

Article

From Norm to System: BERTopic Analysis of Re100's Socio-Technical Institutionalization in Korean News

Hey Jeong An ¹ and Chung Joo Chung ^{2,*}

¹ Department of Energy Convergence and Climate Change, Kyungpook National University, Daegu 41566, Republic of Korea; hjpn12@knu.ac.kr

² Department of Communication, Kyungpook National University, Daegu 41566, Republic of Korea

* Correspondence: cjchung@knu.ac.kr

Highlights

Please indicate how your work links to systems science via your contributions to systems practice, theory, and/or methodology.

- This study conceptualizes RE100 as an evolving Socio-Technical System (STS), integrating institutional systems theory and Geels' Multi-Level Perspective (MLP) to model how a global sustainability norm becomes embedded in a path-dependent national energy regime.
- Methodologically, it applies BERTopic-based topic modeling to a large-scale news corpus (21,901 articles) and empirically tests the sequencing of institutionalization stages, offering a transferable analytical matrix for studying socio-technical transitions in other governance systems.

What are the main findings and/or the implications of the main findings?

- Analysis of Korean news media (2019–2025) reveals RE100 evolving through four interpretive clusters—local government-centered construction, industrial-policy formation, market-based corporate institutionalization, and ESG-driven governance—reflecting a partially sequential path of externalization, objectivation, and internalization.
- Findings show organizational/ESG internalization (symbolic commitment) consistently precedes market-mechanism internalization (e.g., PPAs), suggesting policy-makers and firms adopting successor frameworks (e.g., CF100) should anticipate this “symbolic-before-operational” sequencing when designing transition strategies.

Abstract

This study conceptualizes RE100 not merely as a voluntary corporate campaign, but as an evolving socio-technical system that intersects with South Korea's path-dependent energy regime. Applying BERTopic to a comprehensive corpus of 21,901 raw news articles (reduced to 13,507 analytically relevant articles after noise filtering; January 2019–October 2025), we trace the systemic evolution and multi-layered governance of RE100. Grounded in social constructionism and institutional systems theory, the research reveals how the global sustainability norm becomes structurally embedded in South Korea's industrial and economic systems through a partially sequential evolutionary process of externalization, objectivation, and internalization. BERTopic identified 39 analytically meaningful topics, subsequently organized into four interpretive clusters: (1) local government-centered construction, (2) spatial reconfiguration and industrial-policy formation, (3) corporate-led market institutionalization, and (4) ESG-driven corporate governance. Collectively, these clusters demonstrate that RE100 has evolved from a peripheral international initiative into



Academic Editors: Katerina Fotova
Čiković and Dinko Primorac

Received: 23 June 2026

Revised: 3 September 2026

Accepted: 7 September 2026

Published: 10 September 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

a multilayered governance norm embedded across local governance, industrial infrastructure, corporate strategy, and ESG systems. The findings hold theoretical and policy implications for energy-transition governance and climate communication. This progression is recursive rather than strictly linear: corporate-level internalization (Cluster 4) is conceptually linked to, and may interact with, regional and industrial-policy externalization (Clusters 1–2), consistent with a non-linear reading of institutionalization in which corporate site-selection criteria shape local-government infrastructure competition. Cluster 1's prominence reflects local governments' discursive visibility rather than substantive policy authority. Notably, RE100's "normalization" denotes routinized corporate compliance practices, not political consensus on its legitimacy relative to alternatives such as carbon-free 100% or nuclear power.

Keywords: RE100; BERTopic; socio-technical system; social constructionism; energy transition; Korean media; institutionalization

1. Introduction

Climate change—primarily caused by anthropogenic activities, especially the rapid post-industrial escalation of fossil-fuel use and the consequent increase in greenhouse-gas emissions—has emerged as one of the most pressing governance challenges of the twenty-first century [1,2]. The increase in atmospheric CO₂ concentrations amid surging global energy demand has contributed to extreme weather events, sea-level rise, and ecological disruptions [3]. These phenomena threaten global sustainability and have prompted sustained international efforts to reduce CO₂ emissions [4].

In response, the international community has developed an increasingly institutionalized climate-governance regime. The establishment of the United Nations Framework Convention on Climate Change, endorsed by 154 signatories during the 1992 Earth Summit, marked a foundational milestone. The framework codified the goal of stabilizing atmospheric greenhouse-gas concentrations at non-dangerous levels and enshrined the principle of "common but differentiated responsibilities," establishing differentiated mitigation expectations across nations [5].

However, the limitations of state-centric climate governance became increasingly apparent during the 2010s [6,7], spurring the rise of voluntary, non-state climate-governance mechanisms, with RE100—launched in 2014 by the Climate Group [8] and the Carbon Disclosure Project (CDP)—being one of the most influential. It commits participating firms to sourcing 100% of their electricity from renewable energy and serves as a market-based governance mechanism that institutionalizes private-sector climate action.

Originally driven by global technology companies such as Apple, Google, and Microsoft, RE100 has rapidly diffused across manufacturing, finance, and retail. As supply-chain requirements intensified, the initiative evolved from symbolic environmental signaling into a central pillar of corporate Environmental, Social, and Governance (ESG) management. Across several OECD countries, including South Korea, RE100 has become intertwined with national energy-transition strategies, industrial competitiveness, and the politics of energy-system restructuring. Accordingly, RE100 operates as a complex socio-technical system with significant economic, institutional, and policy implications.

Yet, the meaning of RE100 is neither static nor purely technical. Its social significance is continuously negotiated among governments, firms, civil society, and, critically, the media. It must, therefore, be understood not as a fixed technical or regulatory instrument, but as an evolving socio-technical system (STS) that intersects with South Korea's path-dependent energy regime. Path dependence, in this context, refers to the phe-

nomenon whereby past institutional decisions and established technological trajectories constrain future choices, generating substantial switching costs that make departure from entrenched systems socially and economically difficult [9,10]. South Korea's historically nuclear-reliant and fossil-fuel-dependent energy infrastructure represents precisely such a path-dependent system, with the emergence of RE100 as a global norm creating systemic friction and pressure for transformation.

This study draws on social constructionism [11] and Frank Geels' multi-level perspective (MLP) on socio-technical transitions to conceptualize RE100 as a meso-level institutional mechanism that translates global decarbonization norms into Korea's incumbent energy regime, thereby opening new niches for novel institutional and industrial practices. The MLP framework maps socio-technical change onto three nested analytical levels: the socio-technical landscape (macro-level exogenous pressures), the socio-technical regime (meso-level institutionalized rules and practices), and technological niches (micro-level protected spaces for innovation) [12,13]. From this perspective, the global decarbonization norm functions as a landscape-level pressure that gradually destabilizes the Korean energy regime, with RE100 serving as the intervening transmission mechanism through which this pressure reaches regime and niche actors—local governments, renewable-energy developers, and ESG-oriented corporations—enabling them to establish new institutional pathways.

News media play a central role in this process. Through selection, framing, and narrativization, they shape public and policymaker interpretations of risks, responsibilities, and governance alternatives related to energy transition. Clarifying the trajectory of this discourse requires considering the unique historical context of the South Korean media landscape. Historically, the South Korean news media have been highly centralized and ideologically bifurcated, playing a pivotal role in shaping public sentiment and driving state-led developmental agendas. When RE100 was launched globally in 2014, its permeation into the Korean cultural discourse was initially minimal, overshadowed by state-centric, nuclear-reliant energy paradigms. However, between 2014 and 2022, global supply-chain pressures began to filter through multinational corporations into domestic governance, shifting RE100 from an obscure international initiative to a contested national imperative embedded across corporate and community systems.

In South Korea, RE100 gained marked visibility after the 2022 presidential debates, triggering a surge in media coverage. Corporate commitments and deliberations among government policymakers further accelerated RE100's transformation into a national agenda. However, Korean news media presented RE100 using divergent frames—green innovation, corporate burden, or threat to industrial competitiveness—revealing competing policy imaginaries and institutional interests. Consequently, RE100 became a target of discursive contestation within Korea's energy-policy debate.

Such variation underscores a methodological challenge in climate-communication research. Although qualitative discourse analyses illuminate ideological patterns, they often rely on limited samples. Given the scale and velocity of climate-related news production, computational approaches have become increasingly essential for uncovering structural discursive patterns and temporal dynamics.

BERTopic, a transformer-based topic-modeling method, is particularly well suited to this task. By combining contextual embeddings with semantic clustering, BERTopic captures linguistic nuance with minimal preprocessing and traces shifts in discourse over time. As such, it is especially appropriate for analyzing RE100, a rapidly evolving, multi-actor policy arena shaped by complex socio-technical interactions.

This study applies BERTopic to 21,901 news articles to examine how Korean media construct and institutionalize RE100. By interpreting topic clusters through an STS and

media-policy lens, the study illuminates how RE100 is framed, contested, and normalized in Korea's evolving energy-transition environment. The study poses the following research questions:

RQ1. What thematic and institutional dimensions structure Korean news media's discursive construction of RE100, and how do these dimensions map onto distinct governance domains?

RQ2. How do the identified topic clusters relate to the socio-technical institutionalization process (externalization, objectivation, internalization) and reflect the multi-level socio-technical dynamics (landscape pressures vs. regime transitions)?

2. Theoretical Background

2.1. RE100 as an Evolving Socio-Technical System

Conventional analyses of RE100 have placed it primarily within the domains of voluntary corporate governance, energy policy feasibility, and ESG disclosure frameworks [14–16]. While valuable, these approaches treat RE100 as a relatively bounded policy instrument whose evolution can be understood through the decisions of discrete organizational actors. This study departs from that tradition by reconceptualizing RE100 as an evolving STS—a complex, multi-actor assemblage of technologies, institutions, regulations, and practices that collectively shape societal functions such as energy provision and industrial production [17,18].

Geels' MLP provides the analytical scaffold for this reconceptualization [19,20]. From an MLP standpoint, the global institutionalization of decarbonization norms—exemplified by the Paris Agreement and sustainability-oriented supply-chain mandates—constitutes a landscape-level pressure that destabilizes incumbent energy regimes [21]. RE100 operates as a key transmission mechanism of this landscape pressure, channeling global decarbonization expectations into national and firm-level governance through market-based supply-chain requirements and ESG reporting standards. In South Korea, where the incumbent energy regime relies heavily on coal and nuclear power, centralized grid infrastructure, and state-dominated energy planning [22–24], RE100 arrives as a landscape-induced shock that regime actors simultaneously resist, negotiate, and partially absorb.

The terms landscape pressure, transmission mechanism, and socio-technical system are not interchangeable; they denote three analytically distinct levels. At the MLP landscape level, the decarbonization norm itself (e.g., the Paris Agreement and sustainability-linked supply-chain mandates) constitutes the landscape-level pressure. RE100 is not this pressure but functions as its transmission mechanism—a meso-level institutional conduit that channels landscape-level expectations into regime-level governance. As this transmission process becomes institutionalized, RE100 progressively acquires the properties of an evolving STS, comprising actors, technologies, and rules. Finally, the term “discursive formation” refers to the empirical object observed in the present study. Because the data consist of media discourse rather than direct observation of technologies or institutions, we analyze RE100's discursive instantiation—that is, the layer of the STS that is analyzable through our method rather than the entire STS.

Applying STS theory to RE100 yields two analytical advantages. First, it directs attention to the multi-scalar and multi-actor dynamics through which RE100 is institutionalized, spanning global governance bodies, national industrial policy, regional development strategies, and corporate ESG systems. Second, it frames media discourse as a constitutive element of socio-technical transition rather than a mere reflective surface.

We further distinguish two dimensions of media agency that operate jointly rather than as alternative descriptions of the same phenomenon. Korean news media do not establish the substantive positions at stake in RE100 discourse; rather, these positions are sup-

plied by other actors (e.g., local governments, corporations, policymakers). In this sense, the media function as an arena—a discursive space in which these externally generated claims are staged and contested [25,26]. Still, the media do not function as a neutral or passive arena: through selective amplification, sequencing, and legitimation, the media exercise a structuring, selective-amplification function that shapes which claims gain traction, when, and at what institutional level [27–29]. We refer to this dual role as the media’s function as a “selective amplifier and legitimating arena.”

2.2. Systemic Institutionalization: Externalization, Objectivation, and Internalization

To analyze how the STS dynamics of RE100 are discursively produced and reproduced, this study draws on Berger and Luckmann’s [11] social constructionist framework, reinterpreted through the lens of systemic institutionalization. Berger and Luckmann conceptualize the construction of social reality as a cyclical process of externalization, objectivation, and internalization. Externalization refers to the projection of subjective experience into the social world, where it is transformed into shared meanings through linguistic and symbolic systems. Objectivation occurs as these meanings crystallize into norms, institutions, and practices that appear as objective reality. Internalization involves the reabsorption of institutionalized meanings into individual cognition and behavior, rendering them naturalized and taken for granted. This constructionist approach has similarly informed discourse analyses of political reality construction in the Korean context [28].

Applied to the socio-technical domain, this tripartite process maps onto distinct phases of systemic institutionalization. Externalization corresponds to the initial discursive articulation of a new governance norm by pioneer actors—governments, firms, or advocacy organizations—who translate global sustainability expectations into locally meaningful claims. This phase is observable in early political declarations, local government pledges, and media coverage that frame RE100 as an aspirational target. Objectivation corresponds to the crystallization of these discursive claims into material and regulatory structures, such as industrial park designations, renewable energy certification systems, public procurement mandates, and physical infrastructure investments. Internalization, finally, corresponds to the deep embedding of RE100 within organizational routines, corporate strategy, ESG reporting systems, and professional norms, such that compliance with RE100 becomes an unremarkable organizational expectation rather than a contested choice.

This systemic institutionalization framework provides a dynamic alternative to static framing analyses. Rather than asking which frames are applied to RE100 at a given moment, it asks how RE100 discourse evolves across institutional levels and actor networks over time, tracing the transition from peripheral norm to systemic governance principle.

2.3. Discursive Dynamics in Energy Transition Systems

Environmental communication encompasses not only the transmission of environmental information but also the broader symbolic processes through which societies define, interpret, and negotiate environmental issues [29]. The environment is simultaneously a biophysical entity and a socially constructed concept shaped through discourse.

Anderson [30] employs a constructionist lens to illustrate how media framing influences public understanding of environmental problems, underscoring that the news media actively produce rather than passively reflect environmental meaning. Previous studies have empirically demonstrated the discursive power of news coverage in the context of climate change. Boykoff and Boykoff [31] showed that U.S. newspapers frequently adhered to the journalistic norm of “balanced reporting,” giving equal weight to climate skepticism despite a strong scientific consensus against it. Subsequent cross-national studies by

Boykoff [32–34] reveal that climate discourse varies substantially across countries due to political, economic, and cultural divergences.

Research in Korea has similarly identified ideological divergence in climate-change discourse. Kim et al. [35] found that conservative newspapers emphasized economic growth and market considerations, whereas progressive outlets highlighted governmental responsibility and environmental protection. More recently, Choi and Cho's [36] structural topic-modeling analysis of 42,881 Korean news articles (2000–2022) revealed persistent ideological bifurcation: conservative media privilege economic, technological, and national-security dimensions, whereas progressive media foreground accountability, governance, and ecological values.

RE100 occupies a distinct position within this landscape for three reasons. First, as a global initiative that rapidly gained attention in Korea after the 2022 presidential debate, RE100 exemplifies the interaction between international governance norms and domestic policy debates. Second, RE100 sits at the intersection of environmental policy, corporate ESG strategy, national industrial competitiveness, and technological feasibility, generating competing frames such as economic burden, market opportunity, and innovation imperative. Third, the emergence and stabilization of RE100 discourse in Korea have occurred rapidly, intensifying discursive contestation.

Empirical analyses of Korean media discourse on RE100—its semantic structures, temporal dynamics, and actor-specific meaning production—remain scarce. The present study addresses this gap through a large-scale, data-driven discourse analysis capable of identifying underlying thematic patterns and mapping discursive evolution across the socio-technical transition.

3. Methods

This study employs a mixed-methods design that integrates large-scale computational topic modeling with a socio-technical and social constructionist discourse-analytic framework. The methodological procedure consists of three components: (1) the selection of a BERT-based language model, (2) the implementation of BERTopic for topic extraction, and (3) data collection, preprocessing, and morphological analysis. Figure 1 illustrates this research framework.

Bidirectional encoder representations from transformers (BERT), introduced by Google in 2018, utilizes a transformer-based architecture to capture bidirectional contextual information during pretraining [37]. By modeling contextual meaning in forward and backward directions, BERT provides richer semantic representations than unidirectional language models, enabling it to capture subtle linguistic relations within sentences.

These characteristics align closely with the analytical requirements of STS-oriented discourse analysis, which necessitate understanding context-sensitive meaning structures across heterogeneous actor networks rather than surface-level word co-occurrences. Because RE100 coverage involves multiple stakeholders and heterogeneous semantic layers, traditional topic-modeling techniques such as latent Dirichlet allocation or latent semantic analysis are insufficient for capturing nuanced semantic relations. Their performance depends significantly on preprocessing quality and often fails to account for deeper contextual associations between terms.

In contrast, BERTopic [38,39], which combines sentence-level BERT embeddings with density-based clustering, is increasingly recognized as a robust alternative for analyzing complex social and policy discourses. BERTopic preserves contextual meaning, reduces preprocessing dependency, and integrates embedding, dimensionality reduction, and density-based clustering to enhance topic coherence. These features make it particu-

larly well suited for large-scale media discourse analysis of evolving socio-technical systems, where semantic nuance and multilayered meaning construction are critical.

BERTopic Modeling

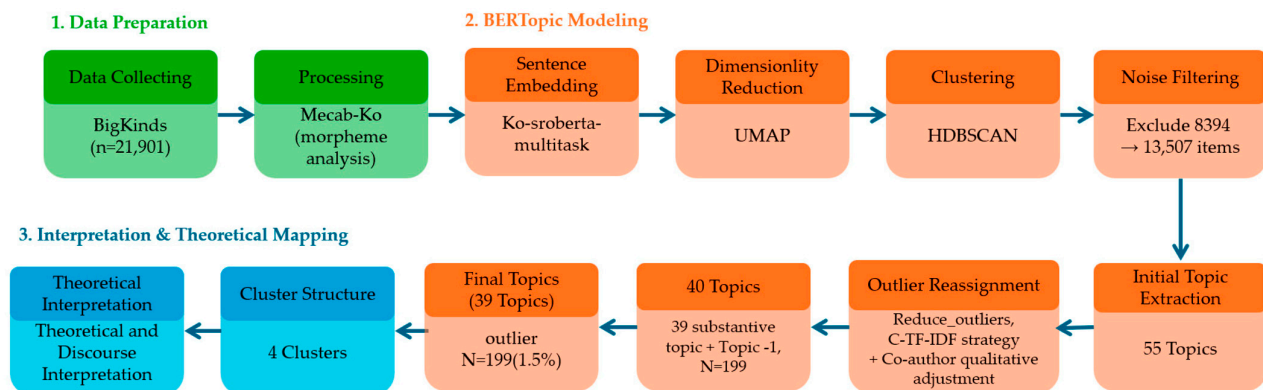


Figure 1. Overall research framework, from data collection and preprocessing through BERTopic modeling (embedding, UMAP, HDBSCAN), noise-topic filtering, and clustering, to theoretical mapping onto the STS institutionalization framework. The BERTopic modeling stage (embedding, UMAP, HDBSCAN) is depicted as a single pipeline for visual clarity; in practice, that pipeline was executed twice—first on the full 21,901-article corpus to identify noise topics, then on the filtered 13,507-article corpus, including the hyperparameter grid search described in Section 3.1 (see below). The 55-topic solution is the output of this second, re-estimated run. Following initial clustering, the 55-topic solution was passed through an outlier-reassignment step using BERTopic’s `reduce_outliers` function with the `c-TF-IDF` strategy, which reassigns each outlier document to its most similar existing topic based on cosine similarity between `c-TF-IDF` vectors. These algorithmic outputs were treated as a starting point rather than a final assignment. Two co-authors subsequently reviewed the candidate reassignments and, where the algorithmic output did not adequately reflect the social and discursive context of individual articles, qualitatively adjusted the topic assignment based on close reading. This combined algorithmic and qualitative process produced a 40-topic solution comprising 39 substantive topics and a residual Topic −1 category containing 199 documents (1.5% of the analytic corpus) that could not be confidently assigned to any existing topic. Excluding Topic −1 yielded the final 39-topic solution. The 199 outlier documents (1.5%) were retained separately, consistent with the Topic −1 category reported in Section 4.

3.1. BERTopic Modeling Procedure

BERTopic incorporates a pipeline consisting of sentence embeddings, dimensionality reduction, density-based clustering, and class-based TF-IDF weighting. This study implemented BERTopic in Google Colab with GPU acceleration to improve computational efficiency. Sentence embeddings were generated using Sentence-BERT (jhgank/ko-sroberta-multitask via the `sentence-transformers` library), producing 768-dimensional dense vectors for each document. These vectors were reduced to a five-dimensional space using uniform manifold approximation and projection (UMAP) with `n_neighbors = 25`, `n_components = 5`, and cosine distance as the metric. Subsequently, HDBSCAN was applied for density-based clustering, with `min_cluster_size = 100` and `min_samples = 20` using Euclidean distance. This multi-stage process yielded coherent topic clusters. This specific hyperparameter configuration was determined through a rigorous grid-search optimization process. While lower values of `min_cluster_size` generated overly fragmented, hyper-local topics and higher values obscured fine-grained thematic transitions, the selected parameters optimized the trade-off between topic coherence and interpretive granularity, ensuring both statistical robustness and conceptual clarity.

Document vectorization was conducted using the transformer model *jhgan/ko-srobeta-multitask*, which is trained on Korean STS and NLI datasets under a multi-task learning framework and is optimized for generating high-quality Korean sentence embeddings [40]. Its capacity to capture fine-grained semantic similarity substantially improves cluster cohesion in BERTopic, making it well suited for analyzing Korean-language media discourse around socio-technical transitions.

The combination of computational modeling via BERTopic and interpretive discourse analysis is particularly appropriate for this study. While traditional qualitative discourse analysis offers deep ideological insights, it is inherently limited by sample size and prone to researcher bias when handling massive datasets. Conversely, purely statistical text-mining often strips away semantic context. This study utilizes the contextual accuracy of transformer-based embeddings to systematically map thousands of articles, while subsequently employing an STS-grounded constructionist framework to interpret the deeper, power-laden narratives within those clusters. This approach ensures both empirical scalability and interpretive rigor.

3.2. Data Collection and Preprocessing

A total of 21,901 news articles were collected from BigKinds using “RE100” as the primary keyword, spanning January 2019 to October 2025. Duplicate entries were removed based on URL matching, and no duplicates remained after verification. Although BERTopic performs adequately even with minimal preprocessing, additional cleaning was undertaken to enhance analytic precision. Particles and inflectional endings (e.g., 은/는/이/가) were removed through a user-defined dictionary. Morphological analysis was conducted using MeCab-Ko, an adaptation of the MeCab system [41] optimized for Korean morphology. Implemented through the KoNLPy library [42], MeCab-Ko applied a conditional random field model to generate context-sensitive morphological segmentation and part-of-speech tagging. Only words belonging to analytically relevant parts-of-speech categories—common nouns, proper nouns, adjectives, and verbs—were retained. Subsequent refinement included the removal of functional words, the recombination of erroneously separated compound nouns (e.g., “RE 100” → “RE100”), and the elimination of media outlet names and irrelevant terms by expanding the stop-word list.

An initial BERTopic run was performed on the full preprocessed corpus of 21,901 articles to identify non-substantive content prior to the core analysis. Among the resulting topics, those capturing routine stock-market updates and generic corporate PR news were treated as non-substantive noise that did not contribute to the socio-technical discourse around RE100.

The exclusion criteria were operationalized as follows. Topics were flagged as non-substantive noise when their top-15 keywords contained two or more terms from a pre-defined stock-market lexicon (e.g., KOSPI, KOSDAQ, share price, trading volume, foreign/institutional net buying) or a corporate-PR lexicon (e.g., award ceremony, MOU signing, commendation, appointment); the full term lists for both lexicons are provided in Table 1. Each flagged topic was manually audited against five representative headlines before exclusion. Articles in these flagged topics (8394 articles; 38.3% of the deduplicated corpus) were then systematically removed, yielding a refined corpus of 13,507 articles for the core analysis. BERTopic was subsequently re-estimated on this 13,507-article corpus, including the hyperparameter grid search reported in Table 2 below. This re-estimation yielded 55 topics. Documents assigned to no topic in this solution ($n = 199$) constitute the Topic −1 outlier category reported in Section 4. The 55 topics were then merged by co-author consensus on the basis of semantic overlap and interpretability within the four-cluster framework presented in Section 4, yielding the final 39 topics reported in Tables 3–6.

Table 1. Excluded Noise Topics, Matched Lexicon Criteria, and Representative Headline Examples.

No	English Translation for Manuscript_Size	Matched Lexicon/Noise Category
1	Investment report maintaining a ‘buy’ rating and target price (KRW 170,000) for Samsung C&T, analyzing stock undervaluation.	Stock Market Lexicon (<i>target price, buy rating, stock undervaluation</i>)
2	Stock market update reporting a 9.52% surge in Shinsung E&G’s share price following an MOU disclosure.	Stock Market and Corporate PR (<i>share price surge, stock trading, MOU disclosure</i>)
3	Daily stock trading update reporting a 2.68% gain in Samsung Electronics’ share price on the stock exchange (KRX).	Stock Market Lexicon (<i>share price gain, stock exchange, daily trading</i>)
4	Corporate PR coverage highlighting a 12.4% YoY surge in new contract orders and order backlog figures.	Corporate PR Lexicon (<i>new contract orders, YoY order growth</i>)
5	Market investment column focusing on stock index movements (KOSPI up 20%, KOSDAQ up 8%) and semiconductor stock allocations.	Stock Market Lexicon (<i>KOSPI, KOSDAQ, stock index, equity investment</i>)

Note: Words in italics are illustrative example keywords matched to each lexicon category, not an exhaustive list.

Table 2. Performance comparison across hyperparameter search configurations. Evaluated metrics include the number of generated topic clusters, the outlier (unassigned-document) rate (%), and per-topic NPMI-based topic coherence (C_{npmi}), computed as the arithmetic mean of pairwise NPMI scores over each topic’s top-10 keywords. The outlier rate ranged from 28.55% to 39.37%, and per-topic C_{npmi} ranged from 0.1031 to 0.1331 across the nine configurations. The chosen optimal configuration (min_cluster_size = 100, n_neighbors = 25) is shown in bold. Note: the reported outlier percentages (28.55–39.37%) reflect the initial HDBSCAN solutions prior to the subsequent c-TF-IDF-based outlier-reassignment step; they are not directly comparable to the 1.5% Topic –1 rate reported for the final 39-topic solution in Section 4.

Min_Cluster_Size	n_Neighbors	n_Topics	Outlier_Pct	per-Topic C_{Npmi}
50	15	125	30.10%	0.1204
50	25	122	28.55%	0.1264
50	35	124	29.87%	0.1331
75	15	72	31.24%	0.1100
75	25	76	32.72%	0.1186
75	35	77	36.54%	0.1267
100	15	55	39.08%	0.1031
100	25	55	38.37%	0.1124
100	35	59	39.37%	0.1184

While this computational design offers significant transparency, it has two limitations in the clustering process. First, transformer-based models are sensitive to hyperparameters, such as HDBSCAN’s minimum cluster size and UMAP’s dimensionality-reduction settings, which can shift topic boundaries. Second, even though the clustering itself is algorithmic, mapping clusters onto STS stages inevitably introduces interpretive subjectivity. Two complementary checks were performed. Sensitivity analyses varied key parameter settings and supported the final specification (min_cluster_size = 100, n_neighbors = 25). Separately, two coders independently evaluated (a) whether the top keywords of each of the 39 topics were semantically consistent with that topic’s assigned label, and (b) whether each of the four cluster names adequately represented its constituent topics. Initial agreement on these 39 topic-label items was 87.3%; remaining disagreements were resolved through discussion between the two coders.

Concretely, the hyperparameter search evaluated a grid of $\text{min_cluster_size} \in \{50, 75, 100\}$ and $\text{n_neighbors} \in \{15, 25, 35\}$, with each of the nine resulting configurations scored on the number of topics produced (range: 55–125 across the grid), the outlier (unassigned-document) rate (28.55–39.37%), and topic coherence (C_{npmi} , computed over each topic's top-10 keywords; range: 0.1031–0.1331); full results are reported in Table 2. The optimal configuration ($\text{min_cluster_size} = 100$, $\text{n_neighbors} = 25$; 55 topics) was selected because it struck the most favorable overall balance among these three criteria relative to coarser and finer settings in the grid. Configurations yielding substantially more topics (e.g., 122–125 topics at $\text{min_cluster_size} = 50$) achieved only marginally higher coherence scores (up to $C_{\text{npmi}} = 0.1331$) relative to the substantial increase in topic count, indicating diminishing returns in coherence beyond a particular granularity, and also achieved somewhat lower outlier rates (28.55–30.10%). Conversely, coarser configurations did not correspondingly improve coherence relative to the selected setting. Although the outlier rate of the selected configuration (38.37%) was comparatively higher than that of the finer-grained alternatives, we deemed this trade-off acceptable because the resulting topic count (55) offered a substantially more interpretable granularity for subsequent qualitative categorization, while coherence ($C_{\text{npmi}} = 0.1124$) remained comparable to other mid-granularity settings in the grid.

To further assess the robustness of the selected configuration to stochastic variation, we conducted a random-seed stability check by re-running the final configuration across five random seeds (0–4). Across seeds, the resulting topic count ranged from 53 to 62, and the outlier rate ranged from 35.1% to 40.5%, indicating modest variation in overall model granularity consistent with the range observed in the hyperparameter grid. To quantify pairwise agreement across seed pairs, we computed the Adjusted Rand Index (ARI), normalized mutual information (NMI), and top-keyword Jaccard similarity. Document-level cluster assignment agreement was moderate ($\text{ARI} = 0.644 \pm 0.041$), reflecting some instability in the precise boundary assignment of individual documents across runs. In contrast, both the overall information-theoretic agreement between clustering solutions ($\text{NMI} = 0.823 \pm 0.015$) and the similarity of topic-defining keyword sets ($\text{Jaccard} = 0.807 \pm 0.028$) were high and exhibited low variance across seed pairs.

These results indicate that while document-level cluster boundaries show some sensitivity to random initialization, the substantive thematic content and structure of the topic model—the primary basis for the subsequent interpretive analysis—remained stable across seeds.

Outlier documents from the initial 55-topic solution (38.37% of the analytic corpus; Table 2) were first reassigned to existing topics using BERTopic's built-in `reduce_outliers` function with the c-TF-IDF strategy. This procedure computes cosine similarity between each outlier document's c-TF-IDF vector and those of existing topics. This algorithmic output served as a starting point rather than a final assignment. Two co-authors subsequently reviewed the reassignment and, where the algorithmic output did not adequately reflect the social and discursive context of individual articles, qualitatively adjusted the topic assignment based on close reading. This combined algorithmic and qualitative process yielded a 40-topic solution comprising 39 substantive topics and a residual Topic –1 category of 199 documents (1.5% of the analytic corpus) that could not be confidently assigned to any existing topic. No similarity threshold was systematically recorded for the algorithmic step, so the exact reassignment criteria cannot be fully reconstructed; this limitation is noted in Section 6. Excluding Topic –1 yielded the final 39-topic solution.

Although this configuration initially yielded 55 topic clusters, overlapping topics were merged based on their analytical utility for socio-technical discourse analysis, result-

ing in a consolidated set of 39 final topics. Consequently, this setting was adopted for all subsequent analyses.

4. Results

After filtering out routine stock updates and corporate PR anomalies, the final refined dataset of 13,507 articles yielded 39 analytically meaningful topics. A complete inventory of all 39 topics—including topic identifiers, document counts, proportions of the corpus, top keywords, and cluster assignments for every topic (not only those discussed in-text)—is provided below. The total document count across these 39 named topics sums to 13,308 articles; the remaining 199 articles (1.5%) of the 13,507-article analytic corpus were assigned to the unclustered outlier category (Topic −1) and are not represented in any of the 39 named topics or the cluster-level tables below. This 13,507-article analytic corpus itself excludes 8394 non-substantive noise articles (38.3%)—routine stock-market updates and corporate PR content—from the initial pool of 21,901 collected articles. These topics were subsequently organized into four overarching interpretive clusters based on their semantic proximity and functional roles within the discourse (Figure 2). Guided by an STS-informed social constructionist framework, this analysis sought to identify how Korean news media discursively reconstructs RE100 and to map the latent semantic structures underpinning its emergence, institutionalization, and contestation. Rather than treating these topics as isolated media events, this study contextualizes them as interconnected socio-technical elements that collectively delineate the macro-level transformation of South Korea’s energy paradigm.

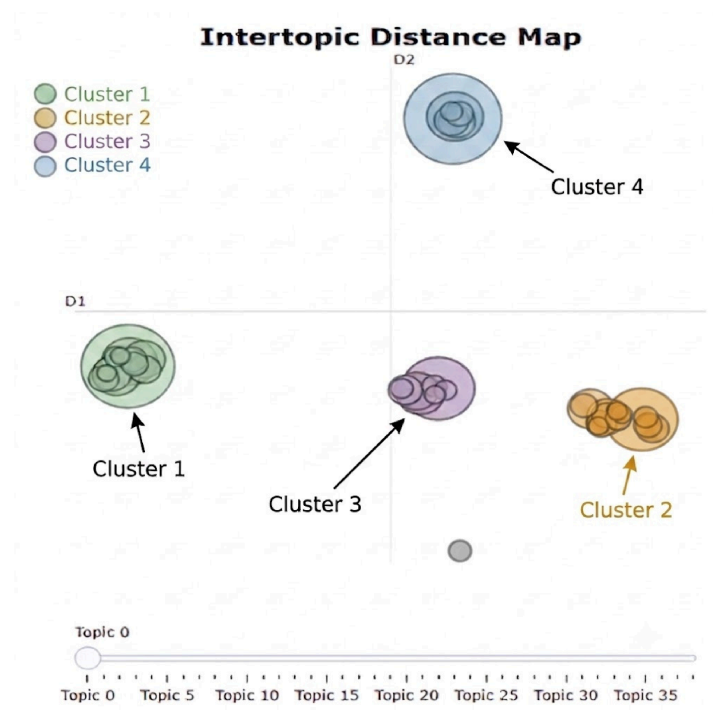


Figure 2. Intertopic distance map of the extracted topics via multidimensional scaling. The map illustrates the distinctness of identified topics. Several tight clusters are observed (e.g., the large group on the far left), suggesting sets of highly correlated topics. Conversely, isolated topics are spatially distant, indicating unique thematic content. Note: The degree of overlap among topic circles within each cluster reflects the semantic proximity of individual topics in the embedding space; overlapping circles indicate closely related topics rather than a loss of information. Exact topic counts are provided in Tables 3–6, and corresponding proportions of the corpus are reported in the text (Section 4). The gray circle indicates an outlier topic that was not assigned to any of the four main clusters.

The structural distribution of these 39 topics reveals a dynamic socio-technical landscape in which global sustainability norms encounter path-dependent domestic regimes. Crucially, the identified clusters do not merely categorize disparate media reports; they correspond to distinct phases of systemic institutionalization. The four clusters are organized below according to their substantive focus—administrative externalization (Cluster 1), industrial-policy objectivation (Cluster 2), market-based internalization (Cluster 3), and organizational internalization (Cluster 4)—corresponding to the framework outlined in Section 2.2. Their temporal ordering is examined empirically in Section 5.1. Read in this way, the four clusters illustrate how a voluntary corporate campaign can intersect with an objectivated institutional structure, realigning the relations between governmental frameworks, industrial systems, market mechanisms, and civil society.

To complement the topic-level statistics with a temporal perspective, Figure 3 plots the monthly volume of news articles assigned to each of the four interpretive clusters. Cluster 1 dominates the distribution and exhibits a marked acceleration beginning in late 2021, followed by a sustained surge toward 2025. This pattern is consistent with the concentration of Cluster 1 volume around the 2022 presidential election discussed in Section 5.1. Clusters 2, 3, and 4 remain comparatively modest and stable in absolute volume for most of the period, although Cluster 3 (market-based internalization) shows a modest late-stage increase toward 2025. This pattern emerges because the visualization displays raw monthly counts rather than each cluster's share of its own mean coverage.

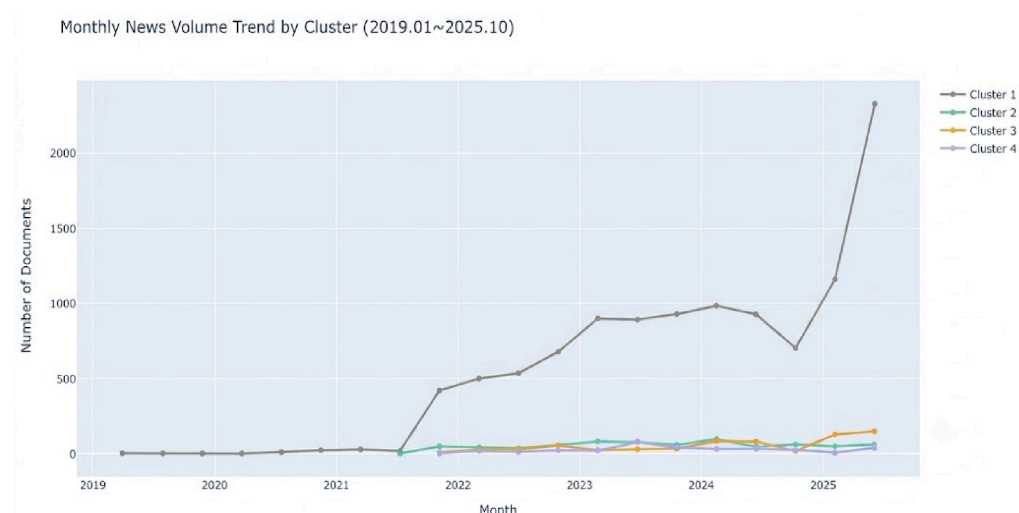


Figure 3. Monthly news volume trend by cluster (January 2019–October 2025).

4.1. Cluster 1: Spatial Sub-System and Administrative Externalization

Cluster 1 represents the largest semantic domain, capturing how local governments shape RE100 across administrative, political, and participatory contexts, as detailed in Table 3. This cluster comprises 13 of the 39 topics (4799 articles, 35.5% of the analytic corpus)—the largest of the four clusters by both topic count and document share. This cluster reflects RE100's localization as a socially embedded governance practice rather than a top-down or corporate-driven initiative. Provinces and municipalities—such as Gyeonggi, Jeolla, Gwangju, and Pyeongtaek—play central roles alongside regional politicians and community actors. From an STS perspective, these local governmental actors constitute the regional sub-system within which the initial externalization of RE100 as a governance norm occurs prior to and independently of broader national regime transformation.

Table 3. Topics, Document Counts, and Keywords for Cluster 1.

Topic	Documents	Keywords
0	1526	Kim Dong-yeon, Northern, Suwon, Provincial Citizens, Heungwon, Hwaseong, Cooperative, Carbon Neutral Green, Action, Green Growth
7	562	Honam, Running for Office, Autonomous Province, Special Autonomy, Preliminary Candidate, Preliminary, Democratic Party Presidential Election, Special Autonomous Province, Pledge Announcement, General Election
8	530	Nuclear Power Plant, Presidential Candidate, Power Supply-Demand, Nuclear Energy, Lee Jae-myung/Democratic Party, TV, Climate Energy, General Election, Resources Minister, Criticism
9	477	Kim Dong-yeon, Governor Kim, Kim Dong-yeon Gyeonggi, Northern, Gyeonggi Governor, Budget Bill, Finance, Cooperation, Popularly Elected, Public Livelihood
13	327	Gwangju City, Metropolitan City, Mobility, Strategic Project, RE100 Promotion, National Strategy, City, Distributed Energy, National Industrial Complex, Technology Institute
14	305	Speech, Democratic Party Leader, Recovery, Public Livelihood, Lee Jae-myung/Democratic Party, City Administration, General Election, Administration, Livelihood Economy, President Lee Jae-myung
16	249	Park, Site Utilization, Maintenance, Perceived Benefit, Surroundings, RE100 Utilization, Environmental Creation, Nuri, City/County, Provincial Citizens
18	173	Materials, Food, Spain, Floating Solar Power, Korea Midland Power, Renewable Energy Development, Korea Midland, Water Resources Corporation, Water Resources, Renewable Energy Adoption
25	152	Semiconductor Industry, Strategic Industry, Gwangju City, Win-Win Cooperation, South Jeolla Region, Strength, Domestically Unique, University President, Wide-Area, Metropolitan City
29	147	Village, Housing, Household, Income, Support Project, Rooftop, Apartment, Pilot Project, Sunlight, Solar Installation
31	126	Broadcast, Debate, Radio, Contest, Difference, TV, CBS, Presidential Candidate, Dialogue, Appearance
33	120	Rural Area, Agriculture, Support Project, Livestock, Village, Food, Project Competition, Target Site, Selected by Competition, North Chungcheong
35	105	Semiconductor Industry, Tax Credit, Tax Amount, Deduction, Presidential Candidate, Special Act, SK Hynix, Hynix, Lee Jae-myung/Democratic Party, Powerful Nation

- Topics 0 and 9: Kim Dong-yeon and the Institutionalization of Gyeonggi's RE100 Agency.
Topics 0 (1526 articles) and 9 (477 articles) jointly account for 2003 articles—14.8% of the entire analytic corpus and the largest single thematic concentration. Both topics are organized around Governor Kim Dong-yeon and Gyeonggi Province's RE100 administration (budgets, provincial cooperatives, carbon-neutral green growth initiatives), indicating that a single regional executive and a single province receive a disproportionate share of media coverage of RE100 in Korea. This concentration qualifies the more general claim that “local governments” as a class externalize RE100: the pattern is substantially carried by one high-profile provincial administration rather than being evenly distributed across municipalities.
- Topics 7, 8, 14, and 31: Presidential Election Discourse and Broadcast Amplification.
Topics 7 (562), 8 (530), 14 (305), and 31 (126) comprise 1523 articles (11.3% of the corpus) organized around the 2022 presidential election—candidate pledges in the Honam region, nuclear-versus-renewable energy debates, Democratic Party leadership speeches, and televised/radio debate formats (TV, CBS)—confirming that electoral contestation is a major discursive channel through which RE100 entered national political debate, though the topic-modeling design used here does not allow us to establish a temporal onset date for this sub-theme specifically.

- Topics 13, 16, 18, 25, 29, 33, and 35: Regional Development, Semiconductor Strategy, and Community-Level Solar.

The remaining seven topics (327, 249, 173, 152, 147, 120, and 105 articles respectively; 1273 articles combined) form a more heterogeneous residual grouping: Gwangju's mobility and semiconductor-industry strategy (Topics 13, 25, 35), municipal park and public-land utilization for RE100 (Topic 16), rural/agricultural support programs (Topic 33), and household- and village-level rooftop solar installation (Topic 29). Compared with the Kim Dong-yeon/Gyeonggi and electoral sub-groups above, these topics point to a more dispersed, place-specific form of administrative externalization that spans provincial industrial policy down to individual rooftop programs, rather than a single coherent narrative.

4.2. Cluster 2: Techno-Economic Reconfiguration and Objectivation

Cluster 2 consists of 10 topics describing how RE100 is rearticulated through regional industrial complexes and large-scale energy infrastructures across Jeonnam, Saemangeum, Naju, Ulsan, Gyeongbuk, and Cheongju, as summarized in Table 4. Cluster 2 comprises 3364 articles (24.9% of the corpus). Unlike Cluster 1, which centers on political-administrative externalization, Cluster 2 demonstrates RE100's objectivation through spatial reconfiguration as part of the national industrial strategy. Data centers (Topic 1), off-shore wind (Topic 27), and hydrogen infrastructure (Topic 22) illustrate the transition from aspirational rhetoric to physical and technical infrastructure, reconfiguring existing industrial spaces. Collectively, the topics in Cluster 2 show that RE100 has evolved beyond an environmental initiative into a strategic framework linking industrial competitiveness, regional restructuring, hydrogen transitions, and energy security.

Table 4. Topics, Document Counts, and Keywords for Cluster 2.

Topic	Documents	Keywords
1	1152	South Jeolla Province, Southwest, County Chief, Special Zone, AI Data, Future City, Welcome, Power Grid, Special Act, Farming
3	737	South Chungcheong, Ulsan, Industrial Complex RE100, RE100 Industrial Complex, Office of the President, RE100 Project, Task Force, Recruitment, RE100 Industrial Complex, Offshore Wind
10	432	Saemangeum, Autonomous Province, Special Autonomous Province, Energy Solution, RE100 Energy, Distributed Energy, Gunsan, Development Corporation, Subcommittee, Energy Specialization
17	213	Jeju, 2035, First in the Nation, Demonstration Project, Energy Great Transition, Agriculture, Launch, Livestock, RE100 Production, Wind Power
24	156	County Chief, RE100 Nation, Hydrogen Production, Saemangeum, Production Base, National Industrial Complex, Base, Budget Securing, Industrial Complex Development, Offshore Wind Power
22	153	Construction Project, Distributed Energy, Hydrogen Fuel Cell, Steel, Resource Organizing, Energy Specialization, Fuel Cell, National Industrial Complex, National Budget, Project Competition
27	148	Naju, Advisory, RE100 Energy, Global Energy, Task Reflection, Investment Attraction, Grid, Energy Sector, Science and Technology, Planning Committee
23	139	Offshore Wind Power, Sinan, Offshore Wind, Offshore Wind Power Generation, GW, Wind Power, Acceptability, Wind, Great Transition Era, Public-Private
30	134	Pyeongtaek, Clean Hydrogen, Hydrogen Production, Project, Production Facility, Briefing, Press, Model Development, Service Contract, Bus
36	100	North Chungcheong, Cheongju, Fluid, Clean Hydrogen, RE100 Membership (100), Hundred, Idea, Club, LNG, Cheonsi

- **Topics 1 and 3: Data-Center-Driven Industrial Complexes and Presidential Task Forces.**
Topics 1 (1152 articles) and 3 (737 articles) jointly account for 1889 articles (14.0% of the corpus), making this the largest sub-theme in Cluster 2. Topic 1 centers on AI and data-center demand driving special-zone and future-city designations in South Jeolla Province; Topic 3 documents the Office of the President’s task-force involvement in designating RE100 industrial complexes in South Chungcheong and Ulsan. Together, they indicate that digital-industry site selection and top-down state task forces are the two principal engines of RE100’s objectivation into physical industrial infrastructure in this corpus.
- **Topics 22, 24, 30, and 36: Hydrogen Infrastructure Objectivation.**
Topics 22 (153), 24 (156), 30 (134), and 36 (100) together comprise 543 articles focused on hydrogen fuel-cell construction projects, national industrial-complex budget allocation, and clean-hydrogen production facilities in Pyeongtaek and Cheongju. This sub-group shows hydrogen infrastructure emerging as a distinct objectivation pathway alongside offshore wind and data centers.
- **Topics 10, 17, 27, and 23: Regional Energy Hubs at Saemangeum, Jeju, Sinan and Naju.**
Topics 10 (432), 17 (213), 27 (148), and 23 (139) center on region-specific renewable-energy hub formation: Saemangeum’s energy-solution development corporation, Jeju’s 2035 energy-transition demonstration projects, and Sinan County’s offshore wind generation. These 932 articles illustrate that objectivation in Cluster 2 is not limited to a single flagship site but recurs across regional hubs with a broadly similar administrative logic (special zones, development corporations, demonstration projects).

4.3. Cluster 3: Market-Based Integration and Systemic Internalization

Cluster 3 comprises eleven subtopics emphasizing market-based, globally aligned RE100 practices (see Table 5). Unlike the administrative or territorial logic of Clusters 1 and 2, Cluster 3 reflects the corporate-led internalization of RE100 as a governance norm within market mechanisms and corporate operating systems. PPA contracts, the integration of financial capital, and the ESG commitments of large conglomerates (Samsung, SK, Hyundai, Lotte) demonstrate how RE100 transitions from external pressure to embedded corporate practice.

Table 5. Topics, Document Counts, and Keywords for Cluster 3.

Topic	Documents	Keywords
2	872	Hanwha, PPA, Plant, LG, Solar Power Plant, Power Trading, Power Purchase, PPA Contract, Power Purchase Agreement, Changwon
4	676	LG, Energy Solution, RE100 Transition, Report, Initiative, Chemicals, RE100 Initiative, Production Facility, ESG Environment, Report Publication
6	609	CF, CF100, Carbon Energy, Coalition, UN, Carbon, General Assembly, International Community, Heatwave, EU
11	405	Hyundai Motor, PPA, Time, Electric Vehicle, CDP, World’s Best, Power Purchase Agreement, Power Purchase, PPA Contract, Supply Expansion
19	188	Expo, Participation, Daegu, Exhibition, Fair, Module, Daejeon, Conference, Future Energy, Smart Energy
20	184	Fund, SK Innovation, Innovation, Power Trading, Domestic RE100, PPA, Wind Power, Global Transition, Gangwon, Energy Agency

Table 5. Cont.

Topic	Documents	Keywords
21	178	LS, Cable, Factory, Carbon Dioxide, Dioxide, Sustainable Management, 2040 Renewable Energy, Office, Renewable Energy Business, Capture
26	140	Eco-Friendly Technology, SK Innovation, Innovation, Logistics, Aviation, Power Purchase Agreement, Power Purchase, Increase, Public Enterprise
28	138	Logistics, Reduction Target, Facility Construction, Major Domestic, Verification, Waste, Vehicle, 2030 Greenhouse Gas, Report, Scale
32	122	Internet, Office Building, Water Resources Corporation, Water Resources, Korea Water Resources, RE100 Membership, Initiative, PPA, Vehicle, Hydropower
37	106	Power Grid, Power Demand, Nuclear Power Plant, Surge, Cloud, AI Data, AI Era, AI Artificial Intelligence, Demand Increase, Nuclear Power

- Topics 2, 11, 20, 26, and 32: PPA-Mediated Corporate Procurement.
Topics 2 (872), 11 (405), 20 (184), 26 (140), and 32 (122) jointly comprise 1723 articles (12.8% of the corpus) organized around power purchase agreements (PPAs): Hanwha and LG solar-plant contracts in Changwon, Hyundai Motor's EV-linked PPAs, SK Innovation's wind-power and logistics/aviation contracts, and Korea Water Resources Corporation's hydropower PPAs. This is the largest and most coherent sub-theme in Cluster 3 and provides the clearest empirical support for describing Cluster 3 as market-mechanism internalization.
- Topic 6: CF100 and International Legitimacy Contestation.
Topic 6 (609 articles; 4.5% of the corpus) is organized around CF100, UN and EU carbon-related coalitions, and international climate summits. Its keyword profile reflects governance tensions and legitimacy contestation surrounding competing decarbonization standards. It remains a standalone, thematically distinct topic rather than being merged with other market-based topics, consistent with treating the CF100-versus-RE100 debate as a discrete, unresolved contestation rather than a routine market practice.
- Topics 4, 19, 21, 28, and 37: Corporate Sustainability Reporting, Emissions Targets, and Digital-Economy Power Demand.
Topics 4 (676), 19 (188), 21 (178), 28 (138), and 37 (106) cover LG's RE100-transition and ESG reporting, corporate 2030 greenhouse-gas reduction-target verification, and AI-driven power-grid demand growth. Together, these 1286 articles (9.5% of the corpus) indicate that market-based internalization in this corpus is not limited to procurement contracts but also includes formal target-setting/reporting and the emerging power-demand pressure from data centers and AI, which overlaps thematically with Topic 1 in Cluster 2 and suggests that digital-economy energy demand is a cross-cutting theme rather than one confined to a single cluster.

4.4. Cluster 4: ESG-Driven Corporate Governance and Deep Internalization

Cluster 4 centers on the deep internalization of RE100 within corporate ESG governance system, as presented in Table 6. Cluster 4 is the smallest of the four clusters, comprising 5 of the 39 topics (1527 articles, 11.3% of the corpus). Through the externalization of global expectations, objectivation via disclosure systems, and internalization into strategy and infrastructure, RE100 alignment is becoming a core organizational principle rather than an optional initiative. Cluster 4 thus demonstrates that Korean corporations are constructing a private-led decarbonization regime in which RE100 is no longer a peripheral voluntary commitment but a constitutive element of organizational identity.

Table 6. Topics, Document Counts, and Keywords for Cluster 4.

Topic	Documents	Keywords
5	633	CDP, Rating, Report, Bio, Acquisition, Securities, Report Publication, Information Disclosure, Sustainable Management, Rating Acquisition
12	377	Environmental Management, 2050 Carbon Neutrality, Strategy Announcement, Vice Chairman, Eco-Friendly Management, Display, Innovative Technology, Domestic Renewable Energy, Target Achievement, RE100 Project
15	299	KT, SK Group, Hynix, Chey Tae-won, SK Hynix, Telecommunications, ESG Environment, Report, Yongin, Consulting
34	118	Kakao, Management Report, Report, Sustainable Management, Initiative, Report Publication, Procurement, Jeju, Renewable Energy Procurement, Global Achievement
38	100	Water Resources, K-water, Korea Water Resources Corporation, Water, Floating Solar Power, Trade Barrier, SK Hynix, Hynix, Barrier, Hydropower

- **Topic 5: CDP Disclosure and Formal Rating Acquisition.**
Topic 5 (633 articles; 4.7% of the total corpus) is the largest topic in Cluster 4 and centers on carbon disclosure project (CDP) assessments, credit-rating and securities-related disclosure, and formal rating acquisition. This topic alone accounts for 41.5% of Cluster 4's total volume, indicating that formal third-party disclosure and rating systems—rather than internal strategy or executive declarations—are the dominant carrier of organizational-level internalization in this corpus.
- **Topics 12, 15, 34, and 38: Executive-Level Strategy Announcements and Corporate ESG Reporting.**
Topics 12 (377), 15 (299), 34 (118), and 38 (100) cover environmental-management strategy announcements by corporate vice chairmen and 2050 carbon-neutrality pledges, SK Group/SK Hynix ESG communications under Chairman Chey Tae-Won, and Kakao's sustainability-management reporting and renewable-energy procurement. These 894 articles (6.6% of the total corpus) support the claim that executive-level declarations function as symbolic acts reinforcing RE100 as an organizational norm, though this sub-theme is now restructured across four topics rather than two (with fewer total articles represented) compared with the original 1886-article version of this cluster.

5. Discussion

5.1. The Temporal Ordering of RE100's Socio-Technical Institutionalization

The four clusters collectively reveal a set of discursive dimensions through which RE100 is socially constructed and institutionalized in Korean news media—administrative externalization (Cluster 1), techno-economic objectivation (Cluster 2), market-based internalization (Cluster 3), and deep organizational internalization (Cluster 4)—broadly reflecting a progression from the periphery of local administrative discourse to the core of national industrial strategy and corporate governance, though only one pairwise ordering of these dimensions (Cluster 4 preceding Cluster 3) is empirically supported as sequential.

Rather than assuming a fixed temporal sequence, the proposed ordering of the four discursive clusters was subjected to a direct empirical test. This test was independently repeated across two full re-estimations of the topic modeling pipeline to assess reproducibility. For each cluster, an onset date was defined as the first month in which its 3-month moving-average share of monthly coverage exceeded 50% of its full-period mean share. Results were robust to threshold choice across a range of 20–80% of the mean in both re-estimations.

One finding was stable across both re-estimations: organizational/ESG internalization (Cluster 4) reached its onset threshold before market-mechanism internalization (Cluster 3) at every threshold tested (100% of thresholds, both runs). This reversal is interpreted in light of institutional decoupling: organizations characteristically adopt legitimating symbolic commitments—such as disclosure practices and executive declarations—before operationalizing them through costlier market mechanisms, such as long-term PPAs. By contrast, the relative onset order of Clusters 1 and 2 was not stable across re-estimations: Cluster 1 preceded Cluster 2 in one run but followed it in the other, indicating that this ordering is sensitive to stochastic variation in the clustering procedure rather than reflecting a robust empirical pattern.

Accordingly, the four clusters are not presented as a uniformly sequential four-stage pathway. Only the ordering Cluster 4 before Cluster 3 is treated as an empirically supported sequential relationship. Clusters 1 and 2 are instead characterized as co-occurring dimensions of externalization and objectivation, whose relative onset cannot be confidently ordered with the current pipeline. Clusters 3 and 4 are characterized as asynchronously ordered internalization pathways, in which symbolic/organizational internalization precedes, rather than follows, market-operational internalization.

This finding extends STS and MLP scholarship in several directions. The volume dominance of Cluster 1 as a whole (13 of 39 topics; 35.5% of the analytic corpus) indicates that local governmental actors were, collectively, the most discursively prominent carriers of RE100 in coverage volume—although, as noted in Section 4.1, this volume is not evenly distributed across municipalities but is substantially concentrated in a small number of high-profile topics (e.g., Topics 0 and 9, centered on a single provincial administration). This pattern supports the “frontrunner” dynamic observed in European energy transitions, where pioneering municipalities and regional governments create institutional precedents that are subsequently adopted at higher governance levels [16]. The temporal surge in Cluster 1 topics following the 2022 presidential debates further illustrates how political events can accelerate externalization, converting a relatively obscure international norm into a domestically contested governance priority almost overnight.

Cluster 2’s emphasis on large-scale physical infrastructure—offshore wind farms, hydrogen hubs, RE100-designated industrial parks, and data-center siting strategies—provides empirical evidence for the objectivation stage. The prominent role of presidential task forces and special legislation in Topics 36 and 22 underscores that objectivation in the Korean context was not a gradual organic process but was actively accelerated by top-down state intervention, reflecting the continued salience of developmental state logic in Korea’s energy-transition governance.

Clusters 3 and 4 illuminate the internalization process at two analytically distinct levels. Cluster 3 documents market-level internalization—the embedding of RE100 within supply-chain PPAs, financial instruments, and corporate operational systems. Cluster 4 documents organizational-level internalization—the incorporation of RE100 into ESG disclosure systems, executive strategy, and corporate culture. The co-presence of both forms of internalization in the corpus suggests that RE100’s institutionalization in Korea has reached an advanced stage in which market pressure and organizational culture are mutually reinforcing. The temporal evidence indicates that Cluster 4’s symbolic/organizational internalization preceded Cluster 3’s market-operational internalization by approximately one year. This finding suggests that mutual reinforcement between these two levels was established only after an initial period in which organizational commitment outpaced market-mechanism adoption.

5.2. Media as Selective Amplifier and Legitimizing Arena in Socio-Technical Transition

A central contribution of this study lies in its reframing of news media as a selective amplifier and legitimating arena—rather than a passive mirror or an autonomous norm entrepreneur—within the socio-technical transition of RE100. The media corpus analyzed here does not originate the substantive positions at stake. Rather, through selective amplification, sequencing, and legitimation, it shapes which externally generated claims gain visibility, when, and at what institutional level. By selectively amplifying certain actor positions (e.g., Gyeonggi Province’s RE100 agency, chaebol ESG commitments), legitimizing specific technological pathways (offshore wind, hydrogen, PPA markets), and circulating contested frames (RE100 vs. CF100 debates), Korean news media function as a discursive arena in which the boundaries of the legitimate RE100 transition are negotiated and stabilized.

This finding aligns with Hansen’s [43] observation that media institutions are not external to environmental politics but are constitutive of the symbolic environment in which environmental governance is produced. It also extends social constructionist media theory by demonstrating that, in socio-technical transitions, media discourse actively constructs the discursive visibility and legitimacy of particular infrastructural pathways and shapes which technological and institutional options become salient and contestable—though the present design cannot establish whether this discursive legitimation translates into material infrastructural outcomes.

5.3. Policy Implications

First, the ideological bifurcation in news coverage documented in prior Korean climate research [35,36] can stall infrastructural energy transitions by creating partisan narratives around RE100 that obstruct the cross-sector consensus necessary for large-scale infrastructure investment. More unified, non-partisan communication frameworks that emphasize industrial survival alongside ecological values may be necessary to sustain momentum.

Second, distinguishing the institutional levels at which RE100 is discursively constructed—regional administrative sub-systems, industrial objectivation, and corporate internalization—provides a transferable analytical matrix for energy-transition governance. Policymakers seeking to accelerate the adoption of successor frameworks such as CF100 or net-zero supply-chain mandates can use this matrix to anticipate points of resistance and identify key actor coalitions at each institutional level. The temporal analysis (Section 5.1) further suggests that symbolic or organizational commitment (e.g., disclosure, executive pledges) may precede operational adoption of market mechanisms (e.g., PPAs). Policymakers may wish to build this asynchrony into the sequencing of interventions rather than assume a uniform administrative-to-corporate rollout.

Third, the prominent role of local governments in Cluster 1 suggests that sub-national administrative actors should be recognized as strategic entry points for new energy governance norms, rather than passive implementers of national policy. Designing institutional support mechanisms—such as RE100 promotion agencies, green public procurement mandates, and regional renewable-energy certification systems—at the local level may accelerate the externalization phase and create the political momentum necessary for national regime transformation.

6. Conclusions

This study conducted a BERTopic-based cluster analysis of 21,901 Korean news articles referencing RE100 and interpreted the results through an integrated socio-technical and social constructionist framework. The findings demonstrate that RE100 functions as a multi-scalar, multi-actor socio-technical system rather than a unified policy narra-

tive. Through interactions among local governments, industrial complexes, large corporations, and public-sector organizations, RE100 is becoming a socially constructed normative framework for Korea's energy transition, interpreted through the theoretical lens of externalization, objectivation, and internalization.

This study makes three key contributions to scholarship on media framing, environmental communication, and energy-transition governance. First, it reconceptualizes RE100 as a discursive formation embedded within an evolving Socio-Technical System, expanding STS-informed and constructionist approaches to energy-transition research. Second, it identifies the differentiated institutional levels through which distinct actors—local governments, industries, corporations, and public institutions—construct and institutionalize RE100 in differentiated ways across administrative, territorial, market, and organizational levels. Third, it clarifies how decarbonization norms become embedded within corporate ESG systems, illuminating the organizational mechanisms through which private-sector actors increasingly shape national energy-transition trajectories.

The multi-layered governance framework developed in this study is highly scalable. It offers a transferable analytical matrix that can be explicitly applied to examine the institutionalization and societal acceptance of other transitional climate initiatives, such as CF100 or global net-zero supply-chain mandates in emerging economies.

Several limitations provide directions for further research. First, while BERTopic reveals broad thematic structures, deeper linguistic insight requires more granular morpheme-based semantic-network analysis to capture co-occurrence dynamics and contextual modifiers. Second, applying actor-network theory could illuminate the hybrid assemblages of human and non-human actors—including technologies, infrastructures, and regulatory instruments—that co-produce the RE100 system. Third, examining the contested narratives surrounding RE100 and CF100 could deepen understanding of legitimacy struggles in Korea's energy-transition politics. Combining sentiment analysis with frame analysis may clarify how these debates reflect broader societal emotions, risk perceptions, and political cleavages. Finally, comparative analysis across multiple national contexts would strengthen the generalizability of the evolutionary institutionalization model proposed here.

Three further limitations should be noted. First, BERTopic operates at the level of aggregate corpus-wide discursive structure; it identifies which actors are discursively prominent. However, it is not designed to trace the strategic intent, competing interests, or entrepreneurial agency of individual actors—applying actor-network theory would help address this gap. Second, our corpus consisted of institutional and elite media discourse (government-, corporate-, and expert-sourced coverage) rather than direct measures of public opinion (e.g., surveys, social media discourse, public consultation records); shifts in underlying public preferences toward RE100 or competing energy pathways are therefore outside the scope of what our media-discourse data can capture, and combining our approach with public-opinion data would be a valuable extension.

Third, approximately 40% of the original corpus ($n = 8394$) was excluded before the core topic modeling as commercial and stock-market-related reporting, based on the lexicon-based criteria described in Table 1. While these criteria were applied systematically and validated through a manual audit of representative headlines, we did not conduct a document-level reassignment of the excluded material onto the final topic structure, nor a comparative analysis of cluster composition with and without this material. Consequently, the cluster proportions, relative onset ordering, and dominance patterns discussed in Sections 4 and 5.1 are descriptive of the filtered analytic corpus and not validated against the full collected corpus. We consider this a substantive rather than incidental limitation given the scale of the excluded subset, and a concrete follow-up analysis—

zero-shot reassignment of the excluded documents into the existing topic space via the trained BERTopic model, followed by a comparison of cluster-level distributions—would extend this study’s findings.

Fourth, a related limitation is the outlier-reassignment step described in Section 3. This step combined the algorithmic output of `reduce_outliers`, whose similarity threshold was not systematically recorded, with subsequent qualitative adjustment by two co-authors based on the social and discursive context of individual articles. The exact reassignment criteria—the unrecorded threshold and the specific qualitative adjustments—cannot be fully reconstructed as a fixed, document-level mapping. Precise reproducibility of the final Topic +1 rate and of the topic assignment of individual reassigned documents is, therefore, limited.

Author Contributions: Conceptualization, H.J.A. and C.J.C.; Methodology, H.J.A.; Software, H.J.A.; Validation, H.J.A.; Formal analysis, H.J.A.; Investigation, H.J.A. and C.J.C.; Data curation, H.J.A.; Writing—original draft, H.J.A. and C.J.C.; Writing—review & editing, C.J.C.; Visualization, H.J.A.; Supervision, C.J.C.; Project administration, C.J.C.; Funding acquisition, C.J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. Jeong, H. Directions for establishing a greenhouse gas inventory in urban planning. *J. Environ. Stud.* **2011**, *50*, 97–113.
2. Yang, A.R.; Seong, E.G. Problems and causes of climate change education in elementary social studies: A case of “A Sustainable Global Village”. *J. Korean Assoc. Geogr. Environ. Educ.* **2022**, *30*, 1–18.
3. Lee, J.W. The study of adaptation model for the hazard by climate change. *Geogr. J. Korea* **2010**, *44*, 635–648.
4. Maslin, M.; Lang, J.; Harvey, F. A short history of the successes and failures of the international climate change negotiations. *UCL Open Environ.* **2023**, *5*, e057. [[CrossRef](#)] [[PubMed](#)]
5. Pereira, J.C.; Viola, E. Climate multilateralism within the United Nations Framework Convention on Climate Change. In *Oxford Research Encyclopedia of Climate Science*; Oxford University Press: Oxford, UK, 2020.
6. Clarke, S.E. Local place-based collaborative governance: Comparing state-centric and society-centered models. *Urban Aff. Rev.* **2017**, *53*, 578–602. [[CrossRef](#)]
7. Dorsch, M.J.; Flachsland, C. A polycentric approach to global climate governance. *Glob. Environ. Politics* **2017**, *17*, 45–64. [[CrossRef](#)]
8. Shin, H.; Park, J.-B. Analysis of Global RE100 Operation Status and Suggestions for Expanding RE100 in South Korea. *Trans. Korean Inst. Electr. Eng.* **2021**, *70*, 1645–1654. [[CrossRef](#)]
9. Arthur, W.B. *Increasing Returns and Path Dependence in the Economy*; University of Michigan Press: Ann Arbor, MI, USA, 1994.
10. Pierson, P. Increasing returns, path dependence, and the study of politics. *Am. Political Sci. Rev.* **2000**, *94*, 251–267. [[CrossRef](#)]
11. Berger, P.L.; Luckmann, T. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*; Anchor Books: New York, NY, USA, 1967.
12. Geels, F.W. *Technological Transitions and System Innovations: A Co-evolutionary and Socio-technical Analysis*; Edward Elgar Publishing: Cheltenham, UK, 2005.
13. Broom, R.; Hanas, I.G. Corporate Power Purchase Agreements—The Preferred Route for Corporates to Secure Renewable Energy Supplies in a Decarbonized World. 2022. Available online: <https://natlawreview.com/article/corporate-power-purchase-agreements-preferred-route-corporates-to-secure-renewable> (accessed on 6 February 2020).

14. Egli, F.; Zhang, R.; Hopo, V.; Schmidt, T.S.; Steffen, B. The contribution of corporate initiatives to global renewable electricity deployment. *Nat. Commun.* **2023**, *14*, 4678. [[CrossRef](#)] [[PubMed](#)]
15. Simoens, M.C.; Fuenfschilling, L.; Leipold, S. Discursive dynamics and lock-ins in socio-technical systems: An overview and a way forward. *Sustain. Sci.* **2022**, *17*, 1683–1699. [[CrossRef](#)]
16. Geels, F.W. From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. *Res. Policy* **2004**, *33*, 897–920.
17. Foucault, M.; Nazzaro, A.M. History, discourse and discontinuity. *Salmagundi* **1972**, *20*, 225–248.
18. Geels, F.W. Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study. *Res. Policy* **2002**, *31*, 1257–1274. [[CrossRef](#)]
19. Geels, F.W.; Schot, J. Typology of sociotechnical transition pathways. *Res. Policy* **2007**, *36*, 399–417. [[CrossRef](#)]
20. Geels, F.W. Regime resistance against low-carbon transitions: Introducing politics and power into the multi-level perspective. *Theory Cult. Soc.* **2014**, *31*, 21–40. [[CrossRef](#)]
21. Joo, Y. News media's political orientation and their framing of climate change in South Korea. *J. Commun. Sci.* **2013**, *13*, 591–626.
22. Lee, T. From nuclear energy developmental state to energy transition in South Korea: The role of the political epistemic community. *Environ. Policy Gov.* **2020**, *31*, 30–42. [[CrossRef](#)]
23. Valentine, S.V.; Sovacool, B.K. The socio-political economy of nuclear power development in Japan and South Korea. *Energy Policy* **2010**, *38*, 7975–7988. [[CrossRef](#)]
24. Gillespie, T.; Boczkowski, P.; Foot, K. *Media Technologies: Essays on Communication, Materiality, and Society*; The MIT Press: Cambridge, MA, USA, 2014.
25. Hoof, F.; Boell, S.K. Culture, technology, and process in 'media theories': Toward a shift in the understanding of media in organizational research. *Organization* **2019**, *26*, 639–654. [[CrossRef](#)]
26. Mamut, P.; Oels, A. The overlooked role of discourse in breaking carbon lock-in: The case of the German energy transition. *Wiley Interdiscip. Rev. Clim. Change* **2019**, *10*, e574. [[CrossRef](#)]
27. Patala, S.; Korpivaara, I.; Jalkala, A.; Kuitunen, A.; Soppe, B. Legitimacy under institutional change: How incumbents appropriate clean rhetoric for dirty technologies. *Organ. Stud.* **2019**, *40*, 395–419. [[CrossRef](#)]
28. Park, S. Discourse analysis of media construction of political reality. *J. Political Sci. Commun.* **2002**, *5*, 209–248. Available online: <https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE01367819> (accessed on 6 September 2026).
29. Geels, F.W. The dynamics of transitions in socio-technical systems: A multi-level analysis of the transition pathway from horse-drawn carriages to automobiles (1860–1930). *Technol. Anal. Strateg. Manag.* **2005**, *17*, 445–476. [[CrossRef](#)]
30. Anderson, A. Source strategies and the communication of environmental affairs. *Media Cult. Soc.* **1991**, *13*, 459–476. [[CrossRef](#)]
31. Boykoff, M.T.; Boykoff, J.M. Balance as bias: Global warming and the US prestige press. *Glob. Environ. Change* **2004**, *14*, 125–136. [[CrossRef](#)]
32. Boykoff, M.T.; Boykoff, J.M. Climate change and journalistic norms: A case study of US mass-media coverage. *Geoforum* **2007**, *38*, 1190–1204. [[CrossRef](#)]
33. Boykoff, M.T. From convergence to contention: United States mass media representations of anthropogenic climate change science. *Trans. Inst. Br. Geogr.* **2007**, *32*, 477–489. [[CrossRef](#)]
34. Boykoff, M.T. *Who Speaks for the Climate? Making Sense of Media Reporting on Climate Change*; Cambridge University Press: Cambridge, UK, 2011.
35. Kim, H.C.; Kim, H.S.; Cho, S.K. Changes in news frames related to climate change in Korean daily newspapers. *Soc. Sci. Res.* **2011**, *19*, 76–106.
36. Choi, E.; Cho, W. Korean Newspapers' Political Orientation and Climate Change Coverage, 2000–2022. *Korean J. Political Inf.* **2023**, *26*, 51–85. [[CrossRef](#)]
37. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019), Minneapolis, MN, USA, 2–7 June 2019; pp. 4171–4186.
38. Grootendorst, M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv* **2022**, arXiv:2203.05794.
39. Ko, Y.; Lee, S.; Cha, M.; Kim, S.; Lee, J.; Han, J.; Song, M. Topic modeling of insomnia social data using BERTopic and development of a deep learning-based automatic classification model. *J. Korean Soc. Inf. Manag.* **2022**, *39*, 111–129.
40. Ham, J.; Choe, Y.J.; Park, K.; Choi, I.; Soh, H. KorNLI and KorSTS: New Benchmark Datasets for Korean Natural Language Understanding. In Proceedings of the Findings of the Association for Computational Linguistics: EMNLP 2020, Online, 16–20 November 2020; pp. 422–430.
41. Kudo, T. MeCab: Yet Another Part-of-Speech and Morphological Analyzer. Available online: <https://taku910.github.io/mecab/> (accessed on 11 June 2026).

42. Park, E.L.; Cho, S. KoNLPy: Korean natural language processing in Python. In Proceedings of the 26th Annual Conference on Human & Cognitive Language Technology, Chuncheon, Republic of Korea, 8–10 October 2014; Volume 6, pp. 133–136.
43. Hansen, A. The Media and the social construction of the environment. *Media Cult. Soc.* **1991**, *13*, 443–458. [[CrossRef](#)]

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.