

Longitudinal Topic Mapping of a National-Language Research Field: Korean Balanced Regional Development Studies in the Korea Citation Index, 2004– 2025

Heonsoo Kim

Korea Social Science Research Lab

Su-Kyeong Park

Chungnam RISE Center

Hee-Su Kim

hskim.1015@gmail.com

Hankuk University of Foreign Studies

Research Article

Keywords: Korea Citation Index, BERTopic, longitudinal science mapping, national-language bibliographic databases, topic stability, regional population decline

Posted Date: October 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11019037/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

National citation indexes capture policy-relevant, national-language scholarship that is often underrepresented in international bibliographic databases. Yet little is known about how such research fields evolve thematically over time. This study maps the evolution of Korean balanced regional development research indexed in the Korea Citation Index (KCI) from 2004 to 2025. Using a screened corpus of 1,127 articles, we fitted a corpus-wide BERTopic model, identified 15 research fronts, and compared their prevalence across four periods; 879 articles were assigned to substantive fronts and 248 were retained as HDBSCAN outliers. Robustness was assessed using alternative corpus definitions, metadata fields, embeddings, clustering settings, period boundaries, and a same-corpus 15-topic LDA baseline. The clearest change was the rise of research on population decline, local extinction, and regional vitality, whose share increased from 1.9% in 2004–2009 to 21.4% in 2021–2025 among BERTopic-classified articles. The upward trajectory persisted across all sensitivity analyses. Notably, a demographic family comprising two LDA topics also rose to 25.5% of all documents despite only modest overlap with the BERTopic front (Jaccard = .294). The findings indicate that demographic concerns became layered onto established traditions of decentralization, innovation, and institutional reform. Methodologically, the study shows that a longitudinal trajectory can remain robust even when alternative models recover different document sets, underscoring the need to report temporal robustness separately from topic-membership stability in national-language science mapping.

1 Introduction

Science maps are shaped by the bibliographic databases from which they are constructed. Web of Science and Scopus provide standardized records and broad international comparability, but their coverage remains uneven across languages, countries, disciplines, and publication systems (Mongeon & Paul-Hus, 2016; Sanz-Casado et al., 2021; Asubiaro, 2026). National citation indexes complement these international databases by capturing scholarship produced for domestic academic and policy communities, including nationally oriented publication channels whose visibility in global indexing systems can vary substantially (Nazarovets et al., 2026). Yet national indexes are not interchangeable sampling frames: the Chinese Social Sciences Citation Index was developed for Chinese-language social-science journals (Su et al., 2001), whereas Japan's Citation Database for Japanese Papers was built from domestic society journals (Negishi et al., 2004). The Korea Citation Index (KCI) likewise provides an observation window onto Korean policy scholarship that may be only partially represented internationally (Kim et al., 2013; Yang et al., 2021). The present study examines this premise through a two-stage Scopus source-screen and article-confirmation diagnostic of the analytical KCI corpus. More broadly, longitudinal science-mapping research has rarely examined the comparability of thematic change under evolving metadata and alternative representations in national-language databases. Addressing that gap can reveal both substantive changes in domestically oriented scholarship and methodological limits that international-database studies may overlook.

Balanced regional development provides a useful setting in which to examine this national knowledge space. Korean balanced-development policy has combined spatial redistribution, regional competitiveness and innovation, decentralization, and, increasingly, responses to population decline and demographic sustainability (Seo, 2009; Choi, 2022). Korean scholarship accordingly spans decentralization, local finance, regional innovation, metropolitan concentration, administrative integration, and demographic sustainability. These concerns are distributed across public administration, law, economics, geography, political science, policy studies, and regional studies rather than concentrated within a single disciplinary specialty. As a result, journal- or discipline-specific reviews may overlook connections among research traditions that address related territorial problems using different conceptual and institutional vocabularies.

Text-based science mapping offers a way to recover these cross-disciplinary relations and trace their evolution over time. Topic-modeling approaches can identify recurring semantic structures across large publication corpora, while embedding-based methods such as BERTopic can capture relations that do not depend solely on exact vocabulary. Previous Korean studies have used computational methods to examine decentralization and balanced-development discourse: Kim and Kim (2024) applied KeyBERT and BERTopic to news discourse, and Jeong and Jun (2025) combined latent Dirichlet allocation with network analysis for balanced regional development in news articles. These studies establish the value of topic modeling for public and media discourse, but they do not map the structure of the scholarly literature, compare common research fronts over a long publication window, or test whether inferred temporal trajectories survive alternative corpus and model specifications. The present study addresses these distinct bibliometric and methodological tasks using KCI-indexed research articles.

What remains less understood is how a national-language scholarly field is thematically organized, how that structure changes over time, and whether apparent longitudinal shifts remain visible under alternative representations of the same corpus. This study therefore asks: (1) Which research fronts organize Korean balanced regional development scholarship indexed in the KCI from 2004 to 2025? (2) How does the prevalence of those fronts vary across four analytical periods? (3) Do the principal longitudinal trajectories, especially the rise of demographic themes, persist across alternative corpus boundaries, metadata fields, embeddings, clustering settings, period definitions, and a same-corpus LDA baseline?

This study contributes in two respects. Substantively, it shows how population decline and local extinction became increasingly central while established institutional and economic traditions persisted. Methodologically, it extends multidimensional topic-model evaluation to longitudinal science mapping in a national-language bibliographic system: rather than asking only whether alternative models recover similar topics, it tests whether temporal conclusions remain stable when overall partition agreement and record-level membership overlap change across corpus, metadata, embedding, clustering, and model-family specifications. Building on model-comparison studies of topic representation and model performance (Egger & Yu, 2022; Benz et al., 2025), this study evaluates temporal robustness in its own right, separately from topic representation and partition similarity.

2 Related work and analytical approach

2.1 Mapping topics in national research systems

Science mapping represents the conceptual, intellectual, and social organization of research through bibliometric relations (Cobo et al., 2011). Established approaches construct and visualize maps from co-citation, bibliographic-coupling, co-word, or other similarity structures, often using tools such as VOSviewer (van Eck & Waltman, 2010). Relations among publications may be defined through citation links, textual similarity, or combinations of the two (Bascur et al., 2025; Xie & Waltman, 2025). In national research systems, the choice among these approaches is also constrained by database infrastructure: citation-based mapping requires temporally consistent record-to-record identifiers, whereas text-based mapping can recover thematic relations across publications that lack interoperable reference links or common disciplinary classifications. This distinction is consequential for national-language databases, whose metadata coverage and standardization may change over time. In the present study, the KCI export did not provide a temporally consistent set of stable record-to-record reference identifiers. Research fronts are therefore defined from textual similarity, while citation information is used descriptively rather than as the basis of a citation network.

Classical probabilistic topic models such as latent Dirichlet allocation (LDA) remain useful baselines for identifying recurring thematic structures (Blei et al., 2003; Griffiths & Steyvers, 2004). Embedding-based approaches such as BERTopic can additionally recover semantic relations that do not depend on exact token overlap (Grootendorst, 2022). Recent field-level applications have used BERTopic alongside classical models to compare topic distinctiveness, coherence, and temporal prevalence (Mishra et al., 2025; Benz et al., 2025), while cross-model comparisons show that different representations need not produce equivalent partitions of the same corpus (Egger & Yu, 2022). Bottom-up topic representations may nevertheless retain meaningful alignment with broader disciplinary structures even when their exact partitions differ from externally defined categories (Hladík & Renisio, 2025). The comparison in this study therefore does not treat agreement with LDA as validation of BERTopic. Instead, a same-corpus, same-k LDA baseline is used to test whether the principal longitudinal trajectory is visible under a different topic representation, while semantic coherence and document-level overlap are reported separately to characterize interpretability and membership similarity.

National-language corpora create an additional challenge because the amount and language composition of available metadata can differ across publication systems (Sanz-Casado et al., 2021; Nazarovets et al., 2026). In a longitudinal corpus, such representation may also change over time. In the present KCI corpus, Korean morphology requires language-aware preprocessing, while English abstracts and keywords become increasingly common in more recent records. The later records may therefore be longer, more multilingual, or semantically richer solely because additional metadata fields are available. Such temporal measurement change can be mistaken for thematic change or can alter cluster membership across periods. We diagnose this threat by examining metadata-field availability and

document length over time and by refitting models after excluding keyword fields and restricting input to Korean-language fields.

Consistent with broader arguments that topic-model quality is multidimensional rather than reducible to a single metric (Ballester & Penner, 2022), a longitudinal topic map is evaluated here along three distinct dimensions: semantic interpretability, partition stability, and temporal robustness. Semantic interpretability concerns whether the terms and representative documents support a coherent substantive label; partition stability concerns whether alternative specifications recover similar document groupings; and temporal robustness concerns whether the direction and magnitude of prevalence change remain visible across those specifications. Partition stability is assessed at two complementary levels: overall partition agreement, summarized by normalized mutual information (NMI), and record-level membership overlap for the focal demographic front, summarized by Jaccard similarity. Semantic interpretability is assessed using topic coherence and evidence-based label review, while temporal robustness is assessed from period-specific prevalence trajectories. None of the three dimensions substitutes for the others: a semantically broad topic can display a robust temporal trajectory, and two models can recover the same directional change despite limited membership overlap. This separation is consistent with evidence that topic-model validation can benefit from stability and corpus-metadata criteria beyond coherence alone (Vázquez et al., 2022). Having outlined the criteria used to evaluate the topic map, we next consider the substantive research traditions that provide context for interpreting the recovered fronts.

2.2 Prior research traditions: innovation, decentralization, and demographic sustainability

For the purposes of interpreting the topic map, prior Korean scholarship points to three recurrent and overlapping policy traditions that are especially relevant to the present analysis. These traditions are not intended as an exhaustive taxonomy of balanced-development research; rather, they summarize recurring substantive orientations in the literature and in Korea's policy evolution. The regional-innovation tradition explains territorial performance through innovation capacity, institutional conditions, and place-specific capabilities, reflecting the longstanding role assigned to regional innovation systems and innovation-city strategies in balanced-development policy (Jang, 2007; Seo, 2009). The decentralization tradition examines how local autonomy, fiscal and organizational authority, spatial planning, and locally led policy implementation shape balanced development (Kim & Jin, 2006; Kim, 2017; Choi, 2022). The demographic-sustainability tradition increasingly reframes population decline and local extinction as problems of regional development, territorial governance, and place viability rather than as demographic change alone (Lim, 2022; Choi & Paek, 2023; Park & Jang, 2024). These traditions provide an interpretive framework for evaluating whether the topic map recovers established lines of scholarship and identifies changes in their relative prominence.

These traditions are analytically distinct but can coexist within the same publication corpus. An increase in the prevalence of demographic-sustainability research does not by itself demonstrate the

displacement of decentralization, innovation, or institutional-reform scholarship. We therefore interpret longitudinal change as a redistribution of classified-document shares across overlapping policy concerns. A demographic reorientation is supported when the demographic front gains sustained prevalence across periods and sensitivity specifications while established fronts remain substantively present. This criterion distinguishes a change in the field's center of attention from the wholesale replacement of an earlier research paradigm.

3 Data and methods

3.1 Retrieval and corpus construction

The corpus was constructed in successive stages to balance retrieval coverage with substantive relevance to Korean balanced regional development. The analyzed KCI export was downloaded on 14 February 2026; records dated 2025 therefore include only articles indexed by that date and may be affected by late indexing. The original account-level query history was not retained. Accordingly, the archived 16-term query was reconstructed from retained project materials and is reported as an approximate retrieval specification rather than as a fully verifiable record of the original search. Corpus-boundary sensitivity analyses described below assess the stability of the principal findings to this retrieval uncertainty.

The initial export contained 3,053 records. Title-based deduplication and metadata-completeness checks removed 143 records, leaving 2,910 unique records. Such preprocessing is consequential because inconsistencies in bibliographic metadata can affect deduplication, record identification, and downstream analysis (Nowakowska, 2025). A preliminary relevance screen of titles, KCI subject classifications, keywords, and abstracts retained 1,939 records. A subsequent semantic screen excluded 812 records that did not establish a substantive relationship between territorial conditions and policy, governance, institutions, public finance, innovation, or demographic sustainability, yielding the 1,127-record primary corpus. The final corpus therefore prioritizes substantive relevance over maximal recall. The preliminary relevance screening was led by one author, with contributions from the other authors, but it was not conducted using a formal independent dual-coding protocol. Accordingly, inter-coder reliability cannot be estimated, and the screening decisions should not be interpreted as independently validated classifications. Online Resource 1, Section S2 reports the operational criteria, exclusion logic, and corpus-boundary checks.

A post-screening diagnostic flagged 246 excluded records that contained strong regional-governance terminology as potential false negatives, not confirmed screening errors. These records were added to the primary corpus to form a targeted sensitivity corpus ($N = 1,373$). A second, broader sensitivity corpus retained all 1,939 records that passed the preliminary screen. The two specifications test different boundary assumptions: the targeted corpus relaxes exclusions for records with strong governance signals, whereas the broad corpus removes the subsequent semantic-screen boundary entirely. Neither

specification substitutes for independent dual coding or establishes the true recall of the primary corpus.

Stable KCI record identifiers were available for linking records across the primary, targeted-sensitivity, and broad screened corpora. These identifiers ensured that the 1,373- and 1,939-record specifications were defined without duplicate or ambiguous record matching. Their use for corpus linkage should be distinguished from the record-to-record reference identifiers required for citation-network construction, which were not consistently available across periods. Table 1 summarizes the corpus boundaries, and Fig. 1 reports the corresponding record flow.

3.1.1 Cross-database diagnostic

A two-stage procedure was used to identify confirmed Scopus overlap. First, the 1,127 KCI records were linked by print or electronic ISSN and publication year to Elsevier’s July 2026 Scopus source list. Exact ISSN matching within the listed coverage years identified 13 journal-year candidates; exact-title-only source matches were reviewed as possible homonyms and were not accepted as evidence of coverage. Second, all 13 candidates were checked at the article level on 8 September 2026 in the institutionally authenticated Scopus document interface. Eight were confirmed by exact DOI, one by exact title, and two by locating the target in the complete source-ISSN and publication-year result set. Two relaxed-title API searches returned hits with Scopus EIDs, but these hits did not establish matches to the target KCI articles; the two candidates were absent from the corresponding complete source-ISSN and publication-year result sets during the 8 September 2026 live recheck and were therefore excluded from the conservative confirmed-overlap count. The resulting count is a conservative confirmed-overlap measure, not an estimate of Scopus coverage or its false-negative rate, and it does not construct an alternative Scopus corpus.

Table 1
Corpus construction and analytical boundaries

Stage	Records	Change	Role
Retrieved corpus	3,053	—	Initial KCI retrieval
Eligible records after deduplication	2,910	−143	Deduplicated corpus
Broad screened corpus	1,939	−971	Preliminary relevance screen
Primary analytical corpus	1,127	−812	Main analysis
Targeted sensitivity corpus	1,373	+ 246	Adds 246 potential false negatives

3.2 Text preparation and the 15-front BERTopic model

The modeling strategy first identifies a common set of research fronts across the full 2004–2025 corpus and then compares their prevalence across periods. This common-partition design allows temporal

change to be interpreted as changes in the prevalence of comparable research fronts rather than as changes in independently estimated period-specific topic systems. For the primary specification, seven metadata fields were concatenated when available: Korean title, Korean author keywords, Korean abstract, foreign-language abstract, English abstract, foreign-language keywords, and English keywords, which are stored as distinct metadata fields in the KCI export. The resulting concatenated text was embedded directly with paraphrase-multilingual-MiniLM-L12-v2 (Reimers & Gurevych, 2019). KoNLPy 0.6.0 with Okt was used for token-based topic representation and coherence analysis rather than for generating the sentence embeddings (Park & Cho, 2014). UMAP used `n_neighbors = 30`, `n_components = 5`, `min_dist = 0.0`, cosine distance, and `random_state = 42`. HDBSCAN used `min_cluster_size = 20`, `min_samples = 1`, Euclidean distance, and EOM selection; low-density documents remained topic - 1 (McInnes et al., 2017; McInnes et al., 2018). Topic representation used CountVectorizer with `min_df = 5` and 1–2-grams followed by c-TF-IDF.

The main BERTopic specification produced 15 substantive research fronts, classifying 879 documents and assigning 248 documents (22.0%) to the HDBSCAN outlier category. The number of fronts was an outcome of the specified UMAP and HDBSCAN configuration rather than a value fixed in advance; the 15-topic LDA model was subsequently set to the same `k` to provide a comparably granular baseline. A single corpus-wide partition was estimated before period-specific prevalence was calculated. BERTopic's dynamic-topic mode was therefore not used, because the analysis compares the prevalence of a common set of research fronts across periods rather than estimating period-specific topic structures. Classical dynamic topic models, by contrast, allow topic-word distributions to evolve smoothly from period to period (Blei & Lafferty, 2006). Such models are well suited to tracking gradual within-topic drift, but the sparse earliest period (Period I, `N = 134`) and the emphasis on a common cross-period partition favored a simpler design. The common-partition approach trades finer-grained within-topic drift for direct comparability across periods and with the same-`k` LDA baseline. Topic labels should be read as concise interpretive summaries of document clusters rather than as mutually exclusive theoretical categories. Table 2 summarizes the main model configuration.

LDA was selected as a deliberately contrasting classical probabilistic baseline rather than as part of an exhaustive model-comparison exercise; the purpose was to test whether the principal temporal trajectory survives a change in model family while holding the corpus and topic count constant. For the 15-topic LDA baseline, the same 1,127 concatenated text records were tokenized with KoNLPy Okt, retaining nouns, verbs, adjectives, and alphabetic tokens after Korean and English stopword removal. Tokens occurring in fewer than three documents or more than 60% of documents were filtered. Gensim LDA used `k = 15` to match the granularity of the BERTopic solution, `random_state = 42`, 10 passes, 200 iterations, and automatic alpha and eta; each document received its highest-probability topic. The population/decline/extinction lexicon and the rule defining a demographic family—at least three lexicon items among a topic's top 20 terms—were defined by inspecting the fitted LDA topics' top terms, necessarily after model fitting but before any period-specific prevalence was computed. We also report the single LDA topic with maximum Jaccard overlap with BERTopic T4 so that the family rule cannot mask a weak one-topic match.

Topic labeling was LLM-assisted but human-controlled. Gemini 2.5 Flash generated candidate wording from a fixed prompt using cluster size, the top 20 c-TF-IDF terms, representative author keywords, and title/abstract evidence windows of $k = 10, 15,$ and 20 . The authors compared each candidate label with the corresponding terms, keywords, and evidence windows, revised labels that overstated or omitted recurrent content, and finalized all labels by consensus. Candidate outputs and supporting label-review materials were retained to document the labeling procedure. LLM assistance did not affect embeddings, clustering, document assignments, or temporal statistics. Because generative models can introduce framing bias, fluent but unsupported labels, and inconsistent suggestions across prompts, the labels are treated as interpretive summaries rather than independently validated classifications; no external domain-expert panel assessed them.

For each research front, the classified-document share was regressed on the ordered period index 1–4. The coefficient is expressed in percentage-point change per one-step movement between analytical periods, not per calendar year; because the four windows contain six, six, five, and five years, it should not be interpreted as a constant annual rate. Document-bootstrap 95% confidence intervals quantify uncertainty in this descriptive slope. The slopes summarize temporal direction and magnitude within the chosen periodization rather than providing categorical tests of emergence or decline.

Documents assigned to HDBSCAN topic – 1 were treated as unclassified observations rather than as substantively irrelevant records. They were excluded from the classified-document denominator and from front-specific slope estimation, but retained in the alternative all-document prevalence denominator and in diagnostics of outlier rates across periods. Both denominators are reported to show whether temporal conclusions depend on assignment coverage. Topic labels were assigned after clustering, and period information was not used in model fitting or labeling.

Table 2
Main BERTopic model configuration

Component	Specification
Runtime	Python 3.12.13
Embedding	paraphrase-multilingual-MiniLM-L12-v2; sentence-transformers 5.2.0
Input fields	Korean title/keywords/abstract + foreign/English abstracts and keywords
Projection	UMAP 0.5.1; n_neighbors = 30; n_components = 5; min_dist = 0.0; metric=cosine; random_state = 42
Clustering	hdbscan 0.8.39; min_cluster_size = 20; min_samples = 1; metric=euclidean; selection = eom
Topic representation	CountVectorizer min_df = 5; ngram_range=(1,2); c-TF-IDF
Topic model	BERTopic 0.17.4
Numerical stack	scikit-learn 1.6.1; NumPy 1.26.4; pandas 2.2.3
Korean processing	KoNLPy 0.6.0 / Okt
Longitudinal design	One global BERTopic partition; period prevalence computed after fitting; dynamic mode not used
LDA baseline	Gensim 4.4.0; k = 15; seed = 42; passes = 10; iterations = 200; alpha/eta=auto; Okt tokens
Coherence	Top 10 terms; Okt corpus with adjacent bi-/trigrams; C_v and NPMI for every BERTopic and LDA topic

3.3 Temporal inference and sensitivity analyses

Temporal analysis addressed two related questions: how the prevalence of each research front changed over time and whether the principal demographic trajectory remained visible under alternative corpus and modeling choices. The primary periods were defined as 2004–2009, 2010–2015, 2016–2020, and 2021–2025. They serve as descriptive aggregation windows, not estimated structural breaks, policy-regime boundaries, or causal intervention points. Periodization sensitivity was evaluated separately by moving 2010 into the first window, using three approximately equal windows (2004–2010, 2011–2017, and 2018–2025), and excluding 2025 from the final window. These checks assess whether the direction of the demographic trajectory depends on the chosen temporal aggregation.

Semantic coherence was calculated for all 15 BERTopic fronts and all 15 LDA topics using the top 10 terms, Okt-tokenized corpus text augmented with adjacent bi- and trigrams, and both C_v and NPMI (Röder et al., 2015). Coherence evaluates term co-occurrence and interpretability; it does not validate temporal ordering or establish that two models recover identical document sets.

The primary outcome is each front's share of classified documents within a period. Because density-based clustering leaves some documents unassigned, prevalence is also reported using all documents in each period as the denominator. For the population-decline/local-extinction front, nonparametric 95% confidence intervals were obtained by resampling documents with replacement within each period 10,000 times and recalculating the corresponding prevalence. A Cochran–Armitage test assessed monotonic change across the four ordered periods using document-level front membership as the binary outcome. Both procedures treat documents as independent observations. They quantify sampling uncertainty and ordered association within the observed corpus, not uncertainty from corpus construction, model fitting, repeated authorship, or causal policy effects.

Sensitivity analyses targeted four distinct sources of instability. Corpus-boundary checks added the 246 potential false negatives and refitted all 1,939 preliminarily screened records. Metadata-representation checks omitted keyword fields and restricted input to Korean-language fields. Model-specification checks used a Korean-specific embedding, five UMAP seeds (42, 7, 99, 123, and 2024), HDBSCAN `min_cluster_size` values of 15 and 25, an alternative 14-cluster solution, and a same-corpus 15-topic LDA baseline. The five seed runs held the sentence embeddings and HDBSCAN specification fixed and varied only UMAP `random_state`; seed 42 was the main specification. Temporal-aggregation checks excluded 2025 and applied alternative period boundaries. Each specification was used to test whether the principal demographic trajectory retained its direction under the corresponding perturbation; topic identifiers were treated as solution-specific.

An author-keyword co-word analysis, following the descriptive co-word tradition (Callon et al., 1991), is reported in Online Resource 1, Section S6. Because author keywords are also included in the BERTopic input, this analysis is treated as descriptive rather than as independent validation.

For sensitivity analyses involving alternative metadata fields, corpus definitions, or embeddings, models were refitted independently rather than projected through the main embedding and clustering geometry. For specifications fitted to the same records, the corresponding demographic cluster was the alternative cluster with the largest Jaccard overlap with BERTopic T4. For specifications with expanded corpora, Jaccard overlap was calculated only on record identifiers shared with the primary corpus, while prevalence used the alternative model's full corpus, assignments, and period-specific denominators. The maximum-overlap rule identifies the closest membership counterpart rather than proving semantic equivalence; candidate terms and representative records were therefore reviewed to confirm that the matched cluster retained a demographic interpretation. Jaccard overlap measures membership similarity, whereas temporal robustness is evaluated from each matched cluster's period-specific prevalence trajectory.

4 Results

4.1 Growth and disciplinary composition

The final corpus spans 2004–2025 and contains 1,127 KCI-indexed articles on Korean balanced regional development. Period totals increased from 134 articles in 2004–2009 to 270 in 2010–2015, 295 in 2016–2020, and 428 in 2021–2025; the corresponding annual means were 22.3, 45.0, 59.0, and 85.6 articles. Figure 2 displays the annual publication trajectory. These counts indicate sustained growth in publication output within the constructed corpus, including after accounting descriptively for the shorter final two periods. Because the KCI export was downloaded on 14 February 2026, the 2025 count includes only records indexed by that date and may be affected by late indexing.

The corpus spans public administration, law, economics, geography, political science, regional studies, and related fields. T4 was likewise distributed across KCI subject classifications: its largest subject counts were international/regional development (10), public administration (9), law (9), economics (9), and geography (9), and its largest single-journal count was six. These distributions show that no single listed subject category or journal accounts for a large share of T4, although differences in KCI classification practices and field-specific outlier rates limit stronger claims about disciplinary representation. Full subject, journal, and outlier profiles are reported in Online Resource 1. Reference lists were available for 92.3%–100% of records across periods, but stable record-to-record reference identifiers were not consistently available; citation information is therefore used descriptively rather than for network construction.

4.2 Cross-database diagnostic

The two-stage diagnostic confirmed 11 of the 1,127 KCI records as Scopus documents (0.98%). The confirmed proportions were 2/134 (1.49%) in Period I, 1/270 (0.37%) in Period II, 2/295 (0.68%) in Period III, and 6/428 (1.40%) in Period IV. One of 101 T4 records (0.99%) and 10 of the other 1,026 records (0.97%) were confirmed. These are observed confirmed-overlap proportions under the stated procedure. They are not estimates of database coverage, and the small number of confirmed records does not establish equivalence between topic or period groups.

4.3 The 15 research fronts

The BERTopic solution organizes 879 classified documents into 15 research fronts; the remaining 248 documents were retained as HDBSCAN topic – 1 and are not included in the percentages reported in Table 3. The four largest fronts by classified-document count are decentralization and comparative local governance (T0), regional innovation systems (T1), fiscal decentralization (T2), and constitutional/legal reform (T3). The population-decline, local-extinction, and regional-vitality front (T4) ranks fifth by overall classified-document share but has the largest positive descriptive slope across the four ordered periods. Thus, overall size and longitudinal change answer different questions: established institutional and economic fronts remain largest across the full window, whereas T4 exhibits the strongest temporal increase. Topic identifiers are specific to each fitted solution and should not be interpreted as equivalent across alternative model specifications.

The 15-front solution separates several clusters that share demographic or decentralization vocabulary but differ in their recurring terms and representative documents. T4 combines Korean population

decline, local extinction, and place-based vitality policies; T6 emphasizes comparative population decline, ageing, and low fertility; and T11 links youth migration to employment and regional settlement. Among institutional fronts, T0 centers on broad and comparative local governance, T2 on fiscal decentralization, T3 on constitutional and legal reform, T10 on inter-local cooperation and administrative integration, and T13 on institutions and perceptions of local autonomy. These distinctions are interpretive summaries supported by c-TF-IDF terms, author keywords, and representative title/abstract evidence rather than tests that the fronts are theoretically exclusive. The map therefore differentiates recurrent policy mechanisms, legal arrangements, economic systems, and demographic problems while allowing substantive overlap among clusters.

Table 3
BERTopic 15-front solution (classified N = 879)

ID	Research front	n	%	Slope [95% CI]
T0	Decentralization and Comparative Local Governance	146	16.6	-2.27 [-4.80, + 0.24]
T1	Regional Innovation Systems and Industrial Clusters	123	14.0	-1.09 [-3.64, + 1.33]
T2	Fiscal Decentralization and Local Public Finance	116	13.2	-1.25 [-3.60, + 0.92]
T3	Constitutional and Legal Reform for Local Autonomy	106	12.1	+ 0.03 [-1.99, + 1.89]
T4	Population Decline, Local Extinction, and Regional Vitality	101	11.5	+ 6.12 [+ 4.50, + 7.75]
T5	Firms, Manufacturing Location, and Regional Economies	60	6.8	-2.40 [-4.43, -0.55]
T6	Comparative Population Decline and Ageing	37	4.2	+ 1.70 [+ 0.50, + 2.81]
T7	Metropolitan Governance, Urban Regeneration, and Entrepreneurship	27	3.1	-0.68 [-2.22, + 0.66]
T8	Chinese Central-Local Relations and Decentralization	27	3.1	-0.22 [-1.26, + 0.70]
T9	Regional Development Policy and Mega-Regional Economies	25	2.8	-1.78 [-3.62, -0.13]
T10	Inter-Local Cooperation and Administrative Integration	25	2.8	+ 0.11 [-1.07, + 1.15]
T11	Youth Migration, Employment, and Regional Settlement	23	2.6	+ 1.88 [+ 1.16, + 2.70]
T12	Evaluation of Balanced Development Programs and Indicators	22	2.5	+ 0.63 [-0.46, + 1.59]
T13	Local Autonomy Institutions and Decentralization Perceptions	21	2.4	+ 0.07 [-1.08, + 1.07]
T14	Administrative Organization, Functional Assignment, and Performance	20	2.3	-0.85 [-2.11, + 0.20]

Note. n and % are based on the 879 documents assigned to substantive BERTopic fronts; 248 topic – 1 documents are excluded. Percentages describe each front’s share of classified documents across the full 2004–2025 corpus. Slope is the percentage-point change in classified-document share for a one-step movement across the ordered periods 2004–2009, 2010–2015, 2016–2020, and 2021–2025; it is not an annual rate. Brackets report document-bootstrap 95% confidence intervals. Topic IDs are solution-specific and must not be compared by number across different fitted solutions.

4.4 Temporal redistribution toward demographic decline

The clearest longitudinal change is the rise of the population-decline, local-extinction, and regional-vitality front. Its classified-document share increased from 2 of 107 records in Period I (1.9%) to 10 of 202 in Period II (5.0%), 18 of 238 in Period III (7.6%), and 71 of 332 in Period IV (21.4%). Document-bootstrap 95% confidence intervals were 0.0–4.8%, 2.1–8.1%, 4.3–11.2%, and 17.1–25.9%, respectively; the wide first-period interval reflects the small count of two documents. The Cochran–Armitage test indicates a strong monotonic association between ordered period and T4 membership within the observed classified corpus ($z = 6.77$, $p = 1.32 \times 10^{-11}$). Descriptively, T4 moved from a marginal share in the earliest period to a prominent late-period research concern. This result establishes temporal redistribution in the corpus, not a causal effect of demographic or policy change on scholarly agendas.

Using all documents rather than only classified documents as the denominator yielded period shares of 1.5%, 3.7%, 6.1%, and 16.6%, and the ordered association remained strong ($z = 6.69$, $p = 2.20 \times 10^{-11}$). The absolute shares are lower because topic – 1 documents remain in the denominator, but the ordering across periods and the pronounced increase in Period IV are preserved. Period IV contains 71 of T4’s 101 documents, showing that most documents assigned to this front were published after 2020. Agreement between the two denominator definitions indicates that the upward pattern is not explained solely by whether HDBSCAN outliers are excluded from prevalence calculations; it does not rule out other sources of model or corpus uncertainty.

T4 has the strongest positive descriptive slope, increasing by 6.12 percentage points per ordered period step (document-bootstrap 95% CI 4.50–7.75). T5 and T9 have negative slopes whose intervals exclude zero, whereas T6 and T11 have smaller positive slopes whose intervals exclude zero; T0 has a negative point estimate with an interval that crosses zero. Figure 3 compares the T4 trajectory with selected fronts.

Because the Cochran–Armitage test treats documents as independent, repeated publication by the same authors may reduce the effective information in the corpus. Author-name matching, which cannot fully resolve homonyms or name variants, identified 148 distinct author strings across T4’s 101 documents. The most prolific author string appeared in 3 T4 documents, and 23 strings (15.5%) appeared in at least 2; the corresponding classified-corpus figures were 1,014 distinct strings across 879 documents, a maximum of 10 documents, and 18.1% repeat contributors. These descriptive comparisons do not indicate unusual concentration within T4, but they are not a cluster-robust test of author dependence. A separate permutation procedure randomly assigned the T4 label to 101 of the 879 classified documents while holding period sizes fixed and recomputed the Cochran–Armitage statistic over 10,000 iterations. The null distribution was centered near zero (mean = -0.01 , SD = 1.00), and no permutation produced a statistic as extreme as the observed $z = 6.77$ (0/10,000; plus-one Monte Carlo $p = .00010$). This permutation result evaluates whether the observed period concentration exceeds random assignment conditional on T4 size and period totals; it does not address model-selection,

corpus-boundary, or author-clustering uncertainty. Full counts and permutation details are reported in Online Resource 1, Section S16.

Table 4
Population-decline/local-extinction front by period, with two denominators

Period	Front n	Classified N	% class.	95% CI	All N	% all	95% CI	Annual output
I (2004-09)	2	107	1.9	0.0–4.8	134	1.5	0.0–3.7	22.3
II (2010-15)	10	202	5.0	2.1–8.1	270	3.7	1.5–6.3	45.0
III (2016-20)	18	238	7.6	4.3–11.2	295	6.1	3.4–8.8	59.0
IV (2021-25)	71	332	21.4	17.1–25.9	428	16.6	13.1–20.1	85.6
Note. Front n is the number of T4 documents. Classified N excludes HDBSCAN topic – 1 documents, whereas All N includes them. Percentages and document-bootstrap 95% confidence intervals are calculated separately with the corresponding period-specific denominator. Annual output is the mean number of all corpus records per calendar year within each unequal-length period (6, 6, 5, and 5 years).								

4.5 Robustness of the demographic trajectory

Sensitivity refits. The demographic trajectory remained upward across changes in corpus boundaries, metadata representation, and model specification, but the degree of document-level correspondence varied. Table 5 summarizes five independently refitted specifications. Relative to BERTopic T4, the matched fronts had Jaccard overlaps ranging from .399 to .754. Their classified-document shares increased from 3.0% to 25.4% without keyword fields, from 2.9% to 24.5% with Korean fields only, from 4.5% to 27.6% when 246 potential false negatives were included, from 3.4% to 23.9% in the broader 1,939-record corpus, and from 1.0% to 33.6% with the Korean-specific embedding. These results support robustness of the upward temporal direction across the tested perturbations, not equivalence of cluster membership or prevalence magnitude.

The Korean-specific embedding showed the weakest membership correspondence with the main demographic front (Jaccard = .399), despite preserving a pronounced upward trajectory. This contrast further indicates that temporal direction is considerably more stable than the exact set of documents assigned to the demographic front.

Partition diagnostics. A fresh five-seed rerun held the sentence embeddings and HDBSCAN settings fixed while varying UMAP random_state. Seed 42 reproduced the main 15-front partition exactly up to numeric cluster labels (ARI = 1.000; NMI = 1.000). Across seeds 7, 99, 123, and 2024, the solutions contained 14–17 substantive clusters and 20.8–27.0% outliers; relative to the main 15-front partition,

ARI ranged from .609 to .701 and NMI from .805 to .822. The maximum-overlap demographic counterparts had Jaccard values of .541–.804. Their classified-document shares rose from 1.0–4.2% in Period I to 17.2–23.4% in Period IV, and every ordered trend was positive (Cochran–Armitage $z = 6.15–7.22$, all $p < .001$). Changing HDBSCAN minimum cluster size and the embedding model produced additional partition variation, but the demographic trajectory remained upward. Thus, temporal direction was more stable than exact partition membership. Full ID-level assignments and cross-specification summaries are provided in Online Resource 1 and the reproducibility materials.

Metadata and periodization checks. Metadata and assignment diagnostics do not show a simple temporal artifact that parallels the T4 increase. English-metadata availability rose from 89.6% for abstracts and 88.8% for keywords in Period I to 99.8% for both in Period IV, and median document length increased from 475.5 to 532 tokens. However, classified and outlier documents had similar median lengths (518 versus 521; Mann–Whitney $p = .740$), while outlier rates varied non-monotonically across periods (20.2%, 25.2%, 19.3%, and 22.4%). Excluding 2025 also left T4 prevalence essentially unchanged: $54/254 = 21.26\%$ for 2021–2024 versus $71/332 = 21.39\%$ for 2021–2025 among classified documents, with corresponding all-document shares of 16.56% and 16.59%. Alternative period boundaries retained ordered associations ($z = 6.55$ and 5.84 ; both $p < .001$). Together, these diagnostics address metadata completeness, document length, outlier allocation, late indexing, and periodization separately; they leave open any effects of metadata change on the exact fitted partition.

Table 5
Targeted sensitivity refits for the demographic front

Specification	Clusters	Classified	Jaccard	Period I %	Period IV %
No keyword fields	14	894	0.610	3.0	25.4
Korean fields only	15	936	0.754	2.9	24.5
Include 246; N = 1,373	19	1158	0.555	4.5	27.6
Broad prefilter refit; N = 1,939	25	1623	0.621	3.4	23.9
ko-sroberta-multitask	12	895	0.399	1.0	33.6

Note. Clusters is the number of substantive clusters produced by each independently refitted specification; HDBSCAN topic – 1 is excluded. Classified is the number of records assigned to those clusters. Period I and Period IV percentages use each specification’s own period-specific classified-document denominators. Jaccard compares membership in the matched demographic cluster with BERTopic T4; for expanded corpora, it is calculated only over records shared with the primary corpus. Jaccard measures membership overlap and does not establish semantic equivalence.

4.6 LDA baseline and semantic coherence

The same-corpus LDA baseline tests whether the demographic rise remains visible under a different topic representation; it does not treat LDA agreement as validation of BERTopic. LDA did not reproduce

BERTopic T4 as a single topic. The closest single topic was LDA T7 (n = 86; overlap = 37; Jaccard = .247), whose all-document share increased from 2.2% in Period I to 11.9% in Period IV (z = 4.78, p < .001). Under this fitted-topic lexicon rule, demographic terms were distributed across LDA T2 and T7. Their combined family contained 154 documents, shared 58 of the 101 BERTopic T4 documents (Jaccard = .294), and increased from 3.7% to 5.6%, 8.5%, and 25.5% across the four periods (z = 8.24, p < .001; all-document denominator, Table 6). On the same all-document basis, BERTopic T4 reached 16.6% in Period IV; its separately reported 21.4% value uses the classified-document denominator in Table 4. The two models therefore agree on the upward temporal direction while differing substantially in the documents and semantic breadth assigned to the demographic domain.

One LDA topic outside the demographic family, T6 (n = 40), had poor coherence (C_v = .273, NPMI = −.208) and several German function words among its top terms, reflecting a small subset of records with German text in the foreign-language metadata fields (Online Resource 1, Table S20). T6 was neither part of the demographic family (T2 + T7) nor the closest single-topic match to BERTopic T4, so this mixed-language artifact did not affect the main LDA comparison.

Across all 15 fronts, BERTopic mean coherence was C_v = .706 and NPMI = .128, with ranges of .367–.905 and −.120–.263, respectively. T4 had comparatively low coherence (C_v = .368; NPMI = −.082), similar to T9 (C_v = .367; NPMI = −.120), whereas the 15 LDA topics averaged C_v = .749 and NPMI = .066. The low T4 values indicate weak co-occurrence among its top terms and therefore limit claims that it constitutes a narrowly bounded semantic topic. Representative terms, author keywords, and document evidence nevertheless support an interpretive label spanning population decline, local extinction, and multiple place-vitality instruments. Coherence does not evaluate temporal ordering, so the front’s low semantic compactness and its robust upward prevalence trajectory are separate findings. T4 is therefore best interpreted as a broad policy-oriented research domain whose temporal expansion is more stable than its exact semantic boundary.

Table 6
Same-corpus 15-topic baseline for the demographic trajectory; all-document denominator

Model/unit	n	I % all	II % all	III % all	IV % all	Jaccard vs T4
BERTopic T4	101	1.5	3.7	6.1	16.6	1.000
LDA best single topic T7	86	2.2	3.3	7.8	11.9	.247
LDA demographic family T2 + T7	154	3.7	5.6	8.5	25.5	.294
Note. All percentages in Table 6 use the same all-document denominators (Period I–IV N = 134, 270, 295, and 428). BERTopic T4 classified-document shares are reported separately in Table 4. Jaccard compares document membership with BERTopic T4.						

5 Discussion

5.1 A demographic reorientation of the policy field

The main finding is not simply that one topic became larger. Within the constructed corpus, the classified-document share of research on population decline, local extinction, and regional vitality increased markedly, while decentralization, regional innovation, fiscal decentralization, and institutional reform remained substantial across the full period. This pattern supports an interpretation of demographic reorientation through thematic layering rather than wholesale replacement. The upward direction persisted across alternative representations, although the magnitude of change and the exact semantic and document-membership boundaries of the demographic front remained model dependent.

Substantively, the document evidence within T4 links demographic sustainability with youth migration, employment, local finance, service provision, institutional capacity, and place-vitality policies. These recurring associations indicate that population decline was increasingly discussed as a territorial governance problem rather than solely as demographic change. The demographic front did not absorb every population-related study: a separate comparative population-decline and ageing front remained identifiable, while youth migration, employment, and regional settlement formed a smaller front. The solution therefore distinguishes a broad Korean place-based policy domain from comparative demographic research, while the comparatively low coherence of T4 cautions against treating these boundaries as theoretically exclusive.

Established literatures persisted: decentralization, innovation systems, fiscal decentralization, and legal reform remained the four largest fronts across the full corpus. Their continuing scale, alongside the late-period expansion of T4, is consistent with layering rather than replacement. Population decline became a prominent late-period lens while administrative, legal, and economic questions continued to structure a large share of the classified literature. This interpretation describes the redistribution of research attention within the corpus and does not establish that changes in the policy environment caused the observed publication pattern. Domestic text-mining studies likewise report increasing attention to balanced development and decentralization in public discourse (Kim & Kim, 2024; Jeong & Jun, 2025), while the present study documents a corresponding pattern in KCI-indexed scholarly publications.

The distinction between share and volume is important for interpreting the field's development. Annual output within the corpus rose from 22.3 articles in Period I to 85.6 in Period IV, while demographic research also captured a larger share of classified documents. T4's subject and journal profiles show that the front spans law, public administration, economics, geography, regional studies, and related areas rather than forming a single disciplinary enclave, although KCI classification practices and field-specific outlier rates limit stronger claims about disciplinary representation. Comparable shrinkage debates elsewhere motivate future comparative designs (Wiechmann & Pallagst, 2012); evaluating whether similar research-field reorientations occur across countries would require harmonized retrieval and metadata rules.

5.2 Methodological implications for national-language science mapping

First, the three evaluation dimensions specified in Section 2.1 should remain distinct in longitudinal topic mapping: semantic interpretability, partition stability, and temporal robustness. In this study, the demographic trajectory remained upward across corpus, metadata, embedding, clustering, and model-family specifications even though matched clusters varied in semantic compactness and membership. Within partition stability, NMI summarized overall partition agreement across complete classified partitions, while Jaccard measured record-level membership overlap for the matched demographic counterpart. The evidence therefore supports temporal robustness more strongly than exact partition recovery.

Second, longitudinal analyses of national-language databases should treat metadata availability and language composition as potential sources of temporal measurement change. In the KCI corpus, English abstracts and keywords became more prevalent in later periods, increasing the amount and multilingual character of the modeled text. Re-estimating models without keyword fields and with Korean-language fields only tests whether the direction of an observed trajectory depends on those representation choices. Such refits do not prove that metadata change has no effect on the partition; rather, they isolate whether the principal longitudinal inference survives specified changes in the available text.

Third, longitudinal density-based topic models should report how unclassified observations enter each analysis. Here, HDBSCAN topic – 1 documents were excluded from front-specific classified-document shares and slope estimates but retained in all-document prevalence and period-specific outlier diagnostics. Reporting both denominators showed that the upward T4 ordering was not explained solely by excluding outliers and clarifies how assignment coverage enters the prevalence estimates.

Fourth, the cross-database diagnostic provides a complementary sampling-frame check. Confirmed Scopus overlap was low for both T4 and the remaining corpus. The 11 confirmed records are interpreted only as conservative confirmed overlap under the two-stage procedure rather than as a coverage estimate or a substitute for a separately constructed Scopus corpus. More broadly, differences in regional and disciplinary representation across bibliographic databases can yield different bibliometric sampling frames (Kumpulainen & Seppänen, 2022; Asubiaro, 2026).

Fifth, the implications extend beyond the Korean case, but transfer requires adaptation to each database's metadata structure and language environment. National citation indexes with evolving field coverage, mixed-language records, or incomplete reference linking face related threats to longitudinal comparability (Kim et al., 2013; Yang et al., 2021). Cross-database comparisons also show material differences in reference coverage across major bibliographic infrastructures (Culbert et al., 2025), while evidence from nationally oriented and national-language journals shows substantial variation in international-index visibility across publication contexts (Sanz-Casado et al., 2021; Nazarovets et al., 2026). The present workflow offers a transferable diagnostic checklist: report classified and all-

document denominators, refit models under alternative metadata-field and language restrictions, vary corpus boundaries and clustering specifications, compare temporal trajectories with partition-stability measures, and retain documentation of the interpretive labeling procedure. For KCI, more consistent record-to-record reference identifiers and explicit language codes for foreign-language abstract and keyword fields would improve future citation-based and multilingual analyses. The checklist supports transparent sensitivity analysis without treating a single fitted partition as a definitive representation of a national research system.

5.3 Limitations

Several limitations define the appropriate scope of interpretation. First, retrieval reproducibility is incomplete because the original account-level KCI query history was not retained; the archived 16-term query was reconstructed from retained project evidence and should therefore be treated as an approximate retrieval specification. Separately, coverage of the final year may be incomplete because records dated 2025 include only articles indexed by the 14 February 2026 export date. Corpus-boundary refits and the exclusion-of-2025 analysis assess the sensitivity of the main trajectory to these two uncertainties, but they cannot reconstruct the original query history or recover records not yet indexed.

Second, corpus construction involves unresolved screening uncertainty. The relevance screening was led by one author, with contributions from the other authors, but no formal independent dual-coding protocol was implemented. Consequently, inter-coder reliability and systematic disagreement cannot be estimated. The diagnostic set of 246 excluded records with strong regional-governance terminology identifies potential false negatives rather than confirmed screening errors. Targeted and broader-corpus refits show that the upward demographic trajectory survives two relaxed corpus boundaries, but they do not establish the sensitivity, specificity, or recall of the primary screening procedure and do not substitute for independent dual coding.

Third, the reconstructed keyword-based retrieval may omit relevant studies that do not use the selected terms, and title-first deduplication is less reliable than identifier-first matching when DOI or KCI record identifiers are incomplete. These limitations affect the boundaries of the KCI corpus itself.

Fourth, external coverage was assessed only through the Scopus diagnostic. An equivalent Web of Science diagnostic was outside the scope of the present study, and no inference about Web of Science coverage is made. The two-stage Scopus procedure depends on consistent source identifiers and bibliographic metadata, and its false-negative rate was not estimated; Scopus-indexed records may therefore remain unidentified. In addition, the diagnostic does not reconstruct a separate Scopus corpus under harmonized retrieval rules and therefore cannot determine whether the observed trajectory would differ across database systems.

Fifth, topic boundaries and statistical uncertainty remain model dependent. Alternative embeddings and clustering settings did not reproduce identical partitions, and the demographic front had comparatively low semantic coherence. The principal inference therefore concerns the persistence and direction of the

temporal trajectory rather than exact topic membership, a uniquely bounded semantic category, or a model-invariant prevalence magnitude. Document-bootstrap intervals and Cochran–Armitage tests condition on the observed corpus and fitted assignments and treat documents as independent. They do not propagate uncertainty from retrieval, screening, model selection, repeated authorship, or labeling, and they do not establish causal effects of policy change on research agendas.

Sixth, topic labeling was LLM-assisted and remains interpretive. Gemini-generated candidates were reviewed and finalized by the three authors, without an independent domain-expert panel or a formal revision-rate log. Only cluster-level evidence already present in the KCI export—top terms, keywords, and title/abstract excerpts—was submitted to the proprietary cloud-hosted model. Commercial model updates may prevent exact reproduction of candidate wording, although embeddings, clustering, document assignments, and temporal statistics do not depend on the LLM. The retained prompts, candidate outputs, and supporting materials document the procedure but do not constitute independent validation of the final labels.

Seventh, the auxiliary analyses answer narrower questions than the main topic model. The author-keyword co-word analysis describes lexical associations but is not an independent validation because those keywords also enter the BERTopic representation. As noted in Section 2.1, the available reference metadata did not support consistent record-to-record citation-network construction. The Conservative Relevance Score reported in Online Resource 1 is exploratory and domain specific and should not be interpreted as a validated screening classifier. Consequently, these analyses supplement interpretation but do not resolve topic validity, citation structure, or retrieval accuracy.

Eighth, the framework is empirically evaluated in one national bibliographic system and one substantive policy field. Its transferability across other KCI domains or national citation indexes remains to be established.

6 Conclusion

This study maps how Korean balanced regional development scholarship in the KCI changed from 2004 to 2025. The clearest substantive shift was the growing prominence of population decline, local extinction, and regional vitality, while established traditions of decentralization, regional innovation, fiscal decentralization, and institutional reform remained substantial. The pattern is therefore best understood as demographic concerns becoming layered onto an existing policy-research structure rather than replacing earlier traditions. The demographic front also spans multiple subject areas and journals, suggesting that this research domain is not well represented by a single disciplinary boundary.

The broader contribution is a three-property framework for interpreting longitudinal topic maps in national-language research systems: semantic interpretability, partition stability, and temporal robustness. Across alternative corpus definitions, metadata fields, embeddings, clustering settings, period boundaries, exclusion of 2025, and a same-corpus LDA baseline, the demographic trajectory remained upward even though prevalence magnitude, semantic compactness, and partition membership

varied. Partition stability can be examined through overall partition agreement (NMI) and record-level membership overlap (Jaccard), while explicit outlier handling clarifies the denominators used for temporal prevalence. The workflow therefore supports transparent sensitivity analysis under evolving multilingual metadata without treating a single fitted partition as uniquely correct. Comparative studies using harmonized retrieval and metadata rules are needed to determine whether similar research-field reorientations occur beyond the Korean case.

Declarations

Funding: No funding was received for conducting this study.

Competing interests: The authors have no competing interests to declare.

Author contributions: Su-Kyeong Park and Heonsoo Kim contributed to the conceptualization of the study. Su-Kyeong Park was primarily responsible for corpus construction, screening, and relevance assessment, with contributions from Heonsoo Kim and Hee-Su Kim. All three authors contributed to data curation, investigation, topic interpretation, and validation. Heonsoo Kim and Hee-Su Kim developed the topic-model methodology and conducted the formal computational analysis, software development, visualization, and Scopus cross-database diagnostic. Heonsoo Kim was responsible for project administration, and Hee-Su Kim provided supervision. Heonsoo Kim and Hee-Su Kim prepared the original draft of the manuscript. All authors contributed to reviewing and editing the manuscript and read and approved the final version.

Data availability: The study uses bibliographic records retrieved from the Korea Citation Index (KCI). KCI-derived titles, abstracts, keywords, and personal metadata are not redistributed. Stable KCI record identifiers and derived analytical outputs supporting the reported results, including the ID-only Scopus source-screen and article-confirmation evidence, are provided in the accompanying public reproducibility package (version 2.6.1). The Zenodo DOI and anonymous reviewer-access link will be inserted after author-controlled upload; this manuscript does not claim a completed deposit.

Code availability: Code used for corpus processing, topic modeling, sensitivity analyses, public Scopus-evidence aggregation, and the reported tables and figures is provided in the accompanying reproducibility materials. Repeating the licensed article-level Scopus searches requires authorized Scopus access; the public package contains the query routes and ID-only results but no credentials or licensed result export.

References

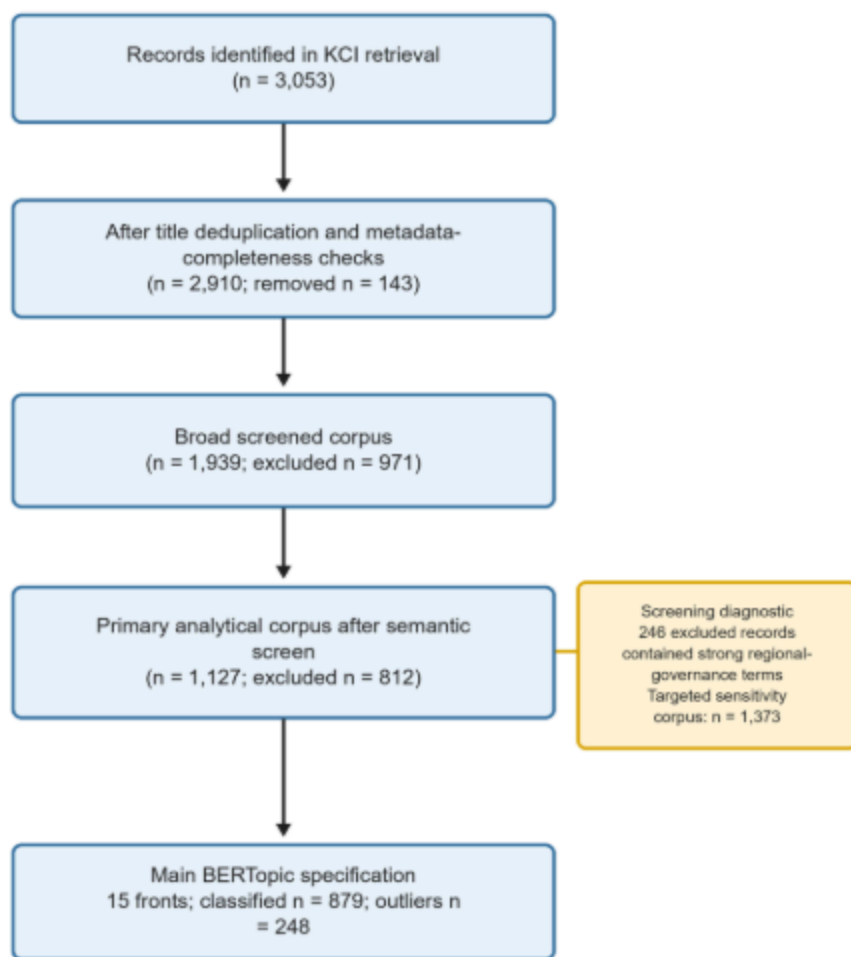
1. Asubiaro, T. V. (2026). Twenty-one years of coexistence: A comparative analysis of disciplinary and regional journal coverage in Web of Science and Scopus (2004–2024). *Scientometrics*, 131, 5277–5299. <https://doi.org/10.1007/s11192-026-05715-w>

2. Ballester, O., & Penner, O. (2022). Robustness, replicability and scalability in topic modelling. *Journal of Informetrics*, 16(1), Article 101224. <https://doi.org/10.1016/j.joi.2021.101224>
3. Bascur, J. P., Verberne, S., van Eck, N. J., & Waltman, L. (2025). Which topics are best represented by science maps? An analysis of clustering effectiveness for citation and text similarity networks. *Scientometrics*, 130, 1181–1199. <https://doi.org/10.1007/s11192-024-05218-6>
4. Benz, P., Pradier, C., Kozłowski, D., Shokida, N. S., & Larivière, V. (2025). Mapping the unseen in practice: Comparing latent Dirichlet allocation and BERTopic for navigating topic spaces. *Scientometrics*, 130, 3839–3870. <https://doi.org/10.1007/s11192-025-05339-6>
5. Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. In *Proceedings of the 23rd International Conference on Machine Learning* (pp. 113–120). <https://doi.org/10.1145/1143844.1143859>
6. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
7. Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research. *Scientometrics*, 22(1), 155–205. <https://doi.org/10.1007/BF02019280>
8. Choi, M.-J., & Paek, Y. (2023). Local extinction in a territorial trap: A critical review of population policy based on administrative region. *The Geographical Journal of Korea*, 57(2), 141–163. <https://doi.org/10.22905/kaopqj.2023.57.2.4>
9. Choi, Y. (2022). The direction of decentralization and balanced regional development policy of the Yoon Seok-Yeol administration. *The Korea Local Administration Review*, 36(4), 3–44. <https://doi.org/10.22783/krila.2022.36.4.3>
10. Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology*, 62(7), 1382–1402. <https://doi.org/10.1002/asi.21525>
11. Culbert, J. H., Hobert, A., Jahn, N., Haupka, N., Schmidt, M., & Donner, P. (2025). Reference coverage analysis of OpenAlex compared to Web of Science and Scopus. *Scientometrics*, 130, 2475–2492. <https://doi.org/10.1007/s11192-025-05293-3>
12. Egger, R., & Yu, J. (2022). A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. *Frontiers in Sociology*, 7, Article 886498. <https://doi.org/10.3389/fsoc.2022.886498>
13. Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(Suppl. 1), 5228–5235. <https://doi.org/10.1073/pnas.0307752101>
14. Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2203.05794>
15. Hladík, R., & Renisio, Y. (2025). Mapping knowledge: Topic analysis of science locates researchers in disciplinary landscape. *Poetics*, 108, Article 101950. <https://doi.org/10.1016/j.poetic.2024.101950>
16. Jang, I.-S. (2007). An analysis of economic effects of regional innovation system. *Seoul Studies*, 8(1), 19–39. <https://doi.org/10.23129/seouls.8.1.200703.19>

17. Jeong, H., & Jun, H.-J. (2025). An exploratory analysis of South Korean balanced regional development: Using LDA topic modeling. *Korean Journal of Public Administration*, 34(4), 148–179. <https://doi.org/10.22897/kipajn.2025.34.4.005>
18. Kim, H. H. (2017). Analysis of issues of Moon Jae-in administration's decentralized and balanced regional development policy focusing on implement system. *The Korea Local Administration Review*, 31(3), 51–80. <https://doi.org/10.22783/krila.2017.31.3.51>
19. Kim, R., & Kim, P. (2024). Analyzing perceptions of local autonomy through topic modeling: Focusing on the relationship between decentralization and balanced development. *Local Administration Review*, 38(3), 253–290. <https://doi.org/10.22783/krila.2024.38.3.253>
20. Kim, S., Choi, T., Yoon, A., & Seol, S. S. (2013). Korea Citation Index and its macro bibliometrics. *Asian Journal of Innovation and Policy*, 2(2), 194–211. <https://doi.org/10.7545/ajip.2013.2.2.194>
21. Kim, S.-B., & Jin, Y.-H. (2006). A study on decentralized regional development system: A political economic analysis. *The Korea Spatial Planning Review*, 49, 55–76.
22. Kumpulainen, M., & Seppänen, M. (2022). Combining Web of Science and Scopus datasets in citation-based literature study. *Scientometrics*, 127, 5613–5631. <https://doi.org/10.1007/s11192-022-04475-7>
23. Lim, H. B. (2022). Crisis of local extinction due to population decline and immigration. *Journal of the Korean Regional Development Association*, 34(4), 15–40. <https://doi.org/10.22885/KRDA.2022.34.4.15>
24. McInnes, L., Healy, J., & Astels, S. (2017). hdbscan: Hierarchical density based clustering. *Journal of Open Source Software*, 2(11), Article 205. <https://doi.org/10.21105/joss.00205>
25. McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform manifold approximation and projection for dimension reduction [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.1802.03426>
26. Mishra, M., Vishwakarma, S. K., Malviya, L., & Anjana, S. (2025). Temporal analysis of computational economics: A topic modeling approach. *International Journal of Data Science and Analytics*, 20, 2325–2339. <https://doi.org/10.1007/s41060-024-00596-9>
27. Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213–228. <https://doi.org/10.1007/s11192-015-1765-5>
28. Nazarovets, M., Laakso, M., & Taşkın, Z. (2026). University journals in the global academic publishing landscape: Mapping over 19,000 diverse titles based on Ulrichsweb data. *Scientometrics*, 131, 703–731. <https://doi.org/10.1007/s11192-025-05535-4>
29. Negishi, M., Sun, Y., & Shigi, K. (2004). Citation database for Japanese papers: A new bibliometric tool for Japanese academic society. *Scientometrics*, 60(3), 333–351. <https://doi.org/10.1023/B:SCIE.0000034378.38698.B2>
30. Nowakowska, M. (2025). A comprehensive approach to preprocessing data for bibliometric analysis. *Scientometrics*, 130, 5191–5225. <https://doi.org/10.1007/s11192-025-05415-x>
31. Park, E. L., & Cho, S. (2014). KoNLPy: Korean natural language processing in Python. In *Proceedings of the 26th Annual Conference on Human & Cognitive Language Technology* (pp. 133–136).

32. Park, S.-H., & Jang, I. (2024). A critique of the 2023 national plan responding to population decline for balanced regional development. *Journal of Policy Development*, 24(2), 1–39. <https://doi.org/10.35224/kapd.2024.24.2.001>
33. Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In *Proceedings of EMNLP-IJCNLP 2019* (pp. 3982–3992). <https://doi.org/10.18653/v1/D19-1410>
34. Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. In *Proceedings of WSDM '15* (pp. 399–408). <https://doi.org/10.1145/2684822.2685324>
35. Sanz-Casado, E., De Filippo, D., Aleixandre Benavent, R., Røeggen, V., & Pölönen, J. (2021). Impact and visibility of Norwegian, Finnish and Spanish journals in the fields of humanities. *Scientometrics*, 126, 9031–9049. <https://doi.org/10.1007/s11192-021-04169-6>
36. Seo, J. K. (2009). Balanced national development strategies: The construction of Innovation Cities in Korea. *Land Use Policy*, 26(3), 649–661. <https://doi.org/10.1016/j.landusepol.2008.08.014>
37. Su, X.-N., Han, X.-M., & Han, X.-N. (2001). Developing the Chinese Social Science Citation Index. *Online Information Review*, 25(6), 365–369. <https://doi.org/10.1108/EUM00000000006535>
38. van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84, 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
39. Vázquez, M. A., Pereira-Delgado, J., Cid-Sueiro, J., & Arenas-García, J. (2022). Validation of scientific topic models using graph analysis and corpus metadata. *Scientometrics*, 127, 5441–5458. <https://doi.org/10.1007/s11192-022-04318-5>
40. Wiechmann, T., & Pallagst, K. M. (2012). Urban shrinkage in Germany and the USA: A comparison of transformation patterns and local strategies. *International Journal of Urban and Regional Research*, 36(2), 261–280. <https://doi.org/10.1111/j.1468-2427.2011.01095.x>
41. Xie, Q., & Waltman, L. (2025). A comparison of citation-based clustering and topic modeling for science mapping. *Scientometrics*, 130, 2497–2522. <https://doi.org/10.1007/s11192-025-05324-z>
42. Yang, K., Lee, H., Kim, S., Lee, J., & Oh, D.-G. (2021). KCI vs. WoS: Comparative analysis of Korean and international journal publications in library and information science. *Journal of Information Science Theory and Practice*, 9(3), 76–106. <https://doi.org/10.1633/JISTaP.2021.9.3.6>

Figures



Retrieval/export date: 14 February 2026; publication years: 2004–2025.

Figure 1

Flow of records from KCI retrieval to the primary and sensitivity corpora. The 1,373-record sensitivity corpus adds the 246 records identified as potential false negatives.

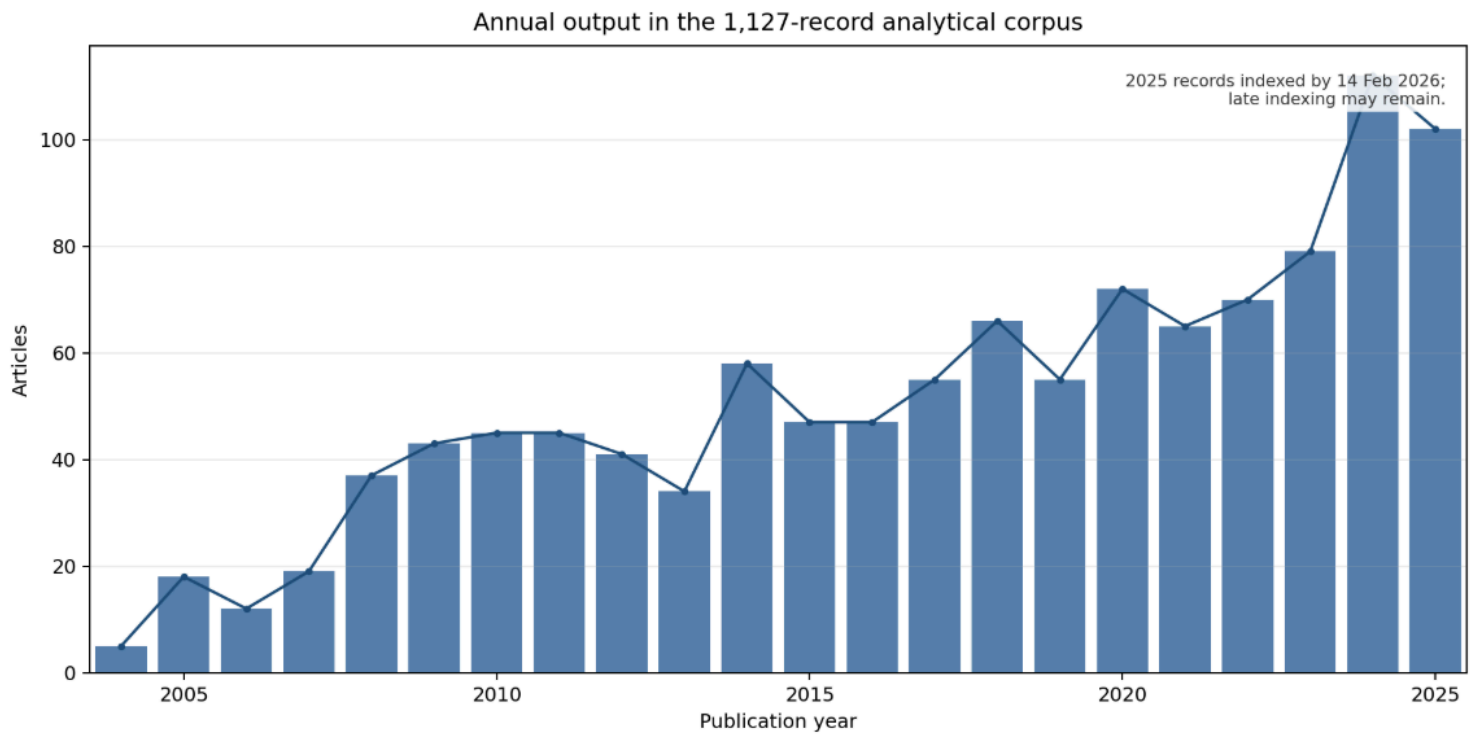


Figure 2

Annual publication output in the 1,127-record corpus, 2004–2025. The KCI export used for analysis was dated 14 February 2026; the 2025 value includes records indexed by that date and may be affected by late indexing.

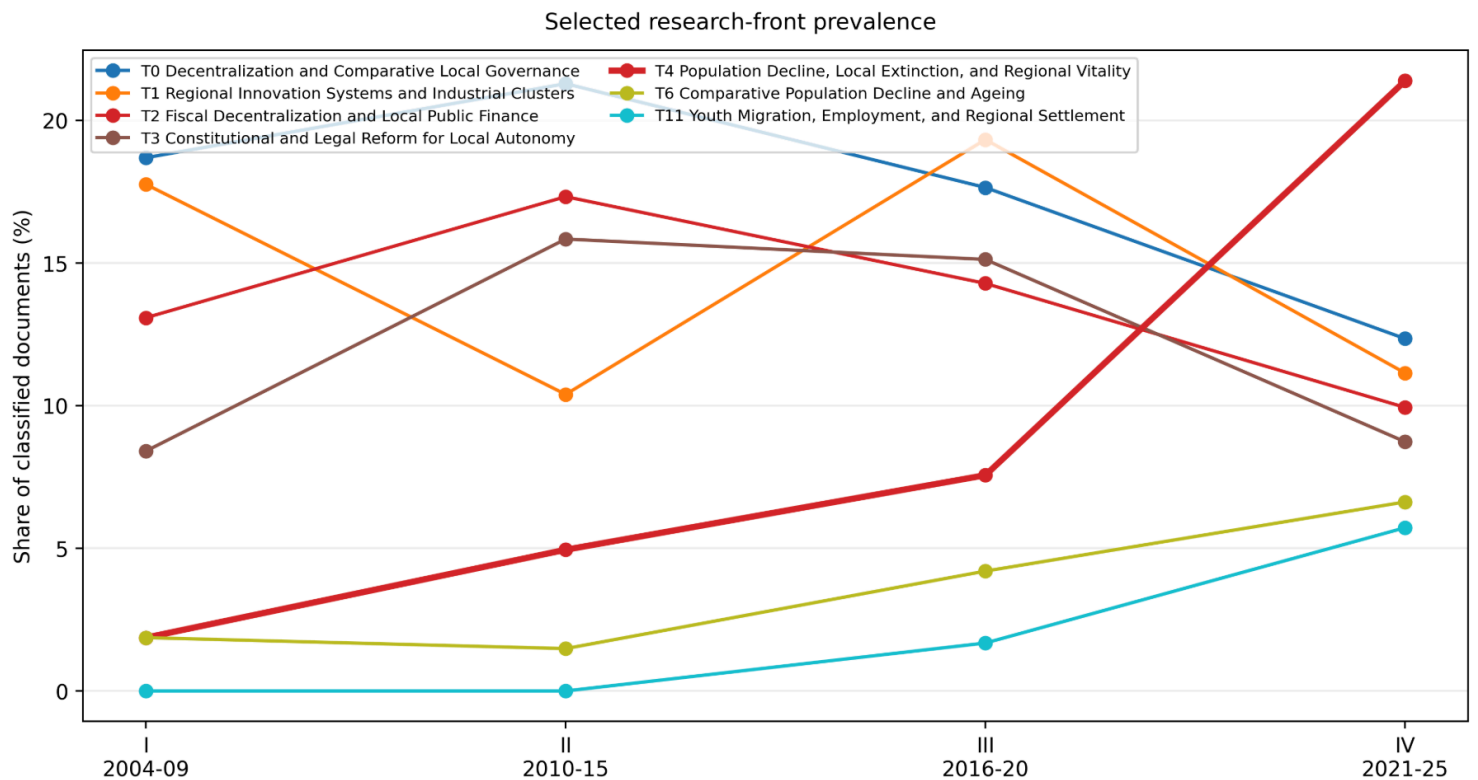


Figure 3

Period prevalence of selected research fronts. The horizontal axis shows four unequal-length windows (6, 6, 5, and 5 years); the vertical axis is each front's percentage of classified documents. The population-decline/local-extinction line rises from 1.9% to 5.0%, 7.6%, and 21.4%. Confidence intervals for T4 are reported in Table 4. Other lines provide scale and do not constitute significance tests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [OnlineResourceScientometrics.docx](#)