International Journal of Law, Ethics and Social Sciences

Review

Research on the Party's Self-Revolution Based on the BERTopic Model: Topic Identification, Evolution Analysis, and Trend Outlook

Qian Wei [1,*]

[1] Party School of CPC Chengdu Committee, Chengdu, 610110, China
* Correspondence: Qian Wei, Party School of CPC Chengdu Committee, Chengdu, 610110, China

Abstract: Research on the Party's self-revolution has developed from a political discourse into an increasingly structured field of academic inquiry. However, existing literature reviews rely largely on keyword co-occurrence and clustering methods, which have limitations in identifying semantic relationships, handling variant expressions, and capturing shifts in research themes. Their analyses of temporal evolution also tend to follow externally defined policy milestones. To address these limitations, this study applies the BERTopic model to analyze abstracts of articles published in Peking University Core and CSSCI-indexed journals. Following data screening, preprocessing, and text vectorization, the study identifies eight research topics and groups them into three major clusters: institutional frameworks supporting large-scale governance; the transition from conceptual principles to operational mechanisms; and organizational identity and sustained governance capacity. The temporal analysis reveals an initial stage of development during 2016–2018, broader expansion after 2021, peak research activity in 2022–2024, and a subsequent decline in 2025–2026. Research attention has increasingly focused on institutional norms, oversight mechanisms, and the challenges of large-scale organizational governance, while foundational value dimensions remain comparatively underrepresented. The analysis also identifies indistinct boundaries among some topics and a relatively isolated position for Topic 4. Overall, existing research demonstrates substantial normative discussion but remains limited in empirical mechanism testing and institutional effectiveness evaluation. Future research should strengthen mechanism-oriented analysis, develop evidence-based evaluation frameworks, improve conceptual clarity and thematic differentiation, and expand comparative studies to promote more systematic theoretical development.

Keywords: self-revolution; BERTopic; topic modeling; party building; political discourse

Received: 23 June 2026
Revised: 17 August 2026
Accepted: 30 August 2026
Published: 14 November 2026

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Recent policy-level discourse has emphasized the importance of institutional self-improvement mechanisms in large-scale governance systems, highlighting two complementary approaches: external accountability frameworks and internal renewal processes. This conceptual duality represents a distinctive contribution to contemporary governance theory and holds significant implications for understanding systemic resilience and adaptive capacity [1]. The internal renewal process-often termed 'self-revolution' in academic literature-has emerged as a central theoretical construct for addressing structural challenges inherent in complex, long-standing institutions and for advancing systematic governance modernization. Since the early 2010s, scholarly engagement with this concept has expanded rapidly, transitioning from normative political statements into a robust academic domain characterized by growing publication volume and increasing thematic diversification [2]. However, as a theoretical idea matures over more than a decade-from initial articulation to a densely populated field of knowledge production-the meta-questions of 'what has been studied, how it has evolved,

and where it is heading' themselves warrant rigorous scholarly attention [3]. Existing literature reviews predominantly employ keyword co-occurrence and clustering techniques to generate macro-level knowledge maps. While valuable for identifying broad structural patterns, such lexical-based methods face inherent limitations: they struggle to resolve synonymy, accommodate variant phrasings, and capture context-sensitive semantic shifts common in evolving governance discourses [4]. Moreover, temporal analyses are frequently aligned with externally defined policy milestones, potentially obscuring the endogenous development rhythm of academic inquiry. Consequently, the field's well-documented tendency toward strong normative argumentation and limited empirical validation is also reflected in how its own intellectual architecture is described and mapped [5]. In contrast, semantic embedding–clustering methods such as BERTopic offer a methodologically grounded alternative: they identify topic associations based on meaning rather than surface-form matching, enabling detection of semantically related yet lexically distinct constructs, and support dynamic modeling of topic emergence, decline, and conceptual drift through probabilistic temporal trajectories [6]. Accordingly, this study applies BERTopic modeling to conduct fine-grained semantic analysis of scholarly texts, enabling robust extraction of latent thematic structures, comprehensive mapping of research foci related to institutional self-renewal, identification of inter-topic relationships and longitudinal trends, precise characterization of core conceptual content and emerging thematic priorities, and evidence-informed guidance for future theoretical development and practical application.

2. Research Design

2.1. Research Roadmap

The overall research roadmap of this study is illustrated in Figure 1. This study applies the BERTopic model to perform topic-based text mining on scholarly texts concerning the Party's self-revolution in China, enabling systematic identification of prominent research themes and their temporal evolution.

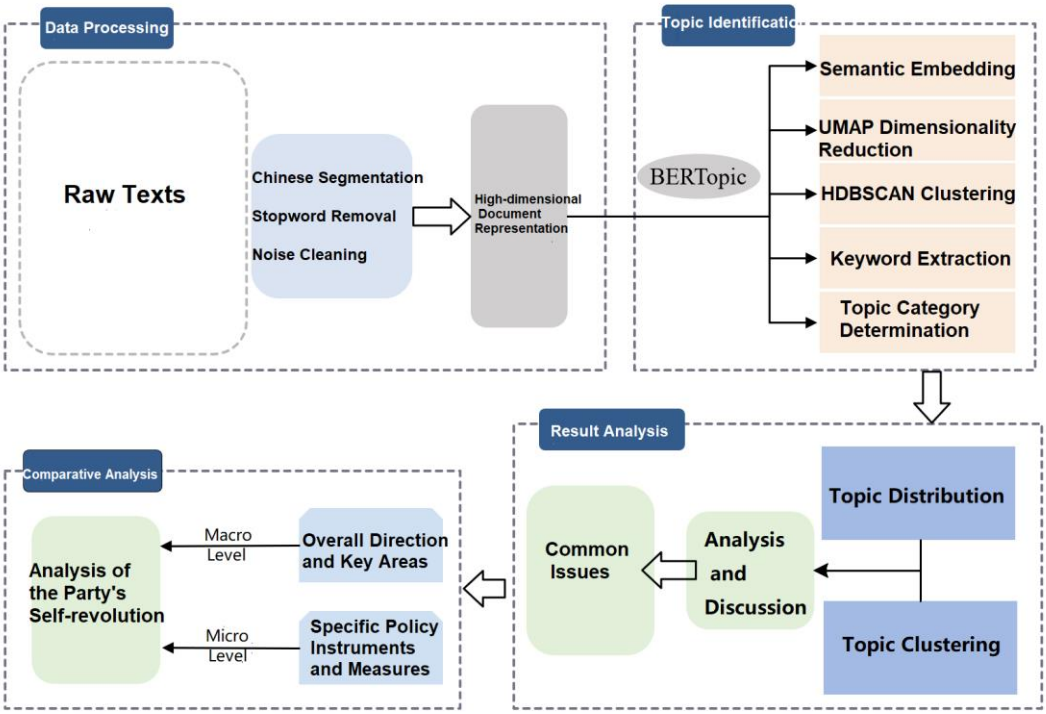


Figure 1. Research Framework

2.2. Research Methods

This study centers on the theme of self-revolution within the context of party building, leveraging the established capabilities of the BERTopic model for textual topic modeling and longitudinal analysis. A structured research workflow is adopted: data acquisition → preprocessing → topic modeling → evolutionary trend analysis → strategic insight derivation. First, abstracts from scholarly literature on party self-revolution are identified; selection criteria include inclusion in authoritative academic journal indexes and relevance to theoretical foundations, implementation approaches, normative significance, and developmental patterns. Abstract content is systematically examined to extract key conceptual and practical dimensions. Subsequently, topic modeling outcomes inform a focused examination of current characteristics and functional demands, enabling the formulation of refined pathways and actionable strategies for advancing self-revolution in contemporary practice. The aim is to contribute rigorous analytical insights that support both theoretical development and institutional application in party-building work.

2.3. Data Preprocessing

2.3.1. Text Screening and Deduplication

First, a preliminary screening was conducted on the raw literature abstract texts obtained from retrieval, removing those weakly related to the theme of self-revolution, lacking substantive theoretical discussion or practical analysis, as well as redundant or low-credibility texts. For texts sourced from different databases, deduplication was performed by comparing titles, publication dates, and core viewpoints to ensure sample uniqueness and validity. Finally, literature abstracts highly relevant to the research theme were retained for analysis [7].

2.3.2. Chinese Word Segmentation and Stopword Removal

Text data preprocessing was implemented using Python. The Excel-formatted text dataset was loaded via the pandas library, and non-string entries were filtered and cleaned. A custom clean_text function was applied to remove numbers, punctuation, and special characters, retaining only Chinese characters. A consolidated stopword list was developed by merging widely used academic and industry resources to eliminate function words and modal particles lacking substantive semantic content [8]. Synonym substitution was performed using a domain-specific synonym dictionary, while key policy-related terms were explicitly preserved through a dedicated terminology list. Word segmentation employed the precise mode of the jieba library, followed by removal of tokens with character length less than one. Keyword frequency counts were derived from the segmented documents, yielding 1,648 sentence-level units to support subsequent topic modeling.

2.3.3. Text Vectorization

Text vectorization employed the pre-trained paraphrase-multilingual-MiniLM-L12-v2 model to generate standardized document vectors. Leveraging deep learning architectures, this model computes context-aware vector representations for input tokens, mitigating semantic ambiguity arising from polysemy and thereby improving both the precision and depth of semantic analysis; the resulting document vectors thus more accurately capture core textual meaning.

2.4. Data Cleaning and Validation

A comprehensive quality check was conducted on the preprocessed data: abnormally short or semantically incomplete samples were removed, and consistency between document vector dimensions and text metadata was verified. Manual sampling assessed word segmentation accuracy, stopword removal effectiveness, and vector representation rationality, ensuring no logical errors or omissions of key information [9]. The resulting dataset fully satisfied the format and quality requirements for input to the BERTopic

model, thereby establishing a robust foundation for subsequent topic identification and evolution analysis.

3. Identification of Research Topics and Trend Analysis of "the Party's Self-Revolution"

3.1. Analysis of Research Topics

In line with the core objective of this study-topic mining on literature abstracts concerning self-revolution-and considering that high-dimensional document vectors tend to reduce the training efficiency of topic models and that some low-information dimensions interfere with clustering performance, dimensionality reduction was applied to the document vectors generated after preprocessing in order to improve model training accuracy and avoid the curse of dimensionality. The UMAP algorithm was adopted for vector dimensionality reduction. This algorithm can learn the global and local structural features of the data through the number of neighbors and the minimum distance parameters, significantly reduce dimensionality while preserving the core features of the original vectors to the greatest extent, and possesses good flexibility in embedding dimensions. The specific dimensionality reduction parameters were set as follows: n_neighbors = 30, n_components = 5, min_dist = 0.0, metric = 'cosine', and random_state = 42. On the basis of dimensionality reduction, the HDBSCAN density-based hierarchical clustering algorithm was used for topic clustering analysis. By constructing a mutual reachability distance graph for data points, this algorithm transforms the original vector space into a density-aware metric space. Its core logic is to first create a minimum spanning tree and convert it into a hierarchical clustering tree, and then automatically extract the optimal clusters based on stability theory. Its advantages lie in requiring no preset number of clusters, being able to identify clusters of arbitrary shapes, and automatically labeling data points in low-density regions as noise points.

In this study, the HDBSCAN clustering parameters were set as follows: minimum cluster size (min_cluster_size) = 30, density sensitivity control parameter (min_samples) = 10, distance metric (metric) = 'euclidean', and cluster selection method (cluster_selection_method) = 'eom'. After clustering was completed, marginal topics with low weight proportions were further removed, ultimately determining topic categories with clear semantic boundaries that reflect core research directions.

Through BERTopic modeling, a total of eight topics were extracted, and their topic feature words and weight distributions are shown in Figure 2. As can be seen from the figure, the research takes self-revolution as its core axis and forms a clearly layered analytical framework:

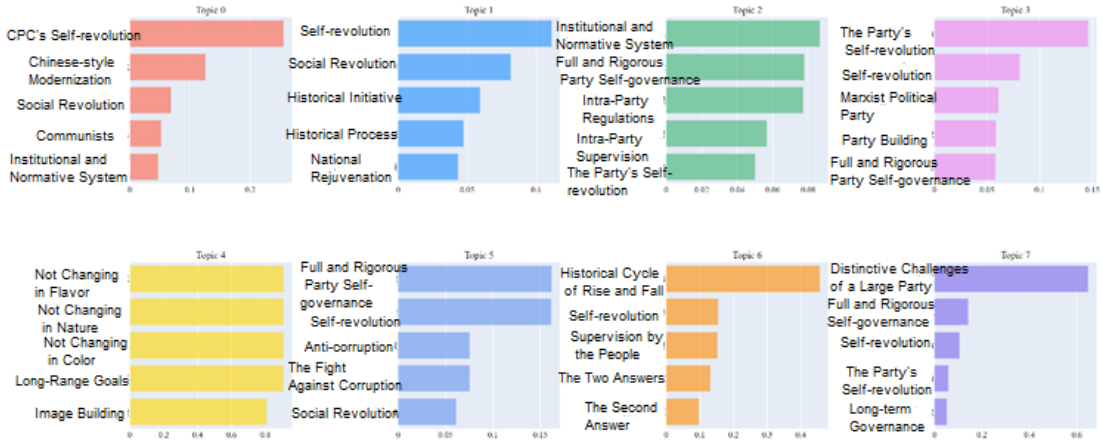


Figure 2. Topic Feature Words and Weights of the BERTopic Model

Topic 0 centers on "revolution," "modernization," and "institutional and normative frameworks," anchoring the fundamental position and institutional foundation of self-revolution in the process of leading national development.

Topic 1, starting from feature words such as "historical initiative" and "long-term development goals," reveals that self-revolution reflects proactive adaptation to historical patterns and serves as an internal driver for advancing broader societal transformation, thereby clarifying its foundational logic and contextual grounding.

Topic 2 focuses on practical pathways such as "comprehensive governance standards," "internal governance frameworks," and "internal oversight mechanisms," demonstrating the specific mechanisms through which self-revolution is advanced via structured and systematic approaches.

Topic 3 returns to theoretical foundations such as "political party theory" and "organizational development," clarifying the conceptual meaning of self-revolution as an intrinsic attribute of political organizations committed to continuous improvement.

Topic 4, with "consistency in purpose, principle, and practice" as its core, emphasizes that the fundamental aim of self-revolution is to uphold organizational identity and strategic continuity.

Topic 5 highlights integrity maintenance as a key domain of application for self-revolution, underscoring the central role of comprehensive governance standards in sustaining organizational health.

Topic 6 revolves around "historical patterns of governance," "public accountability," and "dual mechanisms for sustainability," and, proceeding from historical experience and stakeholder perspectives, explains that self-revolution constitutes a critical mechanism for ensuring long-term institutional resilience, reflecting its underlying value orientation.

Topic 7 ultimately addresses challenges associated with scale and duration of operation, pointing out that self-revolution is a foundational approach for responding to the complexities of sustained organizational leadership.

3.2. Spatial Distribution and Structural Characteristics of Topics

After nonlinear dimensionality reduction of the document embedding vectors using the UMAP algorithm, a two-dimensional vector visualization was generated, as shown in Figure 3. Research on "organizational self-renewal" exhibits a relatively concentrated distribution trend. Topic 0 (organizational self-renewal, development pathways aligned with national context, societal transformation) and Topic 4 (long-term objectives, consistency of foundational principles, continuity of core identity) occupy distinct spatial regions and are semantically distant from other topics, indicating strong thematic uniqueness: the former emphasizes linkages between internal renewal and strategic development frameworks, while the latter centers on enduring aims and preservation of essential characteristics. Topic 1 (self-renewal, societal transformation, historical agency), Topic 2 (institutional architecture for self-renewal, comprehensive governance standards, formalized governance frameworks), Topic 3 (self-renewal, ideological foundations of political organizations), Topic 5 (comprehensive governance standards, integrity assurance mechanisms), Topic 6 (historical patterns of institutional evolution, public oversight mechanisms), and Topic 7 (complexities associated with large-scale organizational structures, sustained institutional stewardship) form a cohesive cluster in the upper-middle portion of the figure. These topics share high semantic affinity and collectively address central themes including "self-renewal," "comprehensive governance standards," and "institutional architecture," reflecting interconnected dimensions such as system design, operational challenges, and historical continuity.

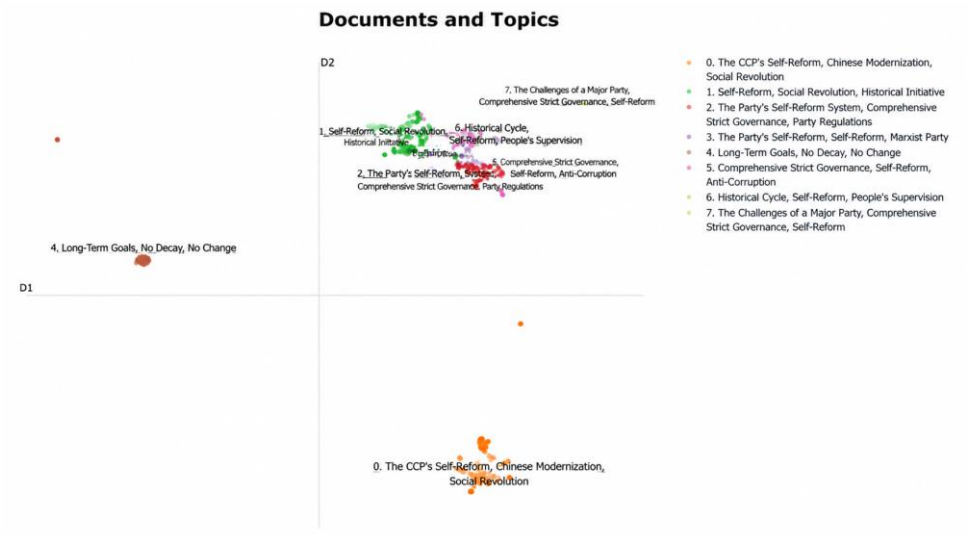


Figure 3. Two-Dimensional Visualization of Topic Distribution in the BERTopic Model

3.3. Topic Association Analysis

The strength of semantic association among topics reflects the degree of knowledge flow and cross-fertilization within a research field. Figures 4 and 5 visually present the association structure among the eight topics through a similarity heatmap and a topic distance map. The similarity heatmap (Figure 4) visually and clearly presents the proximity relationships and clustering potential of different topics in the semantic space. In the figure, darker colors indicate higher semantic similarity between the corresponding topics, meaning that the two are strongly related in semantic expression and keyword distribution, whereas lighter colors indicate lower correlation and greater semantic differences between topics. For example, Topic 2, representing "institutional frameworks, normative systems, and governance mechanisms," and Topic 3, representing "theoretical foundations and organizational identity of political parties," show a relatively deep blue, indicating strong conceptual alignment; Topic 2 and Topic 5, representing "governance mechanisms, institutional frameworks, and integrity assurance practices," also show an evident blue, reflecting close semantic correspondence between structural design and operational implementation. Topic 7, representing "challenges of sustained governance, institutional frameworks, and adaptive capacity," has high similarity with Topic 2, embodying the internal connection between long-term governance requirements and institutional development. Topic 4, representing "enduring objectives, continuity of purpose, and stability of orientation," appears lighter in color against most other topics, and is close to white against Topic 0, representing "transformative governance, modernization pathways, and systemic evolution," indicating a low degree of association.

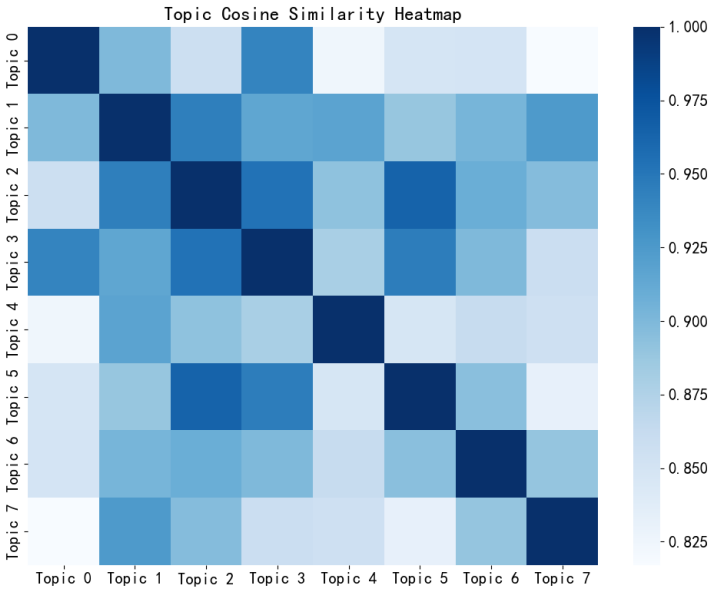


Figure 4. Topic Cosine Similarity Heatmap of the BERTopic Model

As can be seen from the topic distance map (Figure 5), all topics are divided into three relatively independent cluster groups: two dots in the upper left, three dots in the upper right, and three dots in the lower left. Evident spatial separation exists among the groups, reflecting that they belong to different research branches semantically. The dot group in the upper left (comprising Topic 0 and Topic 1) is tightly clustered, and the two are highly similar in focusing on the interplay between transformative governance and modernization pathways. The dot group in the upper right (comprising Topic 2, Topic 5, and Topic 6) forms a core cluster interweaving practice and theory around governance mechanisms and institutional frameworks. The dot group in the lower left (comprising Topic 3, Topic 4, and Topic 7) focuses on continuity of purpose and adaptive governance under extended operational contexts, presenting a relatively independent research perspective. The size of the dots in the figure also implies the document coverage of the topics: larger dots represent topics containing more literature and are the hotspots of the research field.

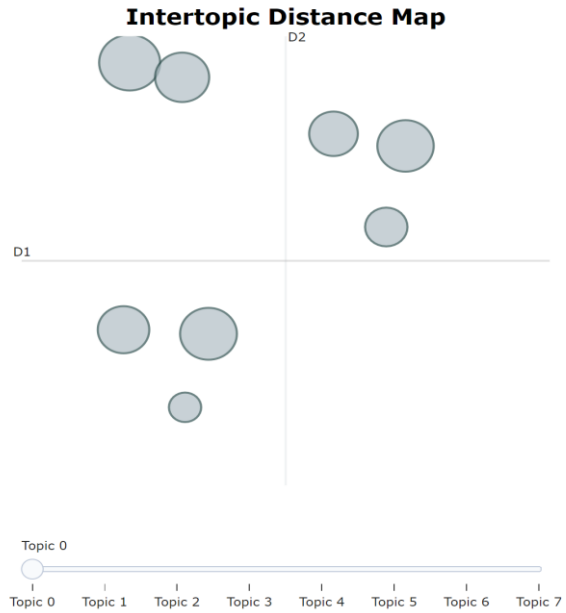


Figure 5. Visualization of Topic Distances in the BERTopic Model

3.4. Clustering and Grouping of Research Topics

Based on the automatic clustering results of BERTopic, this paper conducts manual interpretation and classification of the eight topics by integrating topic feature words, the topic cosine similarity heatmap, the topic distance map, and established thematic frameworks. Cluster naming follows a threefold rationale: "feature words–semantic boundaries–thematic frameworks"-first, identifying the central focus of each topic through its high-frequency feature words; second, assessing semantic boundaries based on proximity relationships among topics in the two-dimensional semantic space; and third, aligning topic groupings with recognized conceptual categories within the broader domain. It should be noted that cluster labels serve as interpretive summaries of the underlying semantic structure rather than definitive or mutually exclusive designations; moreover, because the analysis draws exclusively on literature abstracts, the clustering outcomes retain an exploratory character. As shown in Figure 6, hierarchical clustering combined with manual interpretation consolidates the eight topics into three overarching research directions: the first comprises Topic 0, Topic 1, and Topic 6; the second includes Topic 2 and Topic 5; and the third encompasses Topic 3, Topic 4, and Topic 7. This tripartite grouping aligns broadly with the spatial distribution pattern-"upper left–upper right–lower left"-observed in the topic distance map (Figure 5), suggesting that the semantic proximity relationships among topics exhibit robustness and reflect a relatively stable internal knowledge architecture. Notably, Topic 6 appears adjacent to both Topic 2 and Topic 5 in the topic distance map and displays integrative characteristics; however, given its thematic emphasis on historical patterns of governance, public oversight mechanisms, and foundational theoretical responses, it is assigned to the first cluster in accordance with strategic conceptual positioning and historical-logical coherence, while its associative links to the second cluster are explicitly acknowledged in the interpretive discussion.

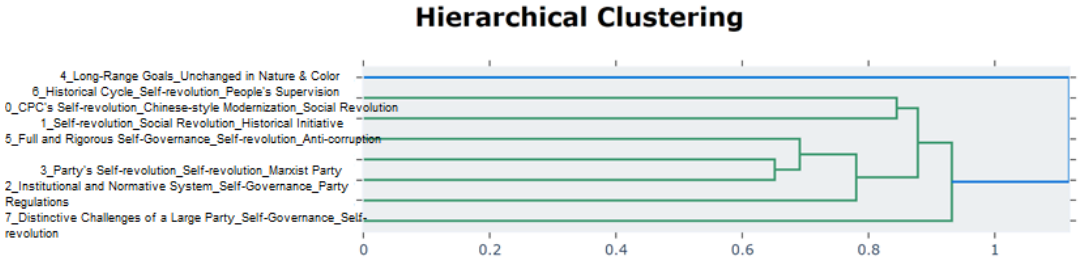


Figure 6. Hierarchical Clustering of Research Topics

1) Studying "the Party's Self-Revolution" from the Strategic Height of Escaping the Historical Cycle of Rise and Fall and Advancing Chinese-Style Modernization

The first cluster consists of Topic 0, Topic 1, and Topic 6. This cluster, with core feature words such as 'self-revolution', 'social transformation', 'modernization with distinctive features', 'historical agency', 'national renewal', 'historical patterns of governance continuity and change', 'public oversight', and 'dual institutional responses', situates self-revolution within the macro framework of modernization with distinctive features, national renewal, and historical patterns of governance continuity and change [10]. It emphasizes that self-revolution is not only an internal requirement of organizational development but also a strategic orientation for guiding broader societal transformation and exercising historical agency. This cluster is mainly divided into three levels:

First, research on the relationship between self-revolution and social transformation. Systematic analyses explain, from theoretical, historical, and practical perspectives, that guiding social transformation through self-revolution represents both a continuation of foundational revolutionary theory and an empirical distillation of long-term

developmental experience; it is further identified as essential for achieving national renewal. Emphasis is placed on strengthening ideological coherence, political alignment, institutional robustness, and organizational effectiveness to realize this guiding function [11]. Other studies examine this relationship through internal logic, operational structure, and implementation pathways, noting that self-revolution and social transformation are interdependent and mutually reinforcing. Effective coordination is achieved via six functional mechanisms-political orientation, conceptual guidance, organizational coordination, conduct alignment, normative adherence, and institutional support. It is underscored that comprehensive organizational strengthening, anchored in unified leadership, is necessary to sustain the guiding role of self-revolution in societal transformation. At the methodological level, coordinated advancement in the current period is proposed to follow five integrative principles: balancing comprehensive analysis with priority setting, integrating strategic planning with iterative implementation, harmonizing structural stability with adaptive responsiveness, combining principled commitment with sustained implementation, and aligning internal capacity building with external engagement-thereby enhancing theoretical clarity and practical coherence in mutual reinforcement between self-revolution and social transformation.

Second, research on the relationship between self-revolution and modernization with distinctive features. Analyses from internal logic, operational structure, and implementation pathways clarify the guiding function of self-revolution in societal transformation, indicating that under evolving conditions, self-revolution serves as a pivotal internal driver for advancing broader transformation and furnishes decisive leadership impetus. From a modernization-capacity perspective, self-revolution contributes to strengthening organizational capability at the agent, conceptual, and governance levels, thereby delivering integration of purpose, articulation of guiding values, and enhancement of administrative efficacy for modernization with distinctive features [12]. Further scholarship situates the role of self-revolution in advancing modernization with distinctive features within the broader scholarly framework of 'self-revolution guiding societal transformation', identifying its contribution through institutional assurance, momentum generation, and strategic consistency.

Third, research on the relationship between self-revolution and public oversight in addressing historical patterns of governance continuity and change. It is observed that the evolution from public oversight to self-revolution reflects a commitment to foundational principles while adapting to new circumstances; the two approaches are logically complementary-functioning respectively as external accountability and internal renewal-and jointly reinforce one another in effect and timing. Together, they illuminate enduring organizational vitality and deepen understanding of governance dynamics, institutional development, and societal progress. Discussions on the evolution of dual institutional responses-from initial emphasis on public oversight to the integrated application of internal renewal and external accountability-highlight a progression toward balanced governance frameworks, encompassing both external input and internal refinement, and reflecting increasing conceptual maturity. Demonstrations of coherence between the two approaches emphasize alignment in underlying objectives, compatibility in functional design, synergy in dynamic interaction, and mutual reinforcement in developmental trajectory. It is proposed that, in contemporary practice, organizational resilience should be reinforced through sustained vigilance, adherence to core commitments, refinement of internal governance systems, and deeper integration of public oversight and self-revolution to ensure continued alignment with historical patterns of enduring governance [13].

2) Studying "the Party's Self-Revolution" from the Transformation of Theoretical Propositions into Institutional Practice

The second cluster consists of Topic 2 and Topic 5. This cluster revolves around themes such as the institutional and normative system of organizational self-renewal, comprehensive and rigorous organizational self-governance, internal rules, internal

oversight, and integrity enforcement, focusing on the institutionalization, standardization, and effectiveness of self-renewal. It centrally embodies the research orientation of transforming self-renewal from a theoretical proposition into institutional practice and supervisory practice. Academic research in this area mainly concentrates on two aspects:

First, research on the institutional and normative system of organizational self-renewal. Previous studies divide the institutional and normative system of organizational self-renewal into three subsystems: fundamental institutional norms, basic institutional norms, and important institutional norms, emphasizing that institutions should be made to work in the same direction through core leadership tenets, the support of the supervisory system, and the implementation of responsibility for managing and governing the organization. Other scholars point out that this system is composed of multiple elements, including fundamental guidelines, the internal rule system, the supervisory system, the dynamic optimization mechanism, and the guarantee of discipline enforcement and accountability, and that it possesses functions such as normative guidance, order construction, and vitality stimulation. Furthermore, adopting a systemic perspective, it is emphasized that handling the relationships between explicit rules and hidden rules, core institutions and supporting institutions, incentive institutions and restrictive institutions, as well as internal rules and national laws, is essential to enhance the integrity and synergy of the institutional system.

Second, taking integrity enforcement as the key practical field of organizational self-renewal, this line of research examines how the campaign against misconduct is regarded as the most thoroughgoing self-renewal. It has been pointed out that eliminating unethical and degenerative factors with a thoroughgoing spirit of self-renewal is the distinctive feature and logical mainline of organizational leadership over integrity enforcement in the contemporary era [14]. Following this mainline to reveal the practical logic of integrity enforcement and to summarize basic experience can provide a reference for pushing these efforts to greater depth. Furthermore, it is proposed that the building of integrity institutions and mechanisms should be guided by the strategic thought of self-renewal: uphold the organizational leadership position; improve the responsibility system under which the organization leads the campaign against misconduct [15]; strictly punish malfeasance in accordance with laws and rules and strengthen the punitive mechanism of deterrence against misconduct; improve integrity laws and rules and strengthen the preventive mechanism of structural prevention of misconduct; strengthen ideological education and consolidate the self-disciplinary mechanism and the ethical defense line of resistance to misconduct; and adhere to a systemic perspective to improve the effective carrier for the integrated and interconnected three-tiered integrity framework. Overall, this cluster has shifted from listing institutional elements to the closed loop of institutional articulation, implementation transformation, and supervisory accountability, taking integrity enforcement and the building of work style and discipline as key practical fields.

3) Studying "the Party's Self-Revolution" from the Perspective of the Party's Essential Attributes and Long-Term Governance Capacity Building

The third cluster consists of Topic 3, Topic 4, and Topic 7. Centered on core themes such as the nature of Marxist political parties, internal capacity building, exercising full and rigorous self-governance, remaining true to the original nature and character of the organization, the unique challenges facing a large political entity, and long-term governance, this cluster focuses on the subjectivity-based foundations that make self-revolution possible. It systematically explores the essential attributes of Marxist political organizations, their advanced nature and purity, the distinct challenges of managing a large party, and the fundamental prerequisites for sustained governance.

First, it examines why self-revolution constitutes an essential attribute and inherent requirement of Marxist political parties. Previous research has thoroughly examined these essential attributes and inherent requirements to demonstrate the intrinsic basis of the organization's self-revolution, explaining its possibility in terms of its fundamental nature and purpose, its strategic goals and aspirations, and its extensive historical experience.

Further studies point out that the discourse on self-revolution inherits and develops Marxist theories of organizational building, carries forward and innovates upon the discourse regarding the ruling party's self-revolution over the past century, and represents a theoretical elevation of the vivid practice of exercising full and rigorous self-governance in the new era. From the perspective of original contributions, it is argued that the central leadership's important thinking on self-revolution has significantly expanded the connotations and scope of the concept, provided a secondary answer to the question of how to break out of the historical cycle of rise and fall under conditions of long-term governance, and successfully broken free from the limitations of Western political discourse [8, 13]. Additional scholarship summarizes the original contributions of this important theoretical framework through three main propositions: the theory of self-revolution as a distinctive hallmark, the theory of the secondary answer, and the theory of the dialectical relationship between the two great revolutions. It is further pointed out that this thinking has profoundly deepened the understanding of the essential attributes of Marxist political parties and the laws governing ruling party administration.

Second, it examines why self-revolution is the fundamental approach for the ruling party to withstand the tests of long-term governance and address the unique challenges facing a large political organization. Drawing lessons from the rise and fall of political parties around the world over the past century, previous research has identified the core challenge facing a large party: how to prevent a Marxist political organization from becoming corrupt and deteriorating under conditions of sustained administration, and how to effectively break out of the historical cycle of rise and fall. From the perspective of long-term governance capacity building, addressing the unique challenges facing a large party is emphasized as a problem-oriented approach. This requires deepening self-revolution in the ideological sphere, improving an organizational system that ensures effective implementation from top to bottom, and preventing the infiltration of vested interests into internal organizational life [16]. Further analysis explains six core persistent questions from the perspectives of value objectives, organizational size, duration of governance, governance advantages, governance performance, and governance risks. A comprehensive set of solutions is proposed, guided by important theoretical thinking and the authority of the central leadership, covering long-term governance capacity, the system for exercising full and rigorous self-governance, self-supervision alongside public oversight, and institutional rules and normative frameworks. Additional studies further develop six long-term mechanisms: sustained driving force, unity and solidarity, enhanced capabilities, renewed morale, early warning and error correction, and self-strengthening and purification. The need to properly address the relationships between focused efforts to tackle key problems and overall coordination, as well as between self-supervision and public oversight, is also strongly emphasized [17].

3.5. Research Trends

As can be seen from Figure 7, in terms of overall trends, the topics were in their initial stage between 2016 and 2018, with literature frequencies generally low. Only the purple curve representing "long-range goals, not changing in nature, and not changing in color" (Topic 4) showed a certain degree of prominence in the early period, reflecting sustained academic attention to foundational questions concerning political consistency [5]. After 2021, the literature frequencies of all topics began to increase significantly, indicating broadening scholarly engagement. The blue curve representing "the distinctive challenges of a large organization, comprehensive internal governance, and transformative renewal" (Topic 0) and the orange curve representing "comprehensive internal governance, transformative renewal, and integrity maintenance" (Topic 5) grew most rapidly, reaching frequency peaks in 2023 and 2022 respectively, underscoring the centrality of organizational scale-related considerations and integrity-related research in contemporary institutional studies [18]. The green curve representing "the structural and procedural framework for transformative renewal, comprehensive internal governance, and formalized standards" (Topic 2) and the red curve representing "transformative

renewal, broader societal development, and historical agency" (Topic 1) also reached their peaks between 2023 and 2024, indicating growing emphasis on systemic design and longitudinal analysis. By 2025–2026, the literature frequencies of the topics show a marked decline; following a phase-specific peak, related scholarship appears to be transitioning toward consolidation and more in-depth theoretical exploration. In summary, the evolutionary trajectory reflects shifting emphases within the field of organizational transformative renewal, while also demonstrating responsiveness to evolving conceptual frameworks in institutional research; from early focus on foundational consistency, to mid-term emphasis on structural development and operational complexity, and finally to later-stage consolidation, the dynamic patterns collectively illustrate the maturation of this research domain [19].

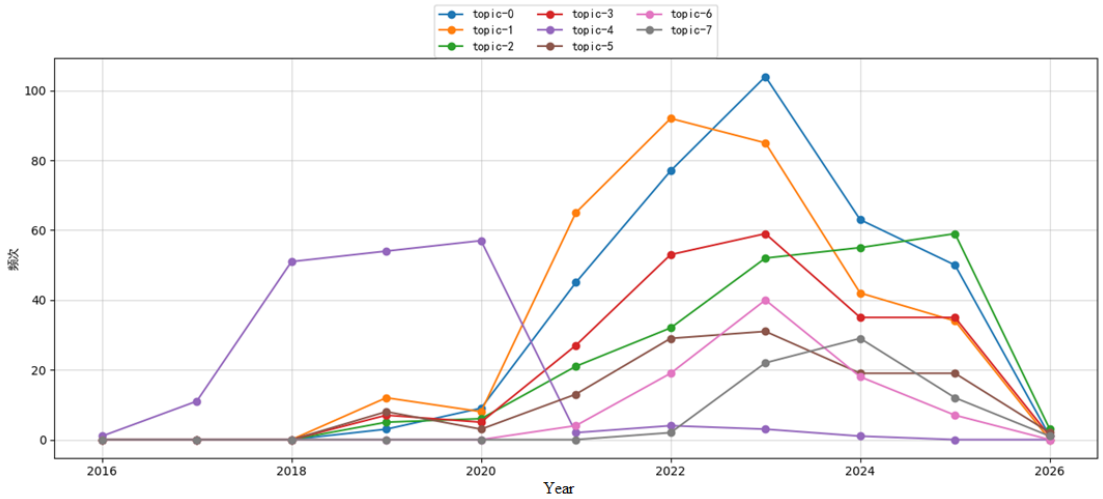


Figure 7. Topic Evolution of the BERTopic Model

4. Conclusion and Outlook

Using the BERTopic model, this paper conducted semantic modeling on abstracts of literature published in Peking University Core journals and CSSCI-indexed journals, identifying eight topics. Based on topic feature words, semantic proximity relationships, similarity heatmaps, topic distance maps, and thematic patterns in the literature, these topics were grouped into three clusters: first, examining self-revolution in relation to institutional frameworks supporting large-scale governance; second, analyzing the transition of self-revolution from conceptual formulation to operational mechanisms; and third, investigating self-revolution through the lens of organizational identity and sustained governance capacity. Topic evolution exhibits stage-specific patterns: an initial emergence during 2016–2018, broad expansion after 2021, peak activity in 2022–2024, and a subsequent decline in 2025–2026. The semantic structure indicates a pronounced shift toward applied concerns-including formalized norms, accountability systems, and challenges inherent to large-scale organizations-while dimensions related to foundational principles or normative aims remain comparatively underrepresented in the academic discourse. Boundaries among Topics 2, 5, and 7 are indistinct, whereas Topic 4 appears relatively isolated, suggesting a potential semantic divergence between widely used policy terminology and scholarly analytical constructs. These findings align with prior observations regarding an 'institutional turn' and a 'governance turn' in the field, and resonate with earlier assessments of a transition from theoretical derivation toward systematic institutional analysis. At the same time, they highlight persistent gaps in current scholarship, particularly concerning empirical mechanism testing, thematic differentiation, integration of normative and functional perspectives, translation between policy concepts and academic inquiry, and methodological coherence.

4.1. Weaknesses in Current Research

First, normative argumentation is relatively strong while empirical testing is weak, and mid-level mechanisms and evaluations of institutional effectiveness are insufficiently addressed. Existing studies mostly elaborate on the theoretical connotations, historical logic, value implications, institutional composition, and practical requirements of self-revolution, producing relatively rich normative results; however, empirical testing of questions such as how the institutions operate, how mechanisms are articulated, and how effectiveness is generated remains insufficient. In the topic identification results of this paper, issues such as institutional norms, comprehensive governance standards, integrity assurance, and the distinctive challenges of a large-scale organization are highly concentrated, indicating that the academic community has discussed "what to do" extensively, but research on mid-level mechanisms-"how it works," "under what conditions it is effective," and "how different mechanisms are combined"-remains weak. Systematic process tracing and comparative analysis are still lacking with respect to the process by which institutional advantages are transformed into governance effectiveness, the articulation mechanism between internal oversight and external accountability, the balancing mechanism between accountability measures and supportive frameworks, and the operating mechanism of self-correction and early warning against deviation; the indicator design and causal identification for evaluating institutional effectiveness also need to be strengthened.

Second, topic boundaries are blurred, concepts tend toward generalization, and knowledge branches have not yet been fully stabilized. The BERTopic results show that Topics 2, 5, and 7 are evidently semantically adjacent, indicating that "the institutional and normative system of self-revolution-comprehensive governance standards-integrity assurance-the challenges of long-term stewardship" have not yet formed clear problem boundaries in existing literature; Topic 4 is relatively isolated, reflecting that although the dimension of "core identity and foundational objectives" is of fundamental significance, it has not been sufficiently integrated with issues such as institutional practice and governance effectiveness. Blurred topic boundaries indicate both the strong interdisciplinary nature of this field and the risk of concept misuse, overlapping issues, and unclear problem levels in related research. Concepts such as "the distinctive challenges of a large-scale organization," "long-term stewardship," "self-revolution," and "broader societal transformation" are highly overarching, but in specific studies they are prone to overly broad extension, unclear analytical levels, and insufficient operationalization.

Third, the value dimension is relatively marginal, and the integration of value logic and institutional practice is insufficient. The semantic structure of this paper shows that although value–historical propositions such as "continuity of purpose and orientation," "accountability to stakeholders," and "dual pathways for sustainability" have dedicated topics, they are relatively marginal in the overall semantic space and have not formed high-intensity associations with practical issues such as institutional norms, internal oversight, and integrity assurance. Existing studies have elaborated considerably on foundational political identity, stakeholder-centered orientation, historical patterns of institutional endurance, and the "second pathway for sustainability," but pay insufficient attention to how value goals are embedded in institutional design, how they are transformed into oversight standards, and how they enter effectiveness evaluation. Two biases may thus arise: one is emphasizing institutional tools while neglecting value integration; the other is emphasizing value declarations while neglecting institutional realization. It has been noted that existing research suffers from a "lag in the construction of discourse systems" and that different discourse dimensions "lack effective connection and integration"; from the perspective of semantic association strength, this paper finds that the "semantic distance" between the value dimension and the institutional dimension is precisely the structural manifestation of this discourse fragmentation. How to achieve the integration of value logic, institutional logic, and practical logic remains an important issue that needs to be deepened in this field.

Fourth, policy-oriented terminology is visibly driving research, academic propositionalization and sustained accumulation are insufficient, and the policy–academic semantic gap has not yet been bridged. The evolution of topics is highly responsive to strategic guidance, indicating that this field has a distinct practical orientation and contextual sensitivity; however, the high frequency of policy-oriented terminology does not necessarily equal a synchronous deepening of theoretical substance. Some topics rapidly decline after forming short-term peaks, suggesting that related research may suffer from following topical trends, conceptual aggregation, repetitive elaboration, and insufficient sustained accumulation. The rhythm of "a peak in 2022–2024 and a decline in 2025–2026" presented by the topic evolution curves in this paper provides semantic-level evidence for this judgment from the temporal structure of knowledge production. Meanwhile, research on the relationship between self-revolution and broader societal transformation and modernization pathways remains rather macro-level: there are many macro-level association judgments but insufficient meso-level path analysis and micro-level evidence supply, which easily turns "guiding principles" into conclusive statements rather than testable propositions. How policy discourse can be transformed into testable, comparable, and operational academic propositions, and how policy issues can be sedimented into a stable theoretical agenda, still require further advancement in conceptual definition, variable extraction, mechanism identification, and evidence standards.

Fifth, the comparative perspective and the building of an academic community are insufficient. Existing studies are mostly conducted within localized institutional discourse and the theoretical framework of organizational development, and comparisons with the governance experience of other large-scale organizations and dialogue with international theories of political organization and institutional design remain limited. In addition, research forces in this field are relatively dispersed: the density of author collaboration networks is low, a core group of authors has not yet formed, inter-institutional cooperation is loose, and phenomena such as blurred topic boundaries and overlapping issues are closely related to the lack of stable cooperation and consensus-building mechanisms in the academic community. The fragmentation of research forces not only affects the efficiency of knowledge accumulation but also constrains the formation of signature concepts, mature research paradigms, and widely cited core achievements.

4.2. Future Research Directions

First, research should shift from descriptive topic analysis toward mechanism-oriented inquiry, adopting a design that bridges macro-level propositions, mid-level processes, and micro-level evidence. Future work should examine how institutional and normative frameworks support the translation of structural features into governance outcomes, focusing on processes such as internal regulatory development, oversight implementation, accountability enforcement, integrity maintenance, ideological alignment, and organizational development. Methodological approaches-including process tracing, multi-case comparative analysis, field-based observation, in-depth qualitative interviews, and longitudinal data analysis-can help identify the contextual conditions, causal pathways, configurational interactions, and boundary conditions associated with these processes. This supports a transition from enumerative reporting to explanatory mechanism analysis. Concurrently, challenges arising from scale and longevity of institutional operation should be reframed as empirically tractable questions. Mechanism-focused investigation can proceed along dimensions including organizational size, duration of institutional continuity, complexity of stakeholder engagement, adaptive oversight capacity, normative reinforcement, and functional capability development. An analytical framework structured around challenge specification, process response, outcome assessment, and iterative refinement can help avoid conceptual vagueness and promote analytical precision.

Second, an empirically grounded evaluation framework for institutional performance should be developed, emphasizing measurable and comparable indicators.

Such a framework should address dimensions including structural coherence, implementation fidelity, integrative oversight functionality, adaptive feedback responsiveness, and outcome generation. Distinct analytical dimensions-such as design quality, implementation intensity, systemic coordination, and perceived functional adequacy-should be differentiated. Evaluation strategies should integrate both subjective and objective measures, and combine qualitative and quantitative methodologies, thereby advancing scholarship from normative assertions about structural strengths toward evidence-based assessments of operational effectiveness.

Third, conceptual clarity and thematic delineation require deliberate attention. To address overlapping problem domains and fragmented analytical foci, future research should explicitly differentiate levels of analysis and substantive domains, cultivating an interconnected yet well-bounded knowledge structure linking formal frameworks, oversight execution, governance coordination, long-term institutional resilience, and foundational objectives. Systematic literature synthesis, historical conceptual analysis, thematic mapping, and structured expert consultation can support standardized definitions of core constructs, consistent classification of issues, and harmonized theoretical scaffolding-reducing conceptual ambiguity and redundant inquiry. Furthermore, foundational normative commitments should be centrally integrated into empirical analysis. Research should investigate how enduring principles are instantiated across institutional architecture, oversight practice, and performance evaluation; how citizen-centered orientations inform standards of accountability and responsiveness; and how complementary governance logics are coherently embedded within unified administrative and policy processes.

Fourth, research should further explore the linkages between internal institutional evolution and broader societal development trajectories, particularly in relation to modernization goals. Within the conceptual lineage connecting internal institutional adaptation and external developmental outcomes, specific transmission channels-including political stability assurance, strategic continuity, organizational coordination, policy coherence, and cross-sectoral integration-warrant empirical validation. Targeted studies in domains such as sustainable economic advancement, institutional capacity building, public service delivery, and systemic risk management can generate a layered evidentiary configuration, integrating high-level strategic analysis with intermediate pathway modeling and granular empirical verification.

4.3. Research Limitations and Room for Improvement

First, the sample and textual scope are limited. Using literature abstracts as analytical units effectively captures the broad topical structure of research domains but offers limited insight into argumentation strategies, theoretical foundations, and mechanistic detail present in full texts. Additionally, constraints on temporal coverage and database selection restrict the ability to trace long-term developmental trends or cross-linguistic scholarly contributions. Future work may broaden scope by incorporating full-text corpora, multiple bibliographic databases, and extended chronological spans to strengthen the representativeness and generalizability of findings.

Second, topic modeling and manual interpretation remain inherently exploratory. The selection of BERTopic parameters-including number of topics and clustering thresholds-as well as subsequent interpretive labeling, relies on iterative judgment rather than deterministic criteria. Although topic naming and classification follow a tripartite framework grounded in feature terms, semantic coherence, and disciplinary relevance, outcomes are not uniquely determined; sensitivity to parameter choices, stability across diverse corpora, and alignment between algorithmic groupings and expert judgment require further empirical validation. Enhancing methodological rigor could involve systematic parameter sensitivity analysis, comparative evaluation across multiple modeling approaches, and structured expert review protocols.

Third, semantic evolution analysis does not substitute for causal mechanism testing. The observed trajectories of topic emergence, persistence, and decline reflect surface-level

lexical shifts rather than empirically verified drivers of scholarly attention. While disparities in conceptual framing have been noted across policy documents, academic publications, and practice-oriented reports, systematic comparative analysis across these text types remains underdeveloped. Future studies may integrate dynamic topic modeling with cross-genre textual comparison and in-depth qualitative investigation of salient thematic clusters to strengthen explanatory depth and theoretical grounding.

In summary, scholarship on self-revolution has expanded rapidly from an initial discursive formulation into a densely populated domain of knowledge production. Its significance lies both in addressing foundational questions about institutional resilience and adaptability in large-scale organizations, and in contributing to broader inquiries into governance modernization. Current work demonstrates substantial progress in mapping topical configurations, documenting institutional practices, and analyzing policy responses. Nevertheless, notable gaps persist in causal mechanism testing, effectiveness assessment, value-laden conceptual integration, operational definition of core constructs, cross-contextual comparison, methodological standardization, and community-building around shared epistemic standards. Advancing the field requires moving beyond descriptive mapping to interrogate underlying mechanisms; shifting from semantic trend detection to substantive theoretical validation; embedding normative considerations within institutional analysis; deepening scholarly accumulation independently of immediate policy cycles; and fostering cumulative, consensus-oriented knowledge development. Such efforts would support a transition from predominantly normative discourse toward a more balanced scholarly paradigm-one that integrates robust description with rigorous explanation, bridges evaluative and structural dimensions, and situates context-specific inquiry within wider theoretical frameworks.

Funding: This paper is a phased achievement of the National Social Science Fund Youth Project "Research on Countermeasures against Privilege-Seeking Thoughts and Behaviors from the Perspective of Self-Revolution" (Project No.: 24CDJ078).

References

1. S. Y. Tang, “Rethinking local and regional governance in China: An institutional design and development perspective,” *Urban Governance*, vol. 1, no. 1, pp. 51–58, 2021.
2. G. Li, "Self-revolution: A Successful Path to Solve the Unique Problems of the Great Party," *Fu Dan Xue Bao. She Hui Ke Xue Ban*, no. 3, pp. 1–8, 2023.
3. Z. E. Frankel and H. M. Levitt, "Postmodern strategies for working with resistance: Problem resolution or self-revolution?," *Journal of Constructivist Psychology*, vol. 19, no. 3, pp. 219–250, 2006.
4. N. Yang and X. Zhang, "Interpretation of the Self-revolution System and Standard System in the New Era: Based on the Social System Research Methods," *Frontiers in Humanities and Social Sciences*, vol. 3, no. 8, pp. 148–152, 2023.
5. Y. Liu and Y. Lv, "Hot topics, evolutionary trajectories, and trend outlook of self-revolution research: A bibliometric analysis based on CNKI from 1993 to 2023," *Observation and Ponderation*, no. 2, pp. 84–96, 2025.
6. L. I. U. Chunming and B. A. O. Yuanxin, " The Centennial Exploration, Basic Experience and Realistic Enlightenment of the CPCs Adherence to Self Revolution," *Journal of Chongqing Jiaotong University Social Sciences Edition*, vol. 23, no. 6, p. 1, 2023.
7. J. Trondal, “‘Let’s organize’: The organizational basis for stable public governance,” *Public Administration*, vol. 101, no. 1, pp. 201–220, 2023.
8. Y. Mao, “What accounts for the different regional responses to COVID-19 in China? Exploring the role of institutional environment, governance capacity and legitimacy,” *Health Policy and Planning*, vol. 38, no. 4, pp. 552–566, 2023.
9. Q. Yu, "Three-Dimensional Exploration on the Construction of the Discourse System of Self-Revolution of the CPC in the New Era," *Journal of Southwest University Social Science Edition*, vol. 50, no. 6, pp. 42–52, 2024.
10. S. Martin-Iverson, "'Revolusi diri' (self-revolution): personal choice, collective identity and subcultural change in the Bandung straight edge scene," in 'Revolusi diri' (self-revolution): personal choice, collective identity and subcultural change in the Bandung straight edge scene, The Australian Sociological Association, 2006, Electronic-Publication.
11. J. Zhao, C. Guo, and L. Yang, "An Analysis of the Proletarian Party’s Self-revolutionary Thought during Lenin’s New Economic Policy Period," *Journal of Yunnan Agricultural University (Social Science)*, vol. 19, no. 5, pp. 171–176, 2025.
12. L. Fang and K. Xie, "Exploring the Modernity Empowerment of Self-Revolution on Chinese-Style Modernization," *Journal of Shandong University (Philosophy and Social Sciences)*, no. 5, pp. 1–12, 2025.

13. Y. Yang and J. Wang, "An empirical study on the impact of internal governance mechanisms on organizational resilience in artificial intelligence companies," *Journal of Hangzhou Dianzi University (Social Sciences)*, vol. 21, no. 1, pp. 1–11, 2025, doi: 10.13954/j.cnki.hduss.2025.01.001.
14. S. Guo, "The practical logic and experience of the anti-corruption struggle from the perspective of self-revolution in the new era," *Theoretical Guide*, no. 12, pp. 90–97, 2019.
15. Z. Jia and Y. Xing, "Advancing the construction of anti-corruption institutional mechanisms under the guidance of the strategic thought of self-revolution," *Theoretical Discussion*, no. 4, pp. 22–28, 2023.
16. G. Wu and D. Tang, "The CPC's Resolution of 'Historical Periodicity' Through Self-revolution in the New Era: Logic, Experience and Methods," *Journal of Southwest University Social Science Edition*, vol. 48, no. 1, pp. 50–61, 2022.
17. G. Hong, "The Great Spirit of Party Building The Key to Leading the Party to Realize Self Revolution," *Journal of Chongqing Jiaotong University Social Sciences Edition*, vol. 26, no. 2, p. 1, 2026.
18. Y. Wu, W. Chen, and H. Yan, "The accountability system for operation and investment and audit quality of state-owned enterprises," *International Review of Financial Analysis*, vol. 96, Art. no. 103680, 2024.
19. S. Dong and S. Fan, "Structural integration of technological implementation: A logical analysis of the operation of digital supervision systems," Learning and Practice, no. 4, pp. 22–33, 2024, doi: 10.19624/j.cnki.cn42-1005/c.2024.04.005.

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 Publisher and/or the editor(s). Publisher 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.