is considered legitimate and relevant disclosure [14].

At the same time, the lasting presence of broad sustainable concepts such as “Environmental Protection” and “Green Development” indicates that institutional change occurred through specification rather than replacement. Instead of eliminating earlier environmental narratives, Chinese-listed companies embedded carbon-focused disclosure within an expanded sustainability framework. This pattern is consistent with institutional gradualism, where new rules and expectations are integrated into existing practices over time rather than replacing them entirely [12].

As a whole, the findings suggest that the “Dual-Carbon” goal functioned as a powerful institutional signal that redefined the cognitive and normative boundaries of corporate sustainability disclosure. Chinese-listed companies responded by restructuring their narratives to emphasize carbon-related strategies, measurement, and governance, to align disclosure practices with evolving institutional expectations.

From Pollution Control to Climate Strategy

The results suggest a clear shift in the narrative logic of corporate sustainability disclosure from traditional pollution control to a more strategic climate-focused framework. From an institutional perspective, this shift may reflect a redefinition of what constitutes legitimate environmental responsibility under the “Dual-Carbon” goal framework.

Chinese-listed companies face increasing pressure to demonstrate alignment with the “Dual-Carbon” goal, which helps explain the rapid rise of carbon-related and policy-linked terms after 2020. This trend is especially visible in the strong growth of policy-linked terms such as “Dual Carbon,” “Carbon Neutrality,” and “Carbon Peaking”, as well as the

significant increase in carbon-related terms such as “Carbon Emissions,” “Emission Volume,” and “Emission Reduction.”

This shift is also clearly observed in the LDA topic modeling results. While the thematic lineage Climate & Emission Strategy rose sharply and became the most prominent topic in 2023, Resource Utilization & Pollution Control declined continuously and even lost semantic coherence in 2022. This trend implies a reclassification that pollution-related issues were absorbed into carbon-oriented narratives. Similar transitions have been observed in contexts where climate policy reframes environmental responsibility from operational control to strategic orientation [30].

The co-occurrence network analysis supports this interpretation in a further respect. By 2023, carbon-related concepts, “Dual Carbon” in particular, became central semantic anchors and then reorganized corporate sustainability disclosure narratives on climate strategy.

Consequently, the findings suggest that pollution control appears to have lost its central position as the primary organizing logic of environmental disclosure, giving way to climate strategy logic shaped by institutional pressure and national policy signals [9].

Digitalization and Governance as Structural Connectors

The results indicate that digitalization and governance appeared to play important structural parts in how Chinese-listed companies articulated and organized their response to the “Dual-Carbon” goal. From an institutional perspective, policy pressure alone does not automatically translate into an effective organizational response. More importantly, firms must rely on specific tools and structures to make compliance observable, credible, and manageable [11].

The rising prominence of terms related to digitalization provides evidence for this interpretation. In the keyword frequency and co-occurrence analyses, terms such as “Digitalization,” “Data Security,” “Cybersecurity,” and “Information Technology” became increasingly connected with carbon-related concepts after 2020. This pattern suggests that Chinese-listed companies increasingly applied digital technologies as infrastructural supports for carbon measurement, monitoring, and reporting, not as generic efficiency tools. Prior research has emphasized that digital systems play a critical role in translating abstract sustainability goals into measurable and auditable practices, thereby enhancing organizational legitimacy under regulatory scrutiny [31].

Meanwhile, terms related to governance gained structural importance across both the LDA and co-occurrence network results. The enhanced semantic links between “Governance” and carbon-related terms indicate that climate issues were increasingly incorporated into formal governance frameworks, not remaining limited to operational or environmental management parts. Institutional theory suggests that embedding new policy demands into governance structures is a central pathway through which organizations stabilize responses to external pressure and signal

long-term commitment [9]. In this sense, governance acted as a coordinating mechanism that aligned the “Dual Carbon” goal with internal decision-making and accountability systems.

Furthermore, the joint rise of digitalization and governance highlights a complementary relationship. Digital tools provided technical support to generate carbon-related information, while governance structures defined how such information was interpreted, disclosed, and acted upon. Together, these findings suggest that digitalization and governance did not merely accompany the “Dual-Carbon” transition, but were firmly integrated into the sustainability reporting strategies of Chinese-listed companies to respond to institutional pressure in a structured and credible manner.

Theoretical Implications

This study provides three main theoretical implications for research on institutional change and corporate sustainability disclosure.

First, our findings contribute to institutional theory by illustrating how corporate sustainability reporting reflects the institutional change under certain policies, which is through gradual realignment of disclosure narratives instead of abrupt transformation. Specifically, companies appear to incrementally enhance or weaken the relative prominence, connectivity and semantic positioning of sustainability-related topics that are significant within the ongoing policy and institutional context. This pattern aligns with views of institutional change as an evolutionary and interpretive process, in which organizations adjust symbolic practices to maintain legitimacy under shifting normative expectations.

Second, the study advances sustainability disclosure literature by highlighting the importance of narrative structure, beyond keyword frequency or disclosure volume. By examining changes in thematic integration and semantic centrality, our analysis demonstrates that disclosure evolution can occur at the level of narrative logic and relational structure, even in the absence of direct evidence of substantive operational change. This perspective complements existing disclosure studies by emphasizing the structural organization of corporate narratives as a key site of institutional response.

Third, our study broadens the use of institutional theory in China’s unique setting by highlighting differences in how companies symbolically align their disclosure narratives with policy signals across various organizational ownership structures. Instead of assuming all organizations respond the same way to coercive pressure, the observed differences between SOEs and non-SOEs suggest that a firm’s level of institutional embeddedness influences the strength and nature of its narrative alignment. In state-led institutional environments, SOEs are more deeply embedded within the political and administrative systems and tend to show more obvious policy-focused disclosure narratives. Meanwhile, non-SOEs adopt more selective and market-driven

approaches. Importantly, this distinction should be viewed as an interpretive pattern in symbolic disclosures rather than as proof of causal mechanisms. It emphasizes that corporate sustainability disclosure develops in response to external institutional signals and also depends on a firm's position within the political and economic hierarchy.

Limitations and Future Research

Several limitations can be seen in this study. First, the analysis focuses on corporate sustainability disclosure but not on firms' actual performance in undertaking sustainable responsibility. While textual disclosure provides important insights into how firms demonstrate their response to institutional pressure and legitimacy concerns, it does not necessarily reflect substantive changes in operational practices [32]. Second, this study examines ownership-based heterogeneity by identifying different disclosure patterns between SOEs and non-SOEs, but it does not further conduct disaggregated analysis across industries. In the Chinese context, institutional pressure and policy influence may be transmitted unevenly across industries with different carbon intensities and regulatory exposure. The absence of such industry-based disaggregated analysis limits the ability to assess how industrial characteristics interact with institutional signals in shaping corporate disclosure strategies. Third, this study emphasizes China's institutional context and does not conduct cross-country comparative analysis. As disclosure regimes, policy timelines, and linguistic environments differ substantially across countries, the findings should be interpreted within the Chinese context. While cross-national comparison could provide additional insights into how other institutional backgrounds shape corporate sustainability disclosure, it would require a distinct comparative research design and further research.

Future research could address these limitations in several ways. One feasible direction is to conduct an industry-based analysis to explore heterogeneous institutional responses under the "Dual-Carbon" goal. Another option is to integrate disclosure analysis with quantitative environmental performance indicators, allowing for a closer examination of the relationship between symbolic disclosure and substantive action. Finally, as international climate-related regulations such as the EU's Carbon Border Adjustment Mechanism continue to evolve, future studies may further investigate how transnational institutional pressures interact with domestic climate policies to shape corporate disclosure behavior in emerging economies. Our keyword frequency analysis already provides initial evidence that the document coverage of "Carbon Border Adjustment Mechanism" have become statistically significant in recent years.

CONCLUSIONS

This study investigates how the sustainability reporting of Chinese listed companies has evolved under the “Dual Carbon” goal from 2019 to 2023. Applying word frequency analysis, LDA topic modeling, and co-occurrence network analysis, the study provides a structural perspective on changes in narrative priority, thematic organization, and semantic connectivity within corporate sustainability disclosure.

The findings indicate that Chinese-listed companies’ sustainability reporting has experienced substantial structural changes under the “Dual-Carbon” goal framework in three key respects. First, the narrative priority of corporate disclosure expanded from universal sustainability concepts toward a clear emphasis on carbon-related policy objectives. Carbon-related keywords experienced a statistically significant rise after 2020. The increased attention to terms regarding measurement and execution orientation was observed as well. This trend suggests a shift toward more concrete and accountable climate disclosure.

Second, the narrative logic of corporate sustainability disclosure evolved from a traditional pollution-control orientation to a climate strategy and emission reduction framework. LDA Topic modeling results show that Climate & Emission Strategy emerged as the most prominent thematic lineage, while Resource Utilization & Pollution Control declined sharply and even lost semantic coherence in 2022. This pattern indicates that pollution-related concerns were increasingly absorbed into carbon-centered climate narratives but not remaining independent organizing logic.

Third, the structural integration of corporate disclosure was strengthened through the growing role of digitalization and governance-related concepts. Co-occurrence network analysis reveals that “Dual Carbon” and “New Energy” functioned as central hubs, while keywords such as “Digitalization” and “Governance” acted as key connectors linking climate strategy, technological pathways, and managerial frameworks. This suggests that digital infrastructure and governance mechanisms have become closely integrated components for organizing and legitimizing climate-oriented disclosure.

From the perspective of potential contribution, this study provides a deeper understanding of how major climate or energy policies are associated with changes in corporate disclosure practices in emerging economies.

DATA AVAILABILITY

The dataset generated and analyzed in the study can be found at uc.cninfo.com.cn.

AUTHOR CONTRIBUTIONS

YX and EK designed the study. YX performed the data collection. YX analyzed the data. YX wrote the paper with input from all authors.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

FUNDING

This research has no funding.

ACKNOWLEDGMENTS

During manuscript preparation, the author used ChatGPT and Gemini to improve the readability and language quality. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

REFERENCES

1. Scricciu S\$, Barker T, Ackerman F. Pushing the boundaries of climate economics: Critical issues to consider in climate policy analysis. *Ecol Econ.* 2013;85:155-65.
2. IPCC. Global Warming of 1.5°C: IPCC Special Report on Impacts of Global Warming of 1.5°C above Pre-industrial Levels in Context of Strengthening Response to Climate Change, Sustainable Development, and Efforts to Eradicate Poverty. Cambridge (UK):Cambridge University Press, 2022.
3. Green JF, Sterner T, Wagner G. A balance of bottom-up and top-down in linking climate policies. *Nat Clim Chang.* 2014;4(12):1064-7.
4. Abbasi S, Ahmadi Choukolaei H. A systematic review of green supply chain network design literature focusing on carbon policy. *Decis Anal J* 2023;6:100189.
5. Valavanidis A. The COP26 UN Conference Promoted Climate to the Top of the Global Agenda Opportunity for pivotal changes in global climate policy to protect the long-term survival of humanity. Available from: <https://www.researchgate.net/publication/356896691>. Accessed 2025 Oct 20.
6. Wen H, Lee C, Zhou F. Green credit policy, credit allocation efficiency and upgrade of energy-intensive enterprises. *Energy Econ* 2021;94:105099.
7. Carroll AB. A three-dimensional conceptual model of corporate performance. *Acad Manag Rev.* 1979;4(4):497-505.
8. Christensen HB, Hail L, Leuz C. Mandatory CSR and sustainability reporting: Economic analysis and literature review. *Rev Acc Stu.* 2021;26(3):1176-248.

9. DiMaggio PJ, Powell, WW. The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *Am Sociol Rev*. 1983;48(2):147-160.
10. Suchman MC. Managing legitimacy: Strategic and institutional approaches. *Acad Manag Rev*. 1995;20(3):571-610.
11. Scott W. *Institutions and organizations: Ideas, interests, and identities*. Thousand Oaks (CA, USA): SAGE Publications, 2013.
12. Mahoney J, Thelen K. *A theory of gradual institutional change*. Cambridge (UK): Cambridge University Press, 2010.
13. Greenwood R, Raynard M, Kodeih F, Micelotta ER, Lounsbury M. Institutional complexity and organizational responses. *Acad Manag Ann*. 2011;5(1):317-71.
14. Marquis C, Qian C. Corporate social responsibility reporting in China: Symbol or substance? *Organ Sci*. 2014;25(1):127-48.
15. Campbell JL. Why would corporations behave in socially responsible ways? An institutional theory of corporate social responsibility. *Acad Manag Rev*. 2007;32(3):946-67.
16. Deegan C. The legitimising effect of social and environmental disclosures—A theoretical foundation. *Acc Audit Acc J*. 2002;15(3):282-311.
17. Cho CH, Michelon G, Patten DM. Enhancement and obfuscation through the use of graphs in sustainability reports: An international comparison. *Sustain Acc Manag Policy J*. 2012;3(1):74-88.
18. Bebbington J, Larrinaga C, Moneva JM. Corporate social reporting and reputation risk management. *Acc Audit Acc J*. 2008;21(3):337-61.
19. Vaara E, Sonenshein S, Boje D. Narratives as sources of stability and change in organizations: Approaches and directions for future research. *Acad Manag Ann*. 2016;10(1):495-560.
20. Zhu Y, Hu Y, Zhu Y. Can China's energy policies achieve the "dual carbon" goal? A multi-dimensional analysis based on policy text tools. *Environ Dev Sustain*. 2024. doi: 10.1007/s10668-024-05190-4
21. Jiang B, Raza MY. Research on China's renewable energy policies under the dual carbon goals: A political discourse analysis. *Energy Strateg Rev*. 2023;48:101118.
22. Zhang J. Evaluation of fiscal policy with text mining under "dual carbon" target in China. *Heliyon*. 2024;10(13):e33466.
23. Feng T, Appolloni A, Chen J. How does corporate digital transformation affect carbon productivity? Evidence from Chinese listed companies. *Environ Dev Sustain*. 2024;26(12):31425-45.
24. Papadimitriou CH, Tamaki H, Raghavan P, Vempala S. Latent semantic indexing: a probabilistic analysis. *Proceedings of the seventeenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*; 1998 June 1-4; Washington, SA, USA.
25. Blei D, Ng A, Jordan MI. Latent dirichlet allocation. *J Mach Learn Res*. 2003;3:993-1022.
26. Callon M, Courtial JP, Turner WA, Bauin S. From translations to problematic networks: An introduction to co-word analysis. *Soc Sci Inf*. 1983;22(2):191-235.

27. Wasserman S. Social network analysis: Methods and applications. Cambridge (UK): Cambridge University Press, 1994.
28. Freeman LC. Centrality in social networks conceptual clarification. *Soc Netw.* 1978-1979;1(3): 215-39.
29. Li W, Zhang R. Corporate Social Responsibility, Ownership Structure, and Political Interference: Evidence from China. *J Bus Ethics.* 2010;96(4):631-45.
30. Delmas MA, Toffel MW. Organizational responses to environmental demands: Opening the black box. *Strateg Manag J.* 2008;29(10):1027-55.
31. Kellogg KC, Valentine MA, Christin A. Algorithms at work: The new contested terrain of control. *Acad Manag Ann.* 2020;14(1):366-410.
32. Marquis C, Toffel MW, Zhou Y. Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organ Sci.* 2016;27(2):483-504.

How to cite this article:

Xie Y, Kim E. The Structural Evolution of Corporate Sustainability Disclosure: Evidence from China's "Dual Carbon" Goal. *J Sustain Res.* 2026;8(1):e260013. <https://doi.org/10.20900/jsr20260013>.