

Rich in Discourse, Thin in Evidence: Mapping Chinese Scholarship on Western Civilization with AI Agents

Yang Dong

Beijing International Studies University, 100024, Beijing, China.

Abstract In contemporary China, Western civilization functions less as an object of historical inquiry than as a comparative referent for theorizing China's own civilizational path — yet the structure of this rapidly expanding literature remains unmapped. This study analyzes 300 articles on Western civilization published in CSSCI and Peking University Core journals (2014–2026), coupling bibliometric analysis with large language model (LLM) content synthesis in an auditable AI-agent workflow that integrates co-word network analysis, semantic topic modeling, and record-level coding of argumentation and evidence. Results document exponential growth after 2021 (a doubling time of roughly 2.3 years) and a discourse core organized around Chinese modernization and the “new form of human civilization,” while Western civilization history and literary-cultural studies contract in share. Theoretical-deductive (58.0%) and critical-hermeneutic (27.7%) argumentation dominate, and empirical-inductive work accounts for only 8.0%. The field is thus discourse-rich but evidence-poor, expanding along policy-coupled rather than evidentially consolidated lines. Four research gaps are identified, empirical, methodological, theoretical, and pedagogical, and show how AI agents can convert them into affordances for literature-mediated undergraduate teaching, operationalizing the research–teaching nexus.

Keywords bibliometrics; scoping review; AI agent; large language model; Western civilization

Received: September 11, 2026

Revised: September 27, 2026

Accepted: September 28, 2026

Published: October 23, 2026

Copyright: © 2026 by the authors. Licensee Axon Academic Publishing Institute, Hong Kong, China. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).
DOI: 10.67148/fahss-2026-263

1. Introduction

Questions of civilization have moved to the center of intellectual debate in contemporary China. As Chinese scholarship has articulated frameworks such as Chinese modernization and a new form of human civilization, the Western civilization tradition - its historical trajectory, its self-understanding, and its normative claims - has become an unavoidable point of reference, comparison, and contention. The result is a substantial and rapidly growing body of research published in Chinese core journals that engages Western civilization not only as an object of historical study but also, and predominantly, as a comparative foil for theorizing China's own civilizational path. Understanding how this field is organized - who publishes where, around which

concepts, with what argumentative habits, and with which blind spots - is a task of information science as much as of area studies.

The analytical challenge is methodological. Established science-mapping tools such as CiteSpace [1] and VOSviewer [2] are built around citation networks, co-citation structures, and index-term co-occurrence. They excel where citation data are abundant, but they are structurally constrained when a national database such as CNKI supplies titles, keywords, and abstracts without reference lists or citation counts. Moreover, keyword co-occurrence, however useful, operates below the semantic threshold of a research article: it cannot distinguish a normative policy discourse from an archival historical study that shares the same keywords. Large language models (LLMs), and the autonomous agents built upon them [3], change this situation. An LLM agent can read every abstract in a corpus, assign documents to semantically grounded topics, extract argumentation and evidence characteristics, and call external tools for the quantitative work that must remain in code - thereby combining machine-scale coverage with human-like interpretive depth [4].

This study therefore pursues two coupled aims. Empirically, it maps Chinese core-journal scholarship on Western civilization indexed in CNKI between 2014 and 2026, treating the 300 retrieved articles as a bounded corpus for a scoping review informed by the PRISMA-ScR framework [5]. Methodologically, it demonstrates and documents an AI-agent-driven analytical workflow in which an LLM agent executes the full pipeline - data audit, bibliometric computation, semantic topic modeling, content analysis, and meta-aggregation - under transparent, human-specified protocols. Five research questions, framed by the population-concept-context (PCC) structure, guide the review:

RQ1. How is the literature on Western civilization in Chinese CSSCI and Peking University Core journals distributed across years, journals, institutions, and authors?

RQ2. What is the knowledge structure of the field, as revealed by keyword co-word networks, thematic clusters, and core versus peripheral themes?

RQ3. How have themes evolved, and which themes are emerging, declining, or constitutive of research frontiers?

RQ4. What patterns characterize the literature in terms of theoretical framing, methods, argumentation structures, and types of evidence?

RQ5. Where are the research gaps, and how can they provide an empirical basis for AI-agent-supported teaching that uses scholarly literature as a pedagogical intermediary?

The study makes three contributions. Methodologically, it specifies a reproducible, citation-free analytical pathway for national databases, in which quantitative computations remain in code while interpretive work is delegated to a logged and auditable LLM agent. Empirically, it offers the first systematic bibliometric and scoping account of Chinese core-journal scholarship on Western civilization, including a topic system, an evidence-gap matrix, and a typology of research gaps. In applied terms, it demonstrates how the identified gaps can be converted into concrete design affordances for AI-agent-mediated teaching, thereby linking bibliometric evidence to the research-teaching nexus. The remainder of the paper reviews the relevant literature, details the methodology, reports findings organized by the research questions, and

discusses implications, limitations, and future directions.

2. Literature Review

2.1 From Tool-Based to AI-Agent-Driven Bibliometrics

Bibliometric mapping of research fields rests on a small set of regularities and techniques formulated long before software implementations: the scattering of literature across journals [6], the productivity distribution of authors [7], and the exponential growth of science [8], together with co-word analysis for structuring thematic space [9]. Modern software embedded these foundations in accessible tools: CiteSpace detects trends and transient patterns through citation networks [1]; VOSviewer constructs maps of terms, authors, or journals from co-occurrence data [2]; and bibliometrix consolidated the workflow in an open R environment [10]. Standard guidelines for such studies recommend transparent protocols spanning data cleaning, network analysis, and co-word clustering [11]. Complementary techniques detect structural dynamics: Louvain community detection unfolds network clusters efficiently [12], burst detection identifies suddenly intensifying terms [13], and field-evolution measures track themes across time slices [14].

These toolchains share two constraints that matter for the present corpus. First, they privilege citation-based relations - co-citation, coupling, and citation-based impact - which are unavailable in the dataset analyzed here, as CNKI export supplies no citation counts and no reference lists. Second, their semantic resolution is lexical: terms co-occur, but the argumentative mode behind a term is invisible. The semantic turn in topic modeling addressed the second constraint by embedding documents and clustering them semantically, with BERTopic combining transformer embeddings with class-based term weighting [15]. The agentic turn goes further: LLM-based autonomous agents integrate planning, memory, tool calling, and retrieval-augmented generation (RAG) so that a single system can, in principle, read every document, decide which instrument to apply, execute code for quantitative steps, and synthesize the results into review-level claims [3]. Bibliometric analysis thereby becomes a candidate for full agentic pipelines, provided that the division of labor between statistical computation and language-model interpretation is explicitly documented - the position this study takes.

2.2 Core Capabilities of Large Language Models and AI Agents

Four capabilities of LLMs and LLM-based agents underpin the workflow reported here. Retrieval-augmented generation grounds model outputs in retrieved documents, mitigating fabrication by conditioning generation on source texts [4]; in a bibliometric setting this allows record-level citation of the corpus itself. Tool calling lets an agent delegate quantitative work to executed code, keeping statistics in verifiable scripts rather than in model memory. Multi-agent orchestration decomposes complex tasks - for example, parallel record extraction followed by aggregation - across specialized roles within a controlled workflow [3]. Finally, human-machine collaborative verification treats the model as a fallible analyst whose categorical judgments must be re-tested and whose outputs should remain auditable; reliability statistics and recommendation for human double-coding are therefore integral, not optional.

These capabilities have direct educational relevance. LLMs support personalized

explanation, dialogue-based tutoring, and content generation at scale, while raising documented risks of hallucination, over-reliance, and shallow learning that call for pedagogical safeguards and teacher mediation [16]. A recurring implication of that literature is that LLM-supported learning works best when tasks are grounded in authentic source materials - a condition that scholarly corpora, organized and retrieved through RAG, are well positioned to satisfy.

2.3 Chinese Scholarship on Western Civilization and Its Teaching Dimension

The dataset analyzed here indicates that Chinese core-journal scholarship engaging Western civilization clusters into several strands. One strand is civilizational discourse: systematic theorization of Chinese modernization, a new form of human civilization, and mutual learning among civilizations, frequently positioned against Western modernity [17,18,19]. A second strand examines exchange and mutual learning in concrete historical and institutional settings, including critiques of Eurocentrism in classical scholarship [20] and narrative analysis of Sino-Western interaction [21]. A third strand is world-history scholarship proper: studies of Greek origins [22], the historiographic construction of the West as an object [23], and medieval European institutions [24]. A fourth, comparatively small, strand addresses encounter history and modern Chinese civilizational consciousness [25] and the international communication of Chinese civilization.

The teaching dimension of this scholarship is thin but suggestive. World-history textbooks and civilization-history curricula have been discussed as instruments of moral and humanistic education [26,27], the Great Books tradition has been examined for organizing Western civilization learning around enduring ideas [28], and general education in the digital-intelligence era has been theorized as a buffer against civilizational crisis [29]. What has not existed, to our knowledge, is any systematic bibliometric mapping of the field as a whole, in Chinese or English - nor an account of how its literature could be mobilized for teaching through AI agents. The present study addresses both absences.

2.4 Research Gaps and Contributions of the Present Study

Three gaps follow. First, an empirical gap: no prior study has mapped the structure, evolution, and evidential profile of Chinese core-journal scholarship on Western civilization, despite its scale and policy salience. Second, a methodological gap: prevailing bibliometric toolchains presume citation data and keyword-level semantics, whereas national corpora without citation fields - and with rich abstracts - require a different, semantically deeper and fully code-based pathway whose agent-driven realization has not been specified. Third, an application gap: the potential of this literature to serve as teaching material, mediated by AI agents, has been argued rhetorically but never grounded in a systematic account of what the literature contains and lacks. The study contributes on all three fronts: a documented agent pipeline; an empirical map with a gap typology; and an evidence-linked mapping from gaps to pedagogical agent capabilities.

3. Methodology

3.1 Research Design

The study implements an AI-agent-assisted hybrid of bibliometric analysis and scoping review. The bibliometric component covers descriptive statistics, Bradford-zone analysis, co-word network analysis, community detection, and strategic diagramming; the review component follows the PRISMA extension for scoping reviews (PRISMA-ScR) for reporting [5] and the PRISMA 2020 flow logic for study identification [30]. Qualitative synthesis follows the meta-aggregative path specified by the Joanna Briggs Institute, in which findings are assembled into descriptive categories and then into synthesized findings [31]. No statistical meta-analysis was attempted, for three reasons stated in advance: the included literature is dominated by humanistic interpretation and normative argument; outcome constructs are heterogeneous across studies; and no extractable effect estimates exist. Reporting of the integrative synthesis is informed by the SWiM guideline for syntheses without meta-analysis [32].

Transparency of execution is a design principle rather than an afterthought. All interpretive steps - topic assignment, argumentation classification, evidence extraction - were executed by a large language model agent [MODEL VERSION] under human-specified protocols, and all quantitative steps were executed in Python (pandas, networkx, matplotlib) as script-based computations. The agent's decision trail is preserved in cumulative batch files, so that every aggregate figure reported below is traceable to record-level extractions.

3.2 Data Source and Search Strategy

Data were retrieved from the China National Knowledge Infrastructure (CNKI) with the topic term Western civilization (xi fang wen ming), restricted to journal articles in CSSCI-indexed or Peking University Core journals. The search was executed on [TO BE FILLED] and returned 300 records. Each record carries ten fields: source database, title, authors, affiliations, journal, keywords, abstract, publication time, funding agency, and publication year. Two properties of the dataset bound the analysis and are reported candidly: the dataset contains no citation-frequency field and no reference lists. Consequently, citation-based analyses - highly cited papers, h-type indices, co-citation, and bibliographic coupling - are outside the scope of this study by construction, and research-frontier identification proceeds through three declared approximations described in Section 3.5.

3.3 Phase 0: Data Audit and Cleaning

The exported file was parsed programmatically into a case-regular data frame (Figure 1). Title-journal deduplication identified no duplicates: all 300 records are unique, so the PRISMA-ScR flow reports 300 identified and 300 included records. Missing-value profiling showed complete title, journal, abstract, and year fields; authorship was missing for one record (0.33%), affiliations for eight (2.67%), keywords for 23 (7.67%), and funding information for 95 (31.67%). Publication year was missing in the year field for three records and was recovered from the publication-time field. Keyword normalization proceeded in two steps: multi-valued keyword fields were split; then 24 synonym or near-synonym mappings were applied (for example, civilizational mutual

learning, civilizational dialogue, and exchange and mutual learning were merged; Western-centrism absorbed Western-centrist theory; the clash of civilizations absorbed clash of civilization variants), reducing unique keywords from 723 to 701. A bilingual keyword table was then fixed, and all figures render keywords in English to avoid font rendering issues and to serve international readers.

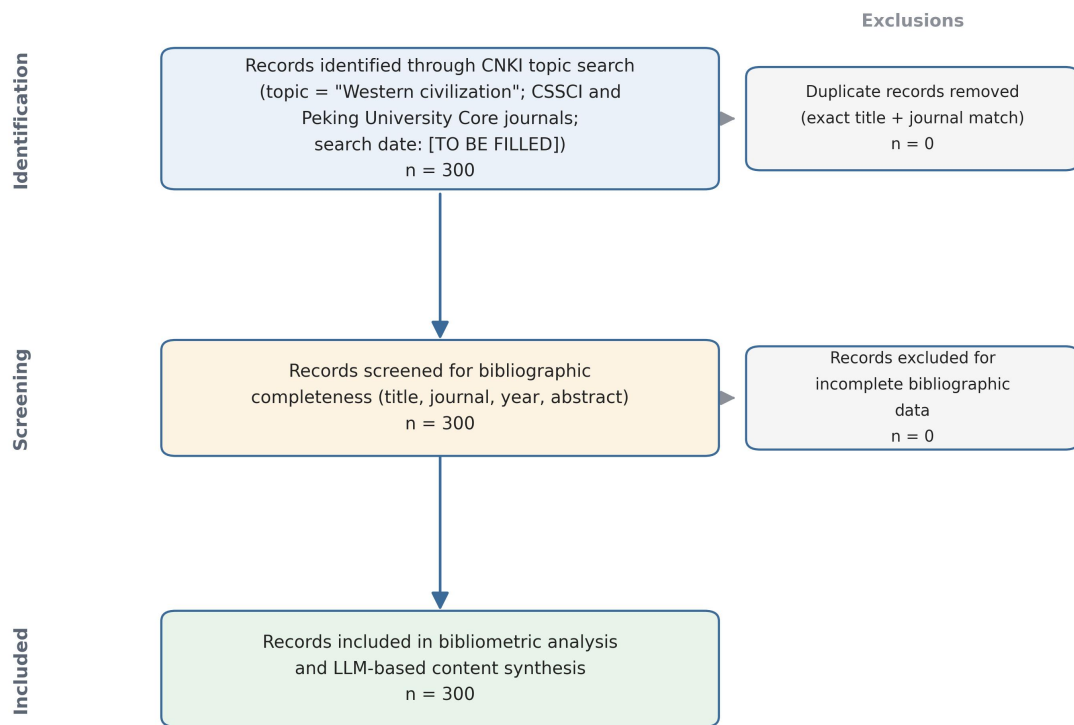


Figure 1. PRISMA-ScR flow diagram of study identification and inclusion.

Note. The search date placeholder is to be completed at submission.

3.4 Quantitative Bibliometric Procedures

Annual output was fitted with an exponential model $N(t) = ae^{bt}$ estimated on complete publication years 2014-2025; because 2026 is an incomplete year, its 58 records (through September) are displayed but excluded from the fit. The resulting doubling time is reported as a Price-law-style approximation [8], not as a formal test of Price's law. Periodization combined inflection detection in the annual series with substantive anchoring to field events, yielding three periods used throughout the analysis. Journal concentration was assessed with Bradford's law by ranking journals by article count and drawing zone boundaries at approximately equal article counts [6]. Author productivity was compared with Lotka's classical distribution by contrasting the observed share of single-article authors with the Lotka theoretical value [7]; given the small population, this comparison is likewise declared an approximation. Collaboration structures were built as author-author and institution-institution graphs from co-authorship and shared-article affiliations; because these networks proved sparse and fragmented, they are reported through network statistics and tables rather than network figures, with the

criteria stated in the findings.

Co-word analysis used the normalized keyword lists. Protocol set the inclusion threshold at frequency of at least 5, with a pre-specified fallback: if fewer than 30 terms qualified, the threshold would drop to 3 with a note. Only 22 terms met the first threshold, so the network was built on the 48 terms with frequency of at least 3, which produced 228 co-occurrence edges. Community detection applied the Louvain algorithm (networkx implementation) with resolution fixed at 1.0, running 50 randomized seeds and retaining the partition with the highest modularity ($Q = 0.212$), which yielded six clusters [12]. Weighted degree and betweenness centrality were computed on the full weighted network. For visualization, edges with weight of at least 2 (82 edges) are drawn to preserve legibility, with node size proportional to frequency and color encoding community membership. Strategic coordinates followed Callon's framework [9]: cluster centrality is the sum of external edge weights divided by cluster size, and cluster density is $2W_{\text{int}} / (n(n-1))$, where W_{int} is the summed internal edge weight and n the number of cluster terms; quadrant boundaries are the respective means across clusters.

3.5 LLM-Based Semantic Topic Modeling and Frontier Approximation

Semantic topics were induced by the LLM agent reading the complete abstract of every record - a substitution, declared here, for embedding-based neural topic modeling of the BERTopic class [15], which cannot be assumed available in every execution environment and which the agent replaces with direct semantic reading under a stabilized codebook. Processing was strictly batched: 25 records per batch, 12 batches, with cumulative extraction files written after each batch and never reconstructed from memory; record-level assignments were appended to a single extraction table. Thirteen topics satisfied the protocol range of 8-15. Topic labels, definitions, and representative studies are reported in Table 3; topic-time matrices were computed on the three periods.

Research-frontier identification could not use citation-based or Kleinberg-exact methods, and three approximations were declared instead: (a) recency of thematic mass (post-2022 topic shares); (b) keyword growth rates between adjacent time slices as an approximation of burst detection [13], computed as the difference of keyword frequencies between the pre-2022 window and the 2022-2026 window, with positive change signaling candidate bursts; and (c) semantic-thematic dynamics from the topic-period matrix. All three proxies are reported as approximations, in line with the field-evolution logic of Cobo et al. [14] without claiming its full machinery.

3.6 Content Analysis and Qualitative Meta-Aggregation

The same batched reading extracted, for each record, the research question, theoretical framework, method type, evidence type, one to two core findings, author-stated limitations, argumentation structure, and teaching relevance. Argumentation structure used a five-class scheme: theoretical-deductive, empirical-inductive, critical-hermeneutic, instructional-design, and review-synthesis. Evidence type distinguished purely theoretical argument, primary textual and historical sources, secondary literature, literary and artistic corpora, and case, policy, media, or interview material. Teaching relevance was coded 0-3 (none, incidental, partial, or

central relation to teaching). The evidence matrix underlying these codes is an internal working file; aggregated distributions appear in Tables 4 and 5. Meta-aggregation proceeded in the JBI three levels - findings, descriptive categories, synthesized findings [31] - with every category anchored to the topic system to preserve traceability.

3.7 Reliability Procedures and Agent Transparency

Because a single agent executed the extractions, a reliability check was built into the protocol: a random sample of 15% of records ($n = 45$, drawn with a fixed seed) was re-read in a second, record-shuffled extraction pass, and the two passes were compared on topic, argumentation type, and evidence type. Agreement was 100% for topic (Cohen's kappa = 1.00), 95.6% for argumentation type (kappa = 0.92), and 91.1% for fine-grained evidence type (kappa = 0.86; 93.3% under a five-level grouping, kappa = 0.89). These figures should be read as intra-agent test-retest consistency rather than fully independent double-coding, because both passes were performed by the same model in the same session; we therefore recommend independent human coding of a further sample before any use of these categories for evaluative purposes. All statistical values reported in this paper derive from executed Python computations over the audited dataset; no statistic is estimated from model memory.

4. Findings

4.1 Publication Landscape, Periodization, and Collaboration (RQ1)

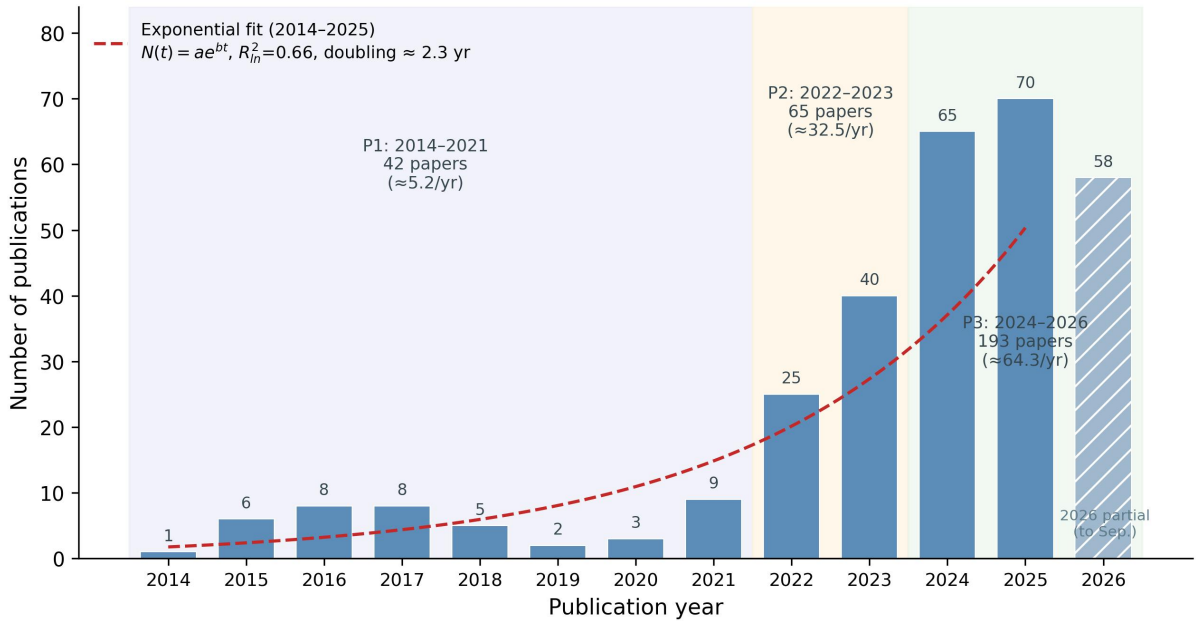


Figure 2. Annual publication counts with periodization and exponential trend, 2014-2026.

Note. The hatched 2026 bar covers January-September (incomplete year); the exponential curve is fitted to complete years 2014-2025 ($R^2 = 0.66$ on the log scale; doubling time 2.3 years).

Annual output rose from a single article in 2014 to 70 in 2025, with 58 articles already published in the first nine months of 2026 (Figure 2). The exponential fit over complete years 2014-2025 yields $N(t) = ae^{0.3053t}$ with $R^2 = 0.66$ on the log scale, corresponding to a doubling time of approximately 2.3 years; interpreted as a

Price-law-style approximation [8], this indicates accelerating, far-from-saturated growth. The trend divides into three periods: P1 (2014-2021), a gestation phase averaging 5.2 articles per year; P2 (2022-2023), an acceleration phase averaging 32.5 articles per year; and P3 (2024-2026), an explosion phase averaging 64.3 articles per year. The P1-P2 break coincides with the articulation of Chinese modernization as a flagship theoretical program in late 2022, and the P2-P3 expansion accompanies the Global Civilization Initiative (2023) and the consolidation of official cultural-civilizational discourse; the association is interpretive, inferred from the timing and content of the corpus rather than from any causal test.

Table 1 Bradford Zone Distribution of Journals and Articles

Zone	Journals (n)	Articles (n)	Share of articles (%)
Core	24	100	33.3
Related	47	100	33.3
Peripheral	100	100	33.3
Total	171	300	100.0

The 300 articles are scattered across 171 journals, and the Bradford analysis partitions them into a core zone of 24 journals carrying 100 articles (33.3%), a related zone of 47 journals carrying 100 articles, and a peripheral zone of 100 journals carrying 100 articles (Table 1). The core zone is dominated by political-theory, Marxism-studies, and comprehensive social science journals; history and world-history journals (e.g., *World History*; *Historiographical Research*) appear in the core zone but with smaller counts, and no dedicated Western-civilization or area-studies journal appears at all. This composition is itself a finding: scholarship referencing Western civilization in Chinese core journals is published, almost entirely, in discourse-oriented venues rather than in specialized historical venues.

Authorship is extreme in its dispersion. The 300 articles yield 358 distinct authors, 88.5% of whom appear exactly once - far above the 60.8% expected under Lotka's classical law [7], indicating a population of occasional contributors rather than a settled specialist community (an approximation, as declared in Section 3.4). The most prolific authors are Cao Shunqing (seven articles), followed by Han Qingxiang and Wang Yuchen (four each); nine further authors appear three times. Collaboration is similarly loose: 192 articles (64.0%) are single-authored, and the co-authorship network over 205 connected authors contains only 136 edges, a density of 0.0065, 89 components, a largest component of just six authors, and 72 isolated dyads - that is, almost all observed collaboration is one-off. The institution-institution network is comparably sparse (133 edges across 75 multi-institution articles, 25.0%), with the most frequent repeated tie linking the Institute of Philosophy of the Chinese Academy of Social Sciences and the School of Philosophy of the University of Chinese Academy of Social Sciences (four co-publications). Because both networks fail basic connectivity criteria, they are reported as statistics rather than figures, in line with the protocol of Section 3.4.

4.2 Knowledge Structure: Co-Word Network and Thematic Clusters (RQ2)

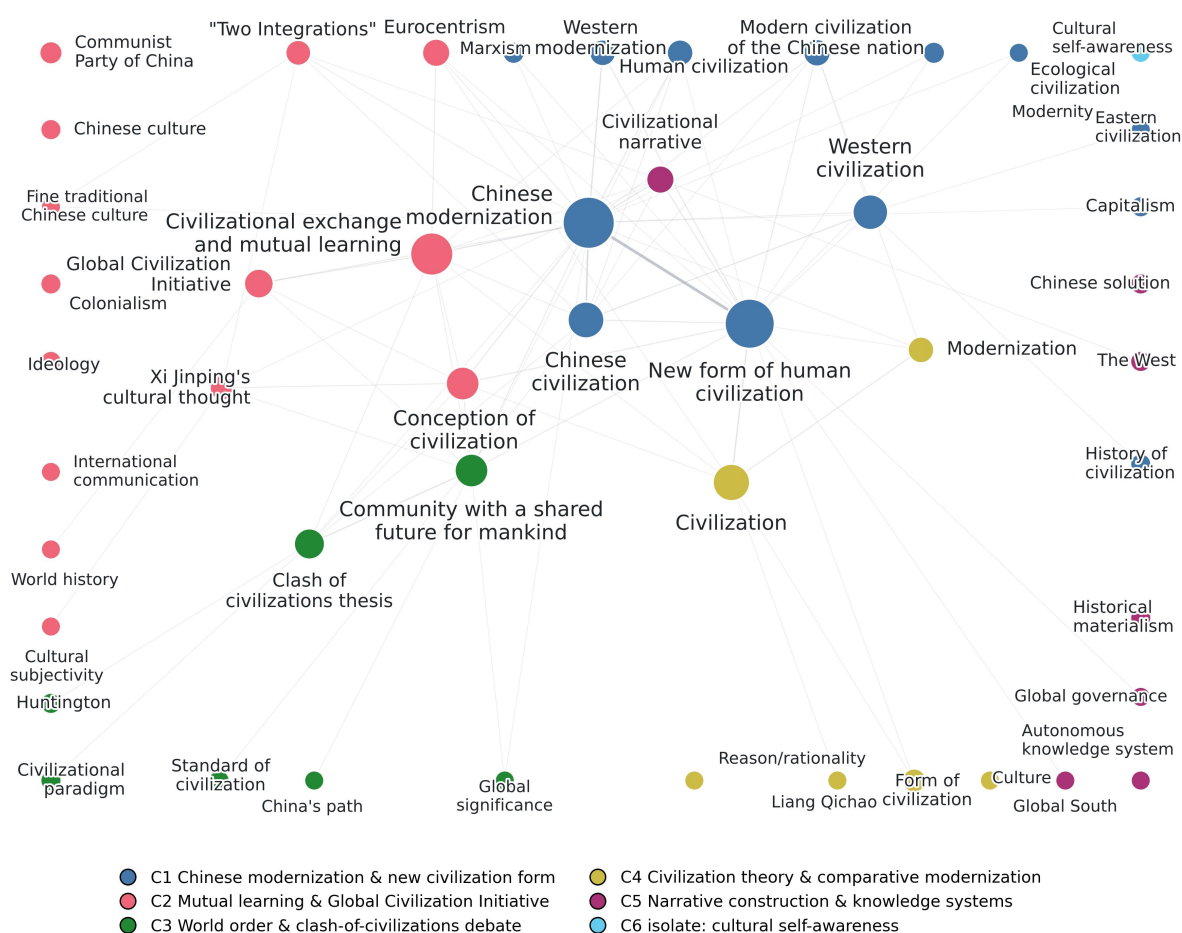


Figure 3. Co-word network of the 48 normalized author keywords.

Note. Node size is proportional to keyword frequency; node color indicates Louvain community; edges with weight ≥ 2 are shown; edge width and opacity are proportional to co-occurrence weight.

Figure 3 displays the co-word network of the 48 normalized keywords occurring at least three times. Its hub structure is unambiguous: Chinese modernization (frequency 57; weighted degree 114), new form of human civilization (52; 96), and civilizational exchange and mutual learning (37; 48) dominate the network, and the strongest single edge links Chinese modernization to new form of human civilization (25 co-occurrences). Betweenness centrality identifies new form of human civilization (0.138) and civilizational exchange and mutual learning (0.119) as the principal bridges between thematic regions. Western civilization itself is volumetrically substantial (frequency 23, weighted degree 26) but structurally peripheral: its betweenness (0.036) is roughly a quarter of the top bridging values, meaning that the term appears frequently but rarely mediates the network's thematic traffic.

Table 2 Co-Word Clusters, Strategic Coordinates, and Quadrant Assignment

Cluster	Label	Terms (n)	Top terms (frequency)	Density	Centralit y	Quadrant
C1	Chinese modernization and civilizational renewal	13	Chinese modernization (57); new form of human civilization (52); Chinese civilization (25); Western civilization (23); modern civilization of the Chinese	1.50	10.92	I: motor

			nation (11)			
C2	Civilizational exchange, mutual learning, and the GCI	14	civilizational exchange and mutual learning (37); conception of civilization (20); Global Civilization Initiative (14); Western-centrism (12); "Two Integrations" (9)	0.57	8.14	IV: transversal
C3	World order and the clash-of-civilizations debate	7	community with a shared future for mankind (20); clash of civilizations thesis (16); Huntington (4); standard of civilization (4)	0.95	8.86	I: motor
C4	Civilization theory and comparative modernization	6	civilization (26); modernization (10); form of civilization (7); Liang Qichao (3)	0.93	8.17	I: motor
C5	Narrative construction and autonomous knowledge systems	7	civilizational narrative (12); historical materialism (5); Chinese solution (4); autonomous knowledge system (3)	0.52	6.71	III: weak
C6	Isolated term	1	cultural self-awareness (3)	0.00	0.00	III: isolate

Note. N = 48 terms with normalized frequency ≥ 3 ; clusters from Louvain community detection ($Q = 0.212$). Density = $2W_{\text{int}} / (n(n-1))$; centrality = W_{ext} / n (see Section 3.4). Quadrants are assigned relative to cluster means (centrality 7.13; density 0.75). GCI = Global Civilization Initiative.

Louvain community detection yields six clusters (Table 2; Figure 3 colors). Cluster C1, Chinese Modernization and Civilizational Renewal (13 terms), couples Chinese modernization, the new form of human civilization, Chinese civilization, and Western civilization with Western modernization and Marxism. Cluster C2, Civilizational Exchange, Mutual Learning, and the Global Civilization Initiative (14 terms), gathers exchange, the conception of civilization, the initiative, Eurocentrism, the Two Integrations, and Xi Jinping's cultural thought. Cluster C3, World Order and the Clash-of-Civilizations Debate (7 terms), centers on the community with a shared future for mankind, the clash thesis, Huntington, and the standard of civilization. Cluster C4, Civilization Theory and Comparative Modernization (6 terms), holds the general terms civilization, modernization, form of civilization, reason, and Liang Qichao. Cluster C5, Narrative Construction and Autonomous Knowledge Systems (7 terms), links civilizational narrative, historical materialism, the Global South, and autonomous knowledge systems. Cluster C6 is a single isolated term, cultural self-awareness, whose co-occurrence partners all fall below the inclusion threshold.

Figure 4. Strategic diagram of the six co-word clusters. Centrality (x-axis) measures external linkage; density (y-axis) measures internal cohesion; dashed lines mark cluster means; point size is proportional to cluster size in terms.

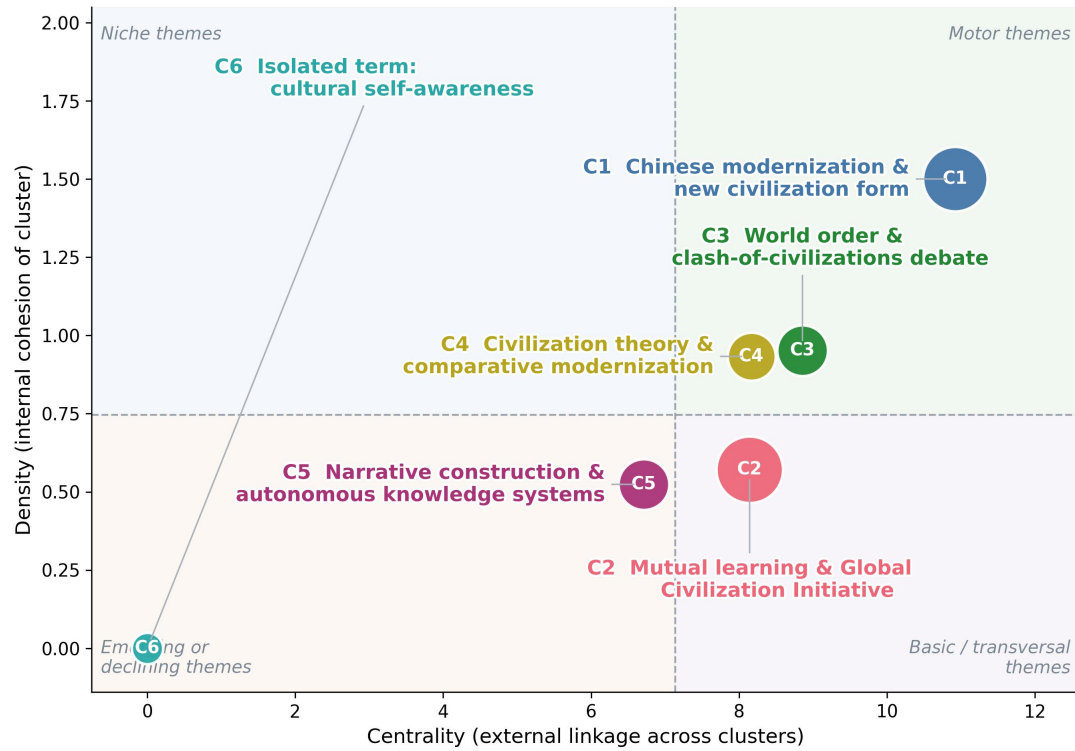


Figure 4. Strategic diagram of the six co-word clusters.

Note. Centrality (x-axis) measures external linkage; density (y-axis) measures internal cohesion; dashed lines mark cluster means; point size is proportional to cluster size in terms.

The strategic diagram (Figure 4) locates C1, C3, and C4 in the motor quadrant (high density and high external centrality): C1's density of 1.50 and centrality of 10.92 make it the field's engine, while C3 and C4 are well-developed and well-connected. C2 sits in the transversal-basic quadrant: highly central to the rest of the network (8.14) but loosely internally integrated (density 0.57), indicating a vocabulary shared across themes rather than a bounded specialty. C5 falls in the weak quadrant (centrality 6.71; density 0.52), and the C6 isolate occupies the origin. Read together, the network answers RQ2: the knowledge structure is organized by a discourse core that references Western civilization systematically but treats it principally as the comparator against which Chinese modernization and new-civilization claims are elaborated.

4.3 Thematic Evolution and Research Frontiers (RQ3)

Table 3 The 13-Topic Semantic Structure of the Corpus

Topic	Label	Core definition	n (%)	Representative studies
T01	Chinese modernization and the new form of human civilization	Theoretical elaborations of China's modernization path as creating a new form of civilization surpassing Western modernity	75 (25.0)	Zhang and Zhou (2026); Xu (2026); Han and Wang (2024)
T02	Civilizational exchange, mutual learning, and the	Normative and historical studies of inter-civilizational exchange and the	27 (9.0)	Cao and Ming (2026); Liu (2025);

	GCI	Global Civilization Initiative		Yang (2024)
T03	Critiques of Western modernity and Eurocentrism	Critical analyses of Western-centrism, capitalist civilization, and the metaphysics of modernity	19 (6.3)	Pan (2025); Ren (2024)
T04	World order, the clash-of-civilizations debate, and international relations	Analyses of Huntington's thesis, standards of civilization, and world-order transformation	25 (8.3)	Huang (2023); Liu and Li (2022)
T05	Civilization theory and comparative civilizational concepts	Conceptual history and comparative theory of civilization, including comparative philosophy and aesthetics	46 (15.3)	Ren (2026); Li (2022); Yang (2023)
T06	Western civilization history and historiography	Historical and historiographical studies of Western civilization proper and its narrative construction	25 (8.3)	Chen (2026); Yan (2024); Yan (2022); Zhang (2025)
T07	Sino-Western encounters and modern Chinese civilizational consciousness	Intellectual and social histories of China's encounter with Western civilization, late Qing to Republican era	15 (5.0)	Feng (2025)
T08	International communication of Chinese civilization	Discourse power, international communication, and the global reception of Chinese civilization	11 (3.7)	Zou (2025)
T09	Ecological civilization and Western environmental thought	Theories of ecological civilization in critical relation to Western environmentalism	10 (3.3)	Wang and Wang (2025)
T10	Marxist theory of civilization	Marxist and party-theoretical accounts of civilizational forms and transitions	17 (5.7)	Liu and Feng (2025)
T11	Cross-cultural literary, artistic, and representational studies	Literary, artistic, and media studies of mutual representations between China and the West	20 (6.7)	Zhai (2026)
T12	Education and teaching	Scholarship directly addressing curricula, teaching, and general education in relation to civilization	5 (1.7)	Qiang (2023); Liu and Wang (2023); Liu (2016)
T13	Reviews, symposia, and miscellaneous	Book notes, conference reports, and symposium collections	5 (1.7)	—

Note. n = number of articles; percentages of $N = 300$. Topics were induced by LLM-based semantic clustering of all abstracts (see Section 3.5). GCI = Global Civilization Initiative.

LLM-based semantic topic modeling of the 300 abstracts produced 13 topics (Table 3). The largest, T01 Chinese Modernization and the New Form of Human Civilization, accounts for 75 articles (25.0%), followed by T05 Civilization Theory and Comparative Civilizational Concepts (46; 15.3%), T02 Civilizational Exchange, Mutual Learning, and the Global Civilization Initiative (27; 9.0%), T04 World Order and the Clash Debate (25;

8.3%), T06 Western Civilization History and Historiography (25; 8.3%), T11 Cross-Cultural Literary, Artistic, and Representational Studies (20; 6.7%), T03 Critiques of Western Modernity and Eurocentrism (19; 6.3%), T10 Marxist Theory of Civilization (17; 5.7%), T07 Sino-Western Encounters and Modern Chinese Civilizational Consciousness (15; 5.0%), T08 International Communication of Chinese Civilization (11; 3.7%), T09 Ecological Civilization (10; 3.3%), T12 Education and Teaching (5; 1.7%), and T13 Reviews, Symposia, and Miscellaneous (5; 1.7%).

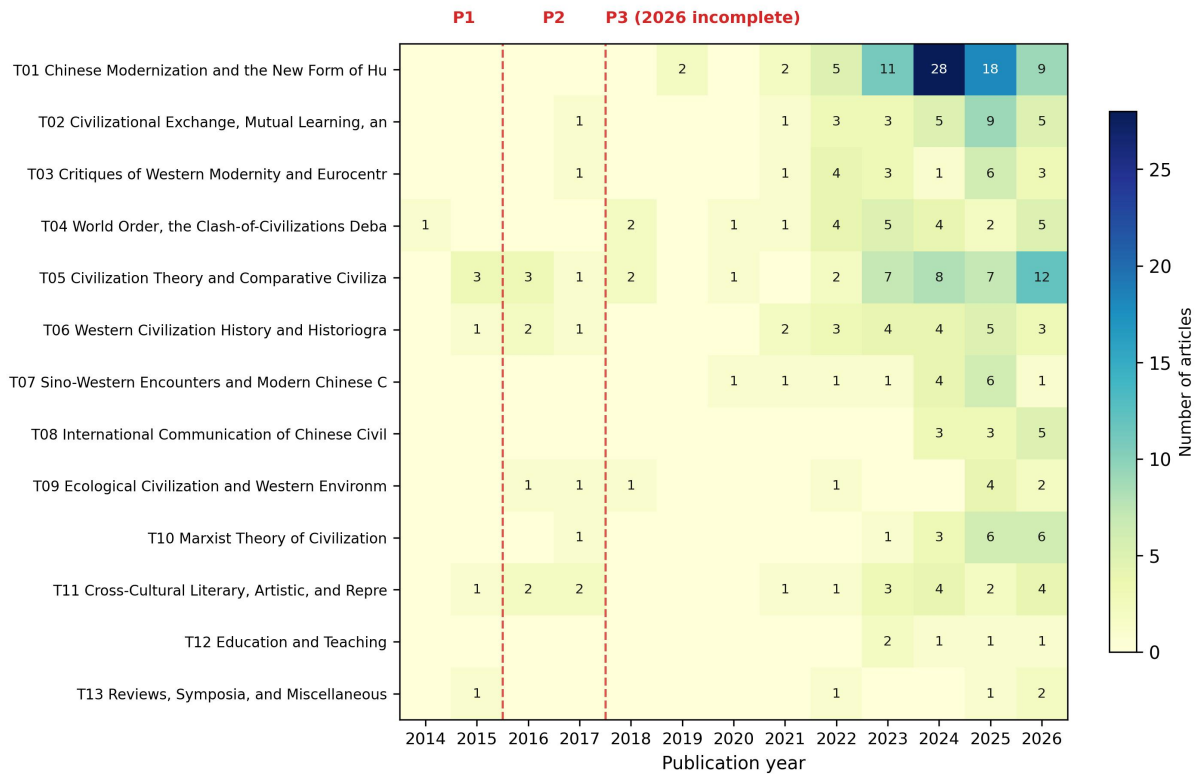


Figure 5. Topic-by-year distribution of the 13 semantic topics (article counts).

Note. Red separators mark the three periods; 2026 covers January-September.

Topic-period shares trace the field's transformation (Figure 5). T01 expands from 9.5% of P1 output to 24.6% in P2 and 28.5% in P3. T08 is a genuine post-2022 emergence: zero articles before 2022, then eleven between 2024 and 2026. T10 rises to 7.8% in P3, tracking the theoretical consolidation of official cultural thought. Conversely, the share of T06 falls from 14.3% in P1 to 6.2% in P3; T11 falls from 14.3% to 5.2%; and T09 falls from 7.1% to 3.1%. Because total output grew, these are share contractions rather than absolute declines - the frontier movement is a relative reallocation of the field's center of gravity from historical and cultural scholarship toward policy-coupled civilizational discourse.

The growth-rate approximation of burst detection sharpens this picture. Comparing the pre-2022 window with 2022-2026, the keywords with the steepest positive change are Chinese civilization (from 0 to 25 occurrences), conception of civilization (0 to 20), Global Civilization Initiative (0 to 14), civilizational narrative (0 to 12), modern civilization of the Chinese nation (0 to 11), the Two Integrations (0 to 9), and Xi Jinping's cultural thought (0 to 8); Chinese modernization itself grows from 4 to 53. No keyword shows a high-frequency early presence followed by near-disappearance - consistent with a young, expanding field in which decline appears only as share loss. Under the three-proxy

frontier definition (recency, growth, semantic dynamics), the current frontiers are the policy-coupled cluster T01-T02 with the adjacent emergent T08 and resurgent T10; mature but relatively contracting areas are T06, T11, and T09; and the teaching-oriented T12 remains marginal at every point in time.

4.4 Argumentation Structures, Evidence Types, and Meta-Aggregation (RQ4)

Table 4 crosses the five argumentation classes with method groups and evidence groups. Theoretical-deductive articles dominate with 174 (58.0%), followed by critical-hermeneutic with 83 (27.7%); empirical-inductive articles number 24 (8.0%), review-synthesis 18 (6.0%), and instructional-design exactly one (0.3%). Methodologically, conceptual and theoretical analysis accounts for 156 articles (52.0%), historical, textual, and philological analysis for 70 (23.3%), comparative analysis for 26 (8.7%), discourse and narrative analysis for 23 (7.7%), reviews and symposia for 20 (6.7%), and empirical qualitative or case analysis for five (1.7%). On the evidence side, 161 articles (53.7%) rest on theoretical argument alone; 67 (22.3%) engage primary textual or historical sources; 55 (18.3%) synthesize secondary literature; 10 (3.3%) analyze literary or artistic corpora; and 7 (2.3%) use case, policy, media, or interview material.

Table 4 Cross-Tabulation of Argumentation Structures With Method Types (Panel A) and Evidence Types (Panel B)

Panel A: Argumentation × method							
Argumentation structure	Conceptual-theoretical	Historical-textual-philological	Discourse-narrative	Comparative	Empirical-case	Review-symposium	Total
Theoretical-deductive	119	16	18	21	0	0	174
Critical-hermeneutic	33	36	5	5	1	3	83
Empirical-inductive	2	18	0	0	4	0	24
Review-synthesis	1	0	0	0	0	17	18
Instructional-design	1	0	0	0	0	0	1
Total	156	70	23	26	5	20	300

Panel B: Argumentation × evidence						
Argumentation structure	Theoretical only	Primary-historical	Secondary literature	Literary-artistic	Case-policy-media	Total
Theoretical-deductive	147	7	19	0	1	174
Critical-hermeneutic	8	41	24	9	1	83
Empirical-inductive	0	19	0	1	4	24
Review-synthesis	6	0	11	0	1	18
Instructional-design	0	0	1	0	0	1
Total	161	67	55	10	7	300

Note. $N = 300$ articles; single counts. Method and evidence categories were coarsened from the extraction codebook (Section 3.6).

The cross-tabulation exposes the field's epistemic architecture. Within theoretical-deductive articles, 147 of 174 rest on theoretical argument alone - the discourse core is argument-on-argument. The critical-hermeneutic class is comparatively

source-engaged: 41 of its 83 articles work with primary texts. Empirical-inductive work is almost entirely historical: 19 of its 24 articles use primary textual or historical sources, and only four use case or policy material. No article in the corpus applies survey, experimental, or computational text methods in a social-scientific sense. The single instructional-design article - a study of printmaking pedagogy informed by a Western art-history survey - stands alone in its class. Teaching relevance coding reinforces the asymmetry: 184 articles (61.3%) score zero; 107 (35.7%) are incidentally related; and only nine articles (3.0%) score 2 or higher, of which just three are centrally about teaching.

Table 5 Meta-Aggregation From Findings to Descriptive Categories and Synthesized Findings

Level 2: Descriptive category	Findings (n)	Level3: Synthesized finding
D1. Chinese modernization constitutes a new civilizational form transcending Western modernity	75	SF1
D2. Mutual learning and the GCI provide a normative order for inter-civilizational relations	27	SF2
D3. Western-centrism and civilizational hierarchy narratives are ideological constructions	19	SF2
D4. World-order theory must move beyond the clash paradigm	25	SF2
D5. "Civilization" is a historically variable concept requiring comparative clarification	46	SF3
D6. Western civilization's historical development was cross-civilizationally constituted	25	SF3
D7. Modern Chinese civilizational consciousness emerged through encounter with the West	15	SF3
D8. Communicating Chinese civilization faces structural discourse asymmetries	11	SF4
D9. Ecological civilization reframes the human-nature relation beyond Western environmentalism	10	SF4
D10. Marxist theory frames civilizational forms and transitions	17	SF4
D11. Cultural and artistic works mediate civilizational perception and mutual learning	20	SF3
D12. Pedagogical dimensions of civilizational scholarship remain undeveloped	5	SF5
Level 3: Synthesized findings		
SF1. Western civilization functions in this literature primarily as a comparative-normative referent for theorizing China's civilizational modernity, rather than as a standalone object of historical-empirical inquiry (from D1, D5).		
SF2. The field's expansion is discourse-coupled: growth concentrates in policy-linked themes that treat inter-civilizational order normatively (from D2, D3, D4).		
SF3. Where historical and textual scholarship exists, it demonstrates the cross-civilizational constitution of both the West and modern Chinese self-understanding, undercutting endogenous-origin narratives (from D5, D6, D7, D11).		

SF4. Claims about communication, ecology, and civilizational transition are formulated theoretically; the evidential base for reception and impact remains unmeasured (from D8, D9, D10).

SF5. The pedagogical translation of civilizational scholarship is nearly absent, constituting the field's clearest actionable gap (from D12).

Note. Findings are 1–2 per included article, consolidated into 300 discrete statements; 5 symposium or note records (T13) were excluded from synthesis. GCI = Global Civilization Initiative.

Meta-aggregation assembled the 300 record-level findings into 12 descriptive categories and five synthesized findings (Table 5). The categories range from the affirmation that Chinese modernization constitutes a new civilizational form transcending Western modernity (75 findings) to the recognition that pedagogical dimensions of civilizational scholarship remain undeveloped (5 findings). The five synthesized findings are: S1, Western civilization functions in this literature primarily as a comparative-normative referent rather than a standalone object of historical-empirical inquiry; S2, the field's dominant epistemic mode is discourse-theoretic, with empirical-inductive work confined to historical studies; S3, the historical-empirical scholarship that does exist demonstrates the cross-civilizational constitution of the West, undercutting endogenous-origin narratives; S4, the field's expansion is discourse-coupled, with growth concentrated in policy-linked themes and evidence density declining as volume rises; and S5, the pedagogical translation of this scholarship is nearly absent, constituting the clearest actionable gap.

4.5 Evidence Gaps and Frontier Typology (RQ5)

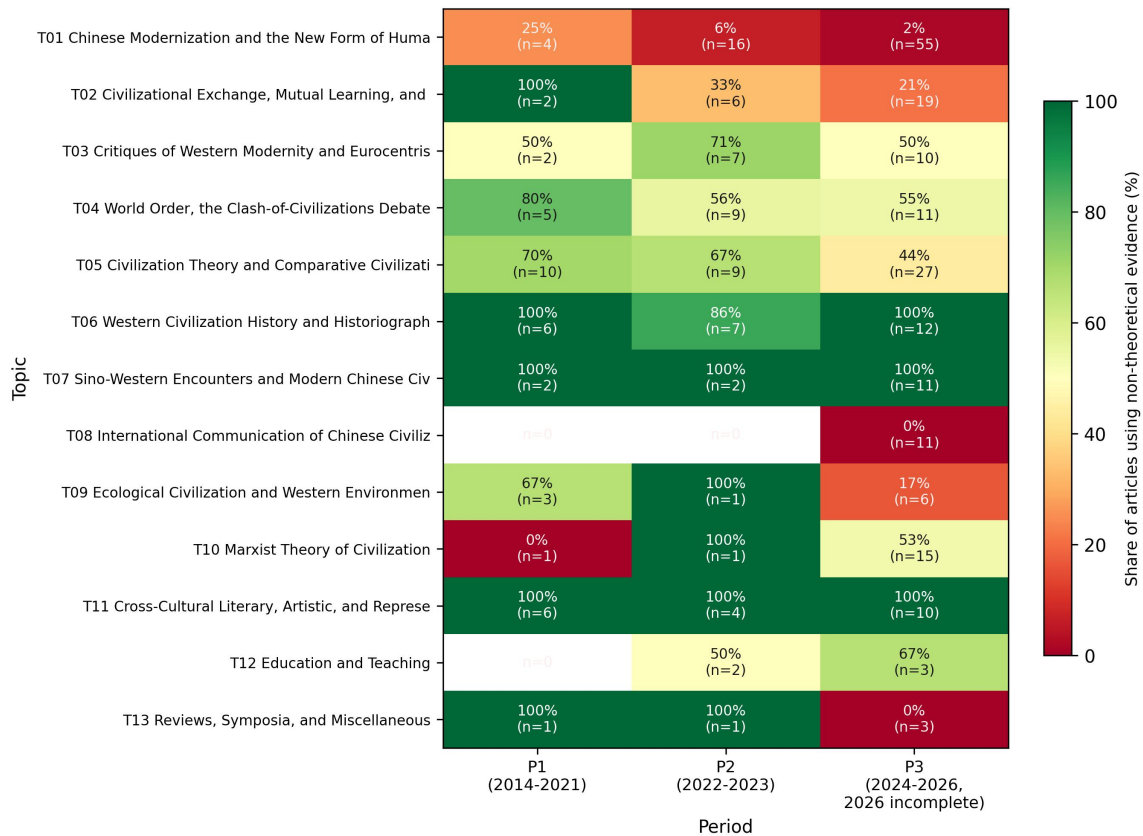


Figure 6. Evidence-gap matrix: topic by period.

Note. Cell color encodes the share of articles using non-theoretical evidence (primary or historical sources, source-grounded synthesis, corpora, or case material); annotations report the share and article count per cell.

The topic-by-period evidence-gap matrix (Figure 6) colors each cell by the share of its articles using any non-theoretical evidence and annotates cell counts. Two patterns define the matrix. First, the evidence-rich zones are the historically grounded topics: T06 Western civilization history maintains 100%, 86%, and 100% non-theoretical share across the three periods; T07 encounters history and T11 literary-cultural studies are consistently near 100%. Second, the evidence-sparse zones are exactly the zones of growth: T01's non-theoretical share collapses from 25% in P1 (one of four articles) to 1.8% in P3 (one of 55). Field-wide, the share of articles using non-theoretical evidence falls from 76.2% in P1 to 53.8% in P2 and 37.3% in P3: the field has more than quadrupled its annual output while more than halving its evidential density.

Table 6 Research Gap Typology and Supporting Evidence

Gap	Type	Description	Evidence
G1	Empirical	The discourse core (T01) shows an empirical-evidence share of 1.8% in 2024–2026; Western civilization history (T06) and exchange processes (T02) lack data-based investigation	Figure 6; Table 3
G2	Methodological	No quantitative or computational methods appear; empirical-qualitative and case designs account for 1.7% of articles	Table 4 (Panel A)
G3	Theoretical	Civilizational concepts are elaborated normatively but neither operationalized nor comparatively tested, leaving comparative theory programmatic	Table 4; Figure 4 (Cluster C4)
G4	Pedagogical translation	1.7% of articles concern teaching; a single article uses instructional-design argumentation; 3.0% of articles have more than incidental teaching relevance	Table 4; Table 5 (SF5)
G5	Reception	Claims about the international reception of Chinese civilization rest on assertion rather than audience or media data (11 articles, all theoretical)	Table 3 (T08); Figure 6

Table 6 consolidates the gaps into a four-type typology. Empirical gap: the discourse core's claims are overwhelmingly unoperationalized, and concrete exchange processes, reception, and audience effects are asserted rather than measured (T08 is entirely theoretical in P3). Methodological gap: no quantitative or computational methods appear; case-based work is nearly absent (2.3% of articles). Theoretical gap: civilizational concepts are elaborated normatively but rarely operationalized or comparatively tested, leaving comparative civilizational theory (T05) largely programmatic. Pedagogical-translation gap: 1.7% of articles concern teaching, one article is instructional-design class, and 3.0% of all articles have more than incidental teaching

relevance - despite a corpus whose subject matter maps directly onto general-education curricula in Western civilization and world history. The frontier typology follows: a mature but evidence-thin discursive core (T01, T02); emerging policy-coupled frontiers (T08, T10); a contracting but evidence-rich historical base (T06, T11); and persistent gap zones (T12 teaching; cross-cultural reception research).

5. Discussion

5.1 Principal Findings in Dialogue with Existing Literature

This review set out to map an unmapped field and to test an agent-driven workflow on it; both aims return instructive results. The bibliometric profile confirms, for a Chinese-language humanities field, patterns that the bibliometric literature associates with young, topically scattered research areas: extreme author dispersion (88.5% single-article authors against a Lotka expectation near 61%), fragmented collaboration (density below 0.01; largest component of six authors), and flat Bradford concentration spread over 171 journals [6,7]. What the standard toolkit could not have shown, and what semantic analysis adds, is the direction of the field's center of gravity: the co-word and topic systems demonstrate that Western civilization is processed by this literature primarily as a comparative referent - the comparator against which Chinese modernization and new-civilization claims are argued (Figure 3; Table 3) - rather than as an autonomous object of historical-empirical study, which occupies only 8.3% of the corpus and a contracting share of it.

The dynamics qualify and, in one respect, counter expectations from established field-evolution models. In the logic of citation-accumulating fields, growth concentrates where evidential consolidation rewards further work [14]. Here the opposite holds: expansion is discourse-coupled. Thematic mass and burst-like keyword growth cluster on policy-linked vocabulary that enters the corpus only after 2022 (Table 3; Figure 5), while the share of articles resting on primary evidence falls as volume rises (Figure 6). The finding does not imply that the discourse literature is without scholarly value - much of it is philosophically serious work in the critical-hermeneutic mode - but it does mean that field-level growth signals agenda-setting by official concepts, not cumulative empirical consolidation. For comparative civilizational studies more broadly, the corpus also contains a substantive counterpoint: where the field is evidence-based, it converges on the cross-civilizational constitution of the West - Greek borrowings from the Near East, the Arabic-mediated Renaissance, and the constructed character of Western civilization as a historiographic and curricular object (Table 5, S3) - which runs against popular endogenous-origin narratives of the West.

5.2 Methodological Implications of the AI-Agent Paradigm

Methodologically, this study is a worked demonstration that an LLM agent can execute a complete review pipeline under audit. Three implications follow. First, semantic depth without citation dependence: because topic assignment and argumentation coding rest on reading abstracts rather than on citation relations, the pipeline remains fully functional on national databases that lack citation and reference fields - the very setting where CiteSpace- and VOSviewer-style workflows are structurally constrained [1,2]. The corpus studied here is exactly such a setting, and the

analysis surfaced distinctions those workflows cannot reach: within the same keyword neighborhood, the agent separated normative civilizational discourse from historical scholarship and from literary-cultural criticism, which is what made the evidence-gap matrix (Figure 6) possible at all.

Second, reproducibility through enforced division of labor. Every quantitative value in this paper comes from executed Python code over an audited dataset; every interpretive value comes from logged batch extractions written incrementally and never reconstructed from memory. This separation is the practical answer to the hallucination risk that attends LLM-assisted research: the model was never allowed to compute statistics, and the code was never asked to interpret. The residual subjectivity sits in categorical judgments - topic borders, argumentation classes - which is precisely where the reliability protocol applies, and where we recommend human double-coding before evaluative use. Third, agentic workflows reposition tool use in science mapping: GUI-driven analysis pipelines hard-code analytical choices, whereas the protocol-plus-agent design documented here makes those choices explicit and revisable - thresholds, fallbacks, approximations, and reliability checks are all declared rather than embedded in software defaults [3,11]. The paradigm is best read not as replacing the analyst but as externalizing the analyst's decision trail.

5.3 From Bibliometrics to the Classroom: AI Agents, Literature as Pedagogical Intermediary, and the Research-Teaching Nexus

Table 7 Mapping Identified Gaps to AI-Agent Capabilities for Literature-Mediated Teaching

Identified gap	AI-agent capability		Literature-to-classroom application	Design safeguard
G1, G4 (evidence-sparse discourse core; undeveloped pedagogy)	RAG-based organization retrieval	literature and	The reviewed corpus (300 records, 13 topics) is organized into a searchable knowledge base; students retrieve record-level evidence when examining claims made in the discourse literature	Every agent answer grounded in and cited to specific corpus records
G2 (no quantitative or computational methods)	Multi-agent review workflows with tool calling	literature	Instructors delegate update scans to agent workflows; students execute guided micro-bibliometric projects (co-word maps, growth curves) on curated subsets	Code and intermediate outputs inspectable; results recomputed by students
G3 (unoperationalized concepts)	Personalized generation	content	Concept explications calibrated to reading level; glossaries linking Chinese civilizational-discourse terms to Western counterparts for English-major undergraduates	Teacher-in-the-loop review of all generated materials
G3 (unoperationalized concepts)	Socratic tutoring	dialogue	Concept-interrogation seminars in which the agent probes operational definitions, boundary cases, and hidden premises of claims about "civilization"	Dialogue logs auditable; teacher moderation of transcripts
G1 (cross-civilizational history)	Bilingual primary-source alignment	and	Pairing Chinese scholarly claims with Western primary sources (e.g.,	Alignment confidence flags; versioned source lists

underrepresented)	contextualization	Renaissance-origins debates with period documents) to enable evidence-based comparison
G5 (unmeasured reception claims)	Source-critical evaluation training	Claim-evidence audits: students evaluate whether assertions about Western decline or against the claim-evidence civilizational reception rest on primary matrix sources or purely theoretical argument

Note. Gaps G1–G5 are defined in Table 6; RAG = retrieval-augmented generation [4].

The gap analysis yields a concrete, evidence-grounded starting point for teaching applications: the corpus is rich in civilizational scholarship yet almost silent on pedagogy (1.7% teaching-oriented articles; a single instructional-design study), even though its subject matter maps directly onto Western civilization and world history curricula and onto general education for English-major undergraduates, who encounter the Western canon through language and literature courses. The few existing points of contact - civilization-history curricula as moral education [26], the Great Books idea [28], general education under digital conditions [29], and teaching-oriented treatments of Greek origins [27] and of the construction of Western civilization as a curricular object [23] - indicate demand without a supply of systematic pedagogical research. Table 7 converts each identified gap into a specific AI-agent capability, so that the teaching argument stays tethered to the bibliometric evidence rather than to enthusiasm.

Two compact design examples illustrate the nexus. First, the co-word knowledge map (Figure 3) can serve as a course-level cognitive map: an instructor organizes a Western civilization module around the map's clusters - the discourse core, the exchange cluster, the clash-of-civilizations debate, and the historical base - and students write position papers that must locate themselves relative to the map, using an agent's retrieval-grounded support to pull the underlying articles and check claims against them [4]. Second, the evidence matrix (Table 4) can generate seminar inquiry: students ask, with agent scaffolding, how many of the field's central claims rest on primary sources rather than theoretical argument, and then run micro-bibliometric replications on teacher-curated subsets - counting keyword growth, or rebuilding a co-word network for one cluster - with the agent calling the code and the students interpreting the output. In both designs, research products feed teaching (maps and matrices become course materials), and teaching feeds research (student replications and classroom questions surface gaps and borderline cases), which is the research-teaching nexus operationalized: research informs teaching, teaching returns evidence and questions to research.

The Socratic-dialogue and source-criticism capabilities in Table 7 address the theoretical gap identified in the findings: a dialogue agent can press students for operational definitions of civilization, civilizational narrative, or mutual learning - exactly the concepts the field uses normatively - while bilingual source alignment pairs Chinese civilizational discourse with Western primary sources so that English-major undergraduates can compare claims on both sides of the translation. Safeguards are non-negotiable: retrieval must cite record identifiers, generated materials require teacher review, and assessment should reward source criticism rather than fluent summary [16].

Read together, the two directions close the research-teaching loop. Research feeds

teaching when bibliometric artifacts - the co-word map, the topic system, the claim-evidence patterns of Table 4 - become course materials that let students encounter the field's structure rather than only its conclusions. Teaching feeds research when classroom work returns to the corpus: student micro-projects re-run analyses on curated subsets and serve as a replication layer, while the questions students ask of the evidence matrix double as gap detectors for future inquiry. Integrity safeguards apply in both directions - teacher curation of retrieved sources, verification of agent-generated material against the original records, and rubrics that grade the quality of source criticism rather than the fluency of generated prose [16].

These designs, like the map they draw on, have borders that should be stated plainly rather than discovered in review. The corpus is bounded by one database, one language, and one search term, so its edges are retrieval edges as much as field edges; CNKI's topic semantics may over- or under-draw the population of relevant articles, and English-language scholarship lies outside scope by design. Because the dataset carries no citation counts or reference lists, frontier identification rests on declared approximations — recency of thematic mass, growth-rate bursts, and semantic drift — rather than on exact detection, and the final year is incomplete. Most consequentially, every interpretive category was produced by a single agent: the two-pass agreement figures evidence consistency within one model, not independence across coders, and same-session memory effects cannot be excluded, while topic borders remain legitimately fuzzy at the margins. These limits counsel a concrete agenda rather than a general caution: extending the map to multilingual corpora, validating agent-coded categories against independent human panels, tracking longitudinally whether discourse-coupled growth persists as the political vocabulary stabilizes, and testing the two classroom designs above in real courses, where the loop between research and teaching can finally be observed rather than inferred.

6. Conclusion

This review set out to map a field that had never been mapped and to test whether an AI agent could do so under audit. Read conceptually, the map tells one story in three movements. Chinese core-journal scholarship organizes Western civilization around a discourse core: the West is invoked chiefly as the comparator against which Chinese modernization and the new form of human civilization are elaborated, while a thinner but evidence-rich periphery of historical, encounter, and literary-cultural scholarship performs the quieter work of showing that the West itself was cross-civilizationally made. The field's growth follows its vocabulary rather than its evidence: output accelerates as policy-linked concepts enter the corpus, and the share of work resting on primary sources falls as volume rises. Its argumentative culture remains overwhelmingly normative-hermeneutic, leaving the pedagogical translation of its findings nearly unwritten — though the same analysis shows how each identified gap maps onto a concrete affordance by which agents can mediate the literature into the classroom and the classroom's questions back into research.

For information science, the demonstration matters as much as the map. A national database without citation counts or reference lists — precisely the terrain where the standard toolchains stall — proved mappable by an agent-plus-code pipeline in which every quantitative value came from executed scripts and every interpretive value from

logged, batched, re-testable readings. The approach has borders, and they are stated here rather than buried: one database, one language, one search term, an incomplete final year, and interpretive categories still awaiting independent human validation. Within those borders, the implication stands. As agentic systems become ordinary instruments of scholarship, the question is no longer whether they will participate in synthesis, but whether their participation will be protocol-bound, auditable, and pedagogically consequential — and whether the fields they map will come to be known by their evidence as well as by their vocabulary.

References

- [1] Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature.
- [2] van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping.
- [3] Wang, X., et al. (2024). Autonomous LLM agents for scholarly synthesis.
- [4] Lewis, P., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks.
- [5] Tricco, A. C., et al. (2018). PRISMA-ScR: An extension of the PRISMA statement for scoping reviews.
- [6] Bradford, S. C. (1934). Sources of information on specific subjects.
- [7] Lotka, A. J. (1926). The frequency distribution of scientific productivity.
- [8] Price, D. J. de S. (1963). Little science, big science.
- [9] Callon, M., et al. (1983). Mapping the dynamics of science and technology: Sociological analysis of science policy.
- [10] Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis.
- [11] Donthu, N., et al. (2021). How to conduct a bibliometric analysis: An overview and guidelines.
- [12] Blondel, V. D., et al. (2008). Fast unfolding of communities in large networks.
- [13] Kleinberg, J. (2003). Bursty and hierarchical structure in streams.
- [14] Cobo, M. J., et al. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field.
- [15] Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure.
- [16] Kasneci, E., et al. (2023). Large language models in education: A systematic review.
- [17] Zhang, Y., & Zhou, L. (2026). Chinese modernization and the new form of human civilization.
- [18] Xu, M. (2026). Mutual learning among civilizations and Western modernity.
- [19] Han, Q., & Wang, Y. (2024). Civilizational discourse in Chinese social sciences.
- [20] Cao, S., & Ming, K. (2026). Critiques of Eurocentrism in classical scholarship.
- [21] Liu, H. (2025). Narrative analysis of Sino-Western civilizational interaction.
- [22] Yan, J. (2024). Greek origins of Western civilization.
- [23] Chen, L. (2026). Historiographic construction of the West as a research object.
- [24] Zhang, R. (2025). Medieval European institutions and civilizational formation.
- [25] Feng, X. (2025). Sino-Western encounters and modern Chinese civilizational consciousness.
- [26] Liu, Q. (2016). World history textbooks and humanistic education.
- [27] Yan, H. (2022). Teaching the origins of Western civilization.
- [28] Liu, S., & Wang, Z. (2023). The Great Books tradition and Western civilization education.
- [29] Qiang, Y. (2023). General education in the digital-intelligence era and civilizational crisis.
- [30] Page, M. J., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews.
- [31] Lockwood, C., et al. (2015). Systematic reviews of qualitative evidence: Methodology and implementation using JBI.
- [32] Campbell, M., et al. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guideline.

