Quantifying Cultural Evolution: A Computational Analysis of Intellectual History in Chinese and Latin Traditions

Anonymous ACL submission

Abstract

Understanding how intellectual traditions evolve across cultures remains a fundamental challenge in cultural analytics. We present a computational framework for quantifying cultural evolution by extracting idea graphs from classical texts using LLMs, detecting thought communities via graph algorithms, and discovering cross-temporal idea lineages through semantic similarity. Applying this framework to 243 Chinese and 636 Latin texts spanning over 2,500 years, we construct graphs containing approximately 360,000 entities and 390,000 relations. Our analysis reveals fundamentally different evolutionary dynamics: Chinese tradition exhibits asynchronous Politics-Religion dominance with cumulative inheritance, while Latin tradition shows synchronized dynamics with substitutional inheritance. We identify two critical transition periods where diversity trajectories diverge in opposite directions, and through structural analysis of thought communities and lineages, reveal why these patterns emerge and how they shaped divergent modern trajectories. Our quantitative findings align with historiographical narratives about Chinese continuity versus European transformation, providing measurable evidence for cultural analytics and cross-civilizational comparison.

1 Introduction

The comparative study of intellectual history across cultures has long been central to understanding human cultural and social development. From Lovejoy's foundational work tracing thoughts evolving from ancient Greece to the 18th century (Lovejoy, 1936) to contemporary comparative philosophy (Stalnaker, 2006), humanistic scholars have sought to trace how ideas and thoughts emerge, transform, and persist across texts and time. However, such analyses have traditionally relied on close reading of literature, limiting both scale and reproducibility.

Recent advances in natural language processing offer new possibilities for large-scale cultural analytics. Large language models can extract structured knowledge from unstructured texts (Wei et al., 2023), graph algorithms can detect community structures in knowledge networks (Traag et al., 2019), and multilingual embedding models enable broad semantic comparison (Chen et al., 2024). Yet applying these computational methods to classical cultural texts poses unique challenges: extracting abstract intellectual concepts from ancient texts requires domain-appropriate operationalization; constructing multi-level thought systems demands careful methodological design; and quantifying cultural evolution patterns and investigating mechanisms lacks established metrics.

Given this gap in computational analysis of classical cultures, we take Chinese and Latin (European) traditions, among the world's most enduring and influential, as our foundation to propose a computational framework for quantifying intellectual evolution. Our approach addresses three fundamental research questions:

RQ1: Dynamics and Diversity. How do ideas distribute and evolve within cultures, and do different cultures exhibit different evolutionary patterns and transitions?

RQ2: Structure and Context. What are the causes of these patterns and transitions? We analyze the internal intellectual composition and the external historical context.

RQ3: Persistence and Regularity. Are these evolutionary patterns related to the cultural persistence, and do different cultures have different persistence patterns and regularity?

To answer these questions, following humanistic scholars' intellectual history research approach, we extracted large-scale idea graphs from 243 Chinese classical texts (771 BCE–1927 CE) and 636 Latin texts (500 BCE–2000 CE) with LLMs. We then discover major thought communities with network algorithms and trace idea lineages across documents and time with a state-of-the-art multilingual embedding model and FAISS (Douze et al., 2025) semantic similarity computing, thereby constructing multi-level, three-dimensional intellectual history networks.

Our quantitative analysis of these networks yields several key findings. First, while both cultures' intellectual production centers on Politics and Religion, they exhibit fundamentally different dynamic patterns: Chinese tradition shows asynchrony dominance, while Latin tradition shows rising and declining in synchrony. Second, we identify two critical transition periods, 200–600 CE and 1400–1800 CE, where Chinese and Latin diversity trajectories diverge in opposite directions. Furthermore, intellectual community structure analysis reveals the root causes: Chinese culture is anchored by Confucianism as an umbrella spanning Politics, Ethics, Education, and other domains, with Buddhism and Daoism representing non-mainstream Religion thought; Latin culture fuses Christian theology with political authority as the long-dominant intellectual orthodoxy, creating synchronized Religion-Politics dynamics. Finally, idea lineage analysis shows Chinese lineages persist roughly twice as long as Latin ones (median 555 vs. 207 years), with survival analysis confirming cumulative inheritance in China versus substitutional inheritance in Latin tradition, and significant turnover occurring during the transition periods.

These findings supplement effective evidence for humanistic research. Historians have long noted Chinese civilization's emphasis on continuity, embodied in Confucian principles, versus European history's periodic ruptures and revivals (Eisenstadt, 1986). Additionally, our work provides quantitative explanations for understanding how Chinese and Latin cultures developed differently in the modern era, as the second transition period profoundly shaped their divergent trajectories into modernity.

Our main contributions are: (1) an effective and efficient methodology for extracting and analyzing intellectual evolution from classical literature, extensible to broader cultural analysis; (2) a large-scale quantitative comparison of Chinese and Latin intellectual traditions revealing distinct evolutionary and persistence dynamics and structural roots; and (3) empirical evidence providing measurable support for understanding culture development.

2 Related Work

2.1 Computational Approaches to Cultural Analysis

Computational cultural analysis has developed significantly over the past decade. Early approaches focused on text mining and topic modeling to trace thematic patterns in literary corpora (Jockers, 2013), but lacked macro-level structural views. Recently, constructing literature or author networks based on semantic similarity or "intertextuality" has enabled quantitative cultural structural analytics. Duan et al. (Duan et al., 2023) visualized tracing networks of ancient Chinese literature and summarized patterns of Chinese intellectual evolution. Becker et al. (Becker and Culotta, 2025) constructed citation networks for Western philosophical history, providing humanities scholars with a distant reading perspective.

However, these works directly construct citation or intertextual networks from texts without extracting ideas or modeling complex thought systems. In contrast, traditional scholars typically induce abstract ideas from texts, examine how related concepts form coherent systems (Stalnaker, 2006), and discover implicit transmission of thoughts (Henke, 2025). Inspired by these paradigms, we model Chinese and Latin culture through three-level intellectual networks, from the "unit-ideas" proposed by Lovejoy, to large idea interaction clusters, and finally to implicit idea lineages across time, achieving more authentic cultural computation.

2.2 Knowledge Graphs and Network Analysis

The emergence of LLMs has enabled more efficient structured representations of cultural knowledge. Recent work demonstrates that LLMs' powerful pretrained knowledge and generalization capabilities perform remarkably well on cultural texts: LLMs can effectively extract complex knowledge from cultural heritage documents (Schimmenti et al., 2025) and construct multi-layered knowledge graphs from historical texts (Zeng, 2024). However, these works stop at graph construction or small-scale network interpretation, lacking advanced graph algorithms to analyze structures and patterns.

Recent NLP research has combined graph algorithms such as community detection (Luo et al., 2025) or hyperedge discovery (Luo et al., 2025) with knowledge graphs, achieving more effective knowledge retrieval. But these methods are primarily applied to general retrieval or knowledge-intensive domains like medicine and finance.

Inspired by these approaches, after extracting knowledge graphs with LLMs, we apply community detection to discover intellectual clusters within cultures, and further use hyperedges to connect semantically similar ideas as "idea lineages" spanning multiple documents, thereby constructing three-dimensional thought graphs that enable quantitative cultural analysis.

2.3 Intertextuality and Semantic Similarity

Computational intertextuality detection has progressed through several generations. First-generation methods focus on n-gram matching; Tesserae (Coffee et al., 2012) identified massive lexical overlaps in Latin poetry for digital humanities studies, but this intertextuality is limited to form matching. Second-generation approaches introduced word embeddings; Yousef et al. (Yousef et al., 2021) developed Latin-specific Word2Vec models, while Burns et al. (Burns et al., 2021) applied LatinBERT for context-aware semantic similarity detection. But these models operate at word or phrase level.

Recent work developed language models for sentence-level similarity. Riemenschneider and Frank (Riemenschneider and Frank, 2023) created SPhilBERTa for classical languages, though evaluation remains limited. The emergence of multilingual embedding models marks new progress. Models like BGE-M3 (Chen et al., 2024), Multilingual-E5 (Wang et al., 2024), and Embedding Gemma (Lee et al., 2025) support 100+ languages with unified representations and powerful retrieval.

While these models excel at multilingual similarity computation, they have not been applied to large-scale cultural analytics. We applied BGE-M3, a state-of-the-art embedding model, to compute semantic similarity across hundreds of thousands of ideas in Chinese and Latin corpora, discovering idea lineages spanning centuries through hyperedge aggregation.

3 Methods

3.1 Data Collection and Preprocessing

To ensure accurate representation of intellectual traditions, we use well-established scholarly compilations recognized by the humanities. For Chinese, we use 243 texts from *Bibliography of Major Chinese Scholarly Works* (), an authoritative compilation organized by leading Chinese scholars across disciplines, representing core classical texts across diverse domains spanning from 771 BCE to 1927 CE. For Latin, we use 636 texts from The Latin Library, a world-leading Latin text database widely used by humanities scholars for Latin cultural and historical research (lat), covering 316 authors from 500 BCE to 2000 CE across diverse genres and periods.

For each document, we inferred composition dates combining web retrieval and Claude Opus 4.5, validated by domain experts. Preprocessing included removing TEI metadata and annotations, normalizing whitespace, converting Chinese to simplified characters with OpenCC (Lin, 2011), and standardizing Latin capitalization. We chunked texts into 2,000 token units with 200 token overlap using tiktoken (OpenAI, 2025b). The resulting corpora are comparable: 35,363 Chinese text units and 33,697 Latin text units.

3.2 Idea Graph Construction

We extract entities and relations using GPT-5.1 API (OpenAI, 2025a) with temperature 0.1. For entities, we define three types: *Idea* (intellectual concepts), *Person* (thinkers, historical figures), and *Work* (texts, treatises). We extract original names and source sentence indices, then generate normalized modern-language names and descriptions (simplified Chinese for Chinese, English for Latin) to enable similarity computing.

For Idea entities, we annotate domain labels

$$\mathcal{D} = \{\text{Ethics}, \text{Politics}, \text{Religion}, \text{Epistemology}, \\ \text{Metaphysics}, \text{Natural Science}, \text{Art}, \text{Education}, \text{Other}\}.$$

We prompt GPT-5.1-mini to establish undirected edges between entities when physical relations exist (e.g., Person proposes Idea, Person authors Work, Work has Idea) or logical relations hold (e.g., comparison, elaboration, causation between Ideas).

To reduce redundancy and ambiguity, we further prompt GPT-5.1 to disambiguate entities and merge based on normalized modern names. We randomly sampled 5% of chunks for expert validation, iteratively refining prompts over two rounds. This yields 186,771 Chinese entities and 195,105 relations; 174,828 Latin entities and 195,296 relations. Figure 2 shows idea domain distributions: both traditions are dominated by Politics (Chinese 33.1%, Latin 32.5%), with Latin showing a higher Religion proportion (23.3% vs. 14.9%).

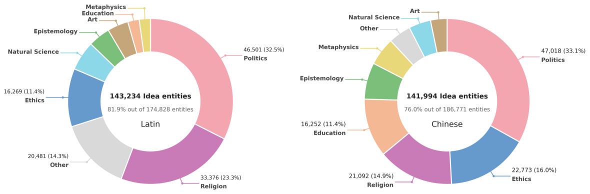

Figure 1: Idea distributions for Latin and Chinese.

3.3 Thought Community Detection

Shared entities across documents and explicit citations form cross-document links, potentially revealing larger-scale intellectual structures. We apply hierarchical Leiden community detection (Traag et al., 2019) with resolution=1.0, detecting 28,548 Chinese and 26,147 Latin communities forming 5-level hierarchies. Most communities are fine-grained (2–5 entities), but a small fraction cover 500+ entities (2% Chinese, 3% Latin) or span 20+ documents, reflecting macro-level intellectual communities.

3.4 Idea Lineage Discovery

Beyond explicit connections, we discover implicit idea lineages, i.e., semantic similarity between ideas across periods, where later ideas inherit earlier ideas. We embed each Idea entity's normalized name and description using BGE-M3:

$$\mathbf{v}_e = \text{Embed}([\texttt{domain}; \texttt{description}]) \quad (1)$$

We build FAISS indices for approximate nearest neighbor search with $k = 500$, retaining edges where $\text{sim}(\mathbf{v}_i, \mathbf{v}_j) \geq 0.75$. Following hypergraph aggregation (Feng et al., 2019), we aggregate pairwise edges into idea lineages, requiring 3–300 ideas per lineage. This yields 536 Chinese and 1,015 Latin idea lineages.

4 Experiments and Results

Our experiments address three research questions through progressive analysis: evolutionary dynamics and patterns, structural and contextual reasons, and inheritance mechanisms.

4.1 Intellectual Evolution Dynamics

We compute normalized intellectual idea density per domain-window pair, assigning documents to 400-year windows. The stream graphs (Figure 2) reveal contrasting evolutionary dynamics between the two traditions. Chinese thought shows *asynchronized* Politics-Religion dominance: Politics dominated through the Classical period, then declined sharply in Late Antiquity as Religion expanded, before reasserting dominance from High Medieval onward. Latin thought shows *synchronized* Politics-Religion dynamics: both domains grew together through Late Antiquity, jointly dominated for over a millennium, then declined together from mid-Early Modern as Education, Natural Science, Art, and Epistemology expanded.

These patterns suggest two potential critical transition periods (marked as vertical dashed lines in Figure 2): first half of Late Antiquity (200–400 CE) and first half of Early Modern (1400–1600 CE), where the two traditions diverge most sharply.

4.2 Intellectual Diversity Trajectories

To quantify these transitions, we compute normalized Shannon entropy over domain distributions and track its local regression slope (using 100-year units with ±2 window smoothing):

$$H_t^* = \frac{-\sum_d p_{t,d} \log_2 p_{t,d}}{\log_2 |\mathcal{D}|} \quad (2)$$

The diversity trajectory (Figure 3) sharpens these observations. In the first critical period (200–400 CE), Chinese diversity growth reaches its positive peak, reflecting the fragmentation of Han political unity and the subsequent flourishing of Buddhism and Daoist metaphysics. Latin diversity simultaneously reached its lowest point during the so-called "Dark Ages," as Christianity became the state religion and intellectual production consolidated around Religion-Politics. In the second critical period (1400–1600 CE), the trajectories reverse: Chinese diversity shifts to negative growth under Ming-Qing Neo-Confucian orthodoxy and autocratic consolidation, while Latin diversity surges as Renaissance and Enlightenment critiques opened space for previously minor domains. These opposite trajectories at both critical junctures align with traditional historical scholarship (Eisenstadt, 1986).

4.3 Thought Community Structure Analysis

We analyze how the composition of major thought communities accounts for the divergent evolutionary dynamics observed. Figure 4 presents the top 10 largest communities in Chinese and Latin traditions and their intellectual composition.

Chinese communities divide primarily into Confucian, Buddhist, and Daoist clusters (Figure X),

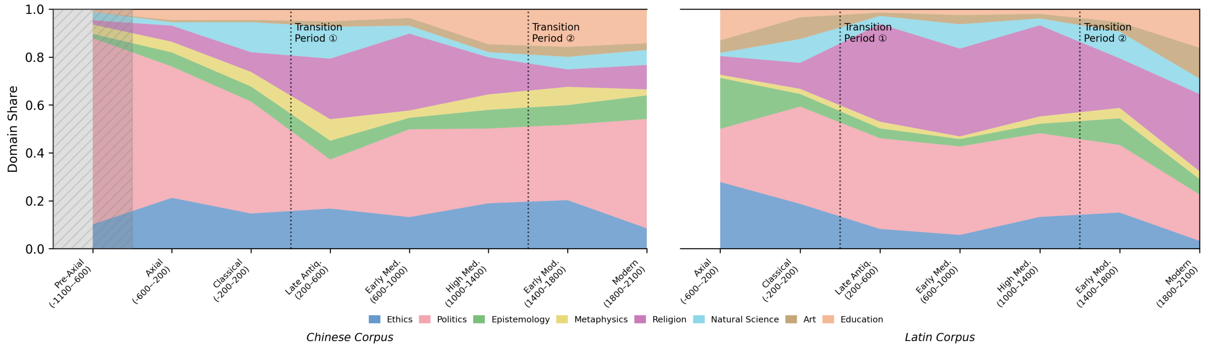

Figure 2: Chinese and Latin intellectual evolution across historical periods, aggregated in 400-year windows. Due to document distribution in corpus, Chinese texts are sparsely represented in the pre-axial period, whereas no Latin texts are present. Dashed lines indicate potential critical transition periods.

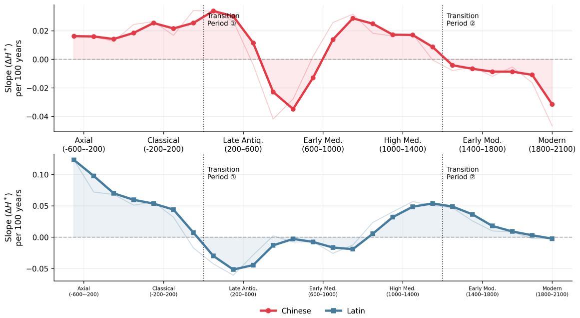

Figure 3: Intellectual diversity growth rate (ΔH^* per 100 years) in Chinese and Latin. H_t^* is normalized Shannon entropy over intellectual domain distributions in 100-year windows, smoothed with a ± 2-window moving average; slopes are estimated via local linear regression. Dotted lines mark the two hypothesized transition periods.

with Confucian communities (Cheng-Zhu Neo-Confucianism, Yangming School, Mencian Ethics, etc.) predominating. Crucially, Politics concentrates within Confucian communities, which function as umbrella frameworks spanning Ethics, Education, Epistemology, and Metaphysics, while Buddhist and Daoist communities show Religion dominance with minimal Politics. This separation means Politics and Religion represent *competing* intellectual streams: when Confucian political orthodoxy weakens during fragmentation periods, religious thought expands into the vacuum, producing the asynchronized dynamics observed above. Latin communities exhibit the opposite structure. Top communities (Classical Roman Jurisprudence, Early Christian Theology, Medieval Papal Canon Law, First Crusade Commanders, etc.) fuse Religion and Politics through institutional coupling, as Christianity served as state religion across Roman, Carolingian, and subsequent dynasties. This fusion means Religion and Politics rise and fall together, producing synchronized dynamics.

Normalized Shannon entropy over each community's domain distribution confirms this structural difference quantitatively (Figure X). In 14 of 20 community pairs, Chinese entropy substantially exceeds Latin. Confucian communities show the high entropy values, confirming their spanning of politics, ethics, education, and other domains, while Latin communities dominated by Religion-Politics fusion show lower entropy, reflecting their concentrated dual-domain structure.

4.4 Idea Lineage Persistence and Inheritance Mechanisms

Finally, we analyze idea lineage persistence to understand inheritance mechanisms. We define temporal span for lineage h as:

$$\text{Span}(h) = \max_{d \in \text{docs}(h)} \text{year}(d) - \min_{d \in \text{docs}(h)} \text{year}(d) \quad (3)$$

The lineage activity visualizations (Figures 5 and 6) reveal distinct temporal signatures: Chinese lineages are longer and flatter, presenting "long tails at both ends, flat plateau in middle"; Latin lineages are shorter and denser, concentrated in narrower temporal pulses from Classical Synthesis through Late Antiquity, becoming sparse by Early Modern. This contrast connects directly to our critical periods: Latin's sparse Early Modern lineages reflect how Renaissance and Enlightenment *critiqued and largely abandoned* previous Religion-Politics dominated thought, while Chinese lineages remain highly active because Neo-Confucianism *revived and developed* earlier Confucian political thought rather than replacing it.

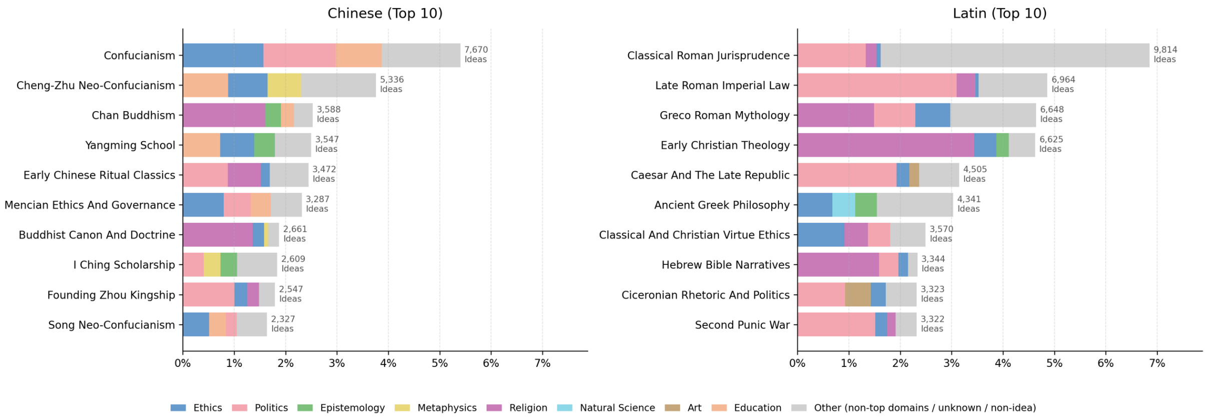

Figure 4: Top 10 largest thought communities in Chinese and Latin traditions and their intellectual composition. Community labels are generated using the GPT-5.1 API based on the top 100 weighted entities and validated by domain experts.

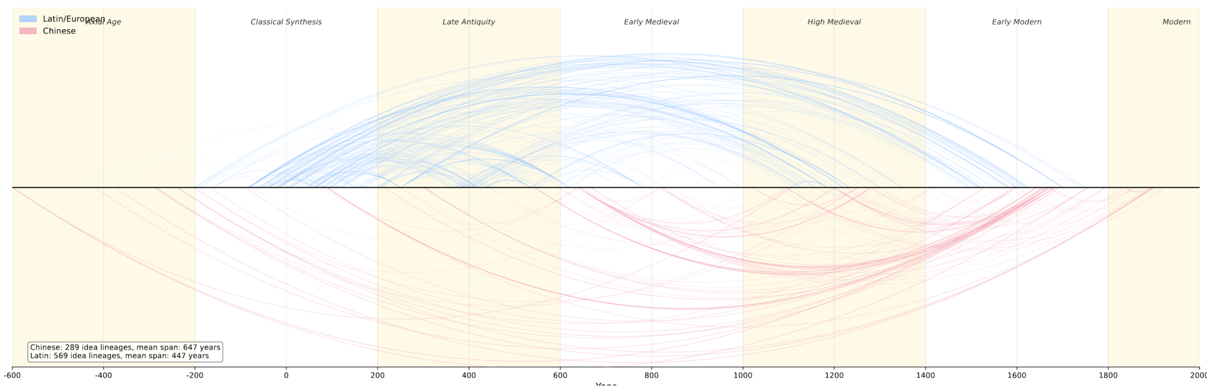

Figure 5: Ideational continuity across civilizations. Each curve represents an idea lineage connecting its earliest to latest documentary attestation. Latin/European lineages (top, blue) show concentrated temporal pulses peaking in Late Antiquity with rapid decline afterward. Chinese lineages (bottom, red) display longer spans with sustained activity across periods. Vertical bands indicate historical periods; the visualization reveals Chinese lineages' characteristic "long tails" versus Latin's temporal concentration.

Quantitative metrics confirm these patterns. The cumulative distribution function (Figure 8) shows Chinese median span of 555 years versus Latin's 207 years, with 53.6% of Chinese lineages spanning ≥500 years versus 28.8% for Latin.

We adopt Kaplan-Meier survival analysis (Kaplan and Meier, 1958), originally developed for biomedical research to estimate time-to-event probabilities under censoring. This method transfers naturally to intellectual history: just as clinical studies track patient survival over time, we track idea lineage "survival" (continued documentary attestation) across historical periods. Each lineage's lifespan—from first to last attestation—parallels a patient's time under observation, enabling rigorous comparison of persistence patterns across civilizations. The survival analysis (Figure 7, left panel) reveals Chinese lineages reach 50% survival at approximately 800 years (4 windows) versus 400 years (2 windows) for Latin; at 800 years, 34.4% of Chinese lineages remain active versus 20.8% for Latin.

We further compute turnover rate per 200-year window as:

$$\text{Turnover}_w = \frac{\text{Birth}_w + \text{Death}_w}{2 \cdot \text{Active}_w} \tag{4}$$

where Birth_w counts lineages first attested in window w, Death_w counts lineages last attested in window w, and Active_w counts all lineages with attestations spanning window w.

Latin shows consistently higher turnover (mean 0.33 vs. 0.29 for Chinese), indicating stronger generational replacement (Figure 7, right panel). Critically, turnover dynamics mirror diversity trajectories: Chinese turnover rises during the first critical period (200–400 CE) when diversity in-

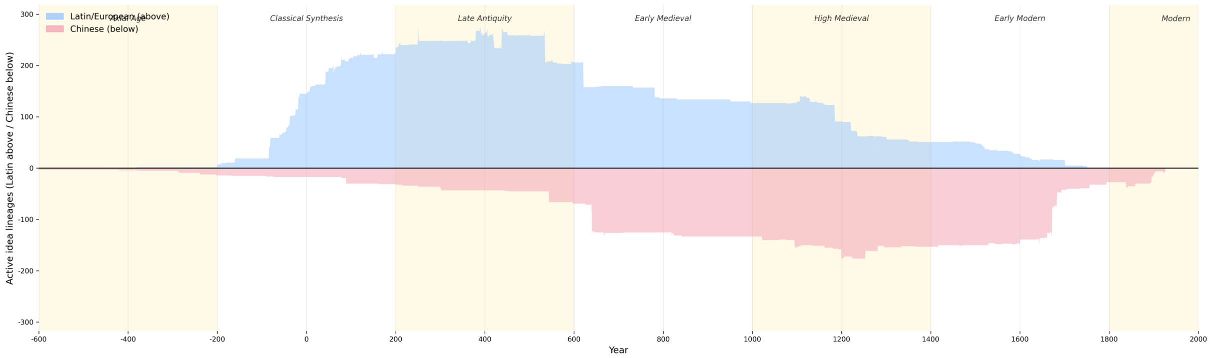

Figure 6: Mirrored visualization of active idea lineage density over time. Latin/European density (above axis) peaks during Late Antiquity (200–600 CE) then gradually declines. Chinese density (below axis, inverted for comparison) shows later growth peaking during High Medieval and Early Modern periods. The asymmetric temporal distributions reflect fundamentally different inheritance mechanisms: Latin lineages concentrate and terminate earlier, while Chinese lineages accumulate and persist longer.

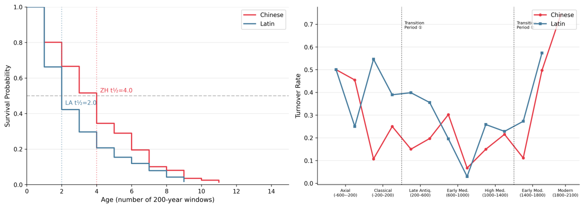

Figure 7: Left: Kaplan-Meier survival curves showing probability of idea lineage persistence across 200-year windows. Chinese lineages (red) reach 50% survival at approximately 4 windows (800 years) versus 2 windows (400 years) for Latin (blue). Right: Turnover rate per 200-year window, measuring the proportion of lineages born or terminated relative to active lineages. Latin shows consistently higher turnover, indicating stronger generational replacement dynamics. Dotted lines mark Transition Periods ① (200 CE) and ② (1400 CE).

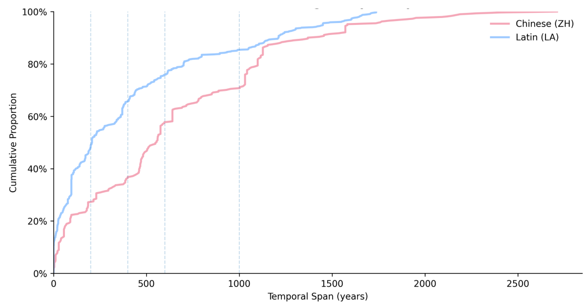

Figure 8: Cumulative distribution function (CDF) of idea lineage temporal spans. Chinese lineages (red) show substantially longer spans: median 555 years versus 207 years for Latin (blue). At the 500-year threshold, 53.6% of Chinese lineages persist versus only 28.8% of Latin lineages. The rightward shift of the Chinese curve indicates systematically longer idea transmission chains.

creased, while Latin turnover decreased; in the second critical period, the pattern reverses. These opposite trends provide quantitative evidence for traditional historiographical conclusions: Europe's Early Modern pluralism and fragmentation nurtured modernity through intellectual replacement (Mokyr, 2016; Jones, 1981), while China's return to unified political centralization preserved continuity, shaping profoundly different modern trajectories.

5 Conclusion

We presented a computational framework for quantifying intellectual evolution across civilizations by extracting large-scale idea graphs using LLMs, discovering thought communities through community detection, and tracing idea lineages through semantic embedding similarity. Our analysis of Chinese and Latin traditions reveals fundamentally different evolutionary dynamics: asynchronized versus synchronized Politics-Religion patterns. Community structure analysis traces these differences to distinct intellectual foundations. Confucianism as the secular ideology competing with Buddhist and Daoist religious thought in China, versus Christian-governance fusion in the Latin tradition. These structural differences produce opposite diversity trajectories at two critical transition periods (200–400 CE and 1400–1600 CE). Idea lineage analysis further reveals how these patterns manifest in inheritance mechanisms: Chinese tradition exhibits cumulative inheritance, with Neo-Confucianism reviving and developing earlier political thought during the second transition period; Latin tradition exhibits substitutional inheritance, as Renaissance and Enlightenment critiques largely abandoned previous Religion-Politics orthodoxy. These contrasting inheritance patterns ultimately shaped the divergent modern trajectories of the two civilizations.

Methodologically, we demonstrate that state-of-the-art NLP methods combined with network analysis and quantitative metrics can effectively support complex cultural analysis. Substantively, our quantitative findings provide measurable evidence for understanding when and how different cultures evolve, aligning with historiographical narratives about Chinese continuity versus European transformation. Future work should address corpus representativeness, develop deeper qualitative integration with humanities scholars, and extend to additional cultures (Arabic, Sanskrit, Nordic, etc.) to discover broader patterns in cultural evolution.

Limitations

Cultural computation attempts to model culture—a highly abstract and complex phenomenon—using quantifiable computational methods, seeking to discover fundamental patterns and explanatory mechanisms. However, the inherent abstraction, complexity, and even subjectivity of the concept "culture" itself pose significant challenges for quantitative analysis, particularly when examining ancient cultures across millennia of history. We have endeavored to employ efficient, reasonable, and expert-validated methods to construct idea graphs and conduct quantitative analysis. Our findings both validate and complement humanistic scholarship, revealing macro-level patterns that are difficult to discern through close reading alone. Nevertheless, several important limitations constrain the interpretation of our findings.

Corpus temporal and compositional imbalances. Our corpora, while canonical, represent elite intellectual production and may not capture folk or oral traditions. The Chinese corpus emphasizes Confucian and Buddhist texts; the Latin corpus over-represents legal and theological documents. More fundamentally, the temporal distributions of Chinese and Latin corpora are not strictly equivalent. European intellectual history witnessed multiple linguistic transitions—from Ancient Greek to Latin to vernacular national languages—meaning Latin gradually fell out of use during the early modern period, resulting in sparser data for this era. Additionally, the Latin Library contains limited material from the Axial Age and earlier periods, notably lacking works by figures such as Socrates (whose ideas survive only through later authors). These imbalances mean our data are not uniformly distributed across historical periods. This raises a methodologically tricky question: temporal sparsity may genuinely reflect periods of cultural discontinuity caused by warfare, natural disasters, or epidemics—in which case uneven data distribution legitimately mirrors uneven cultural evolution. Alternatively, gaps may simply result from editorial decisions in the original corpus compilation, driven by subjective preferences or practical constraints unknown to us. We cannot definitively distinguish between these possibilities.

Composition date uncertainty. Composition dates for ancient texts often span centuries or remain disputed among scholars. We use the latest estimated dates validated by domain experts, but this introduces systematic uncertainty, particularly for early periods with sparse documentation. Different dating choices could potentially affect our temporal analyses.

LLM extraction and potential biases. We employed GPT-5.1 to extract large-scale idea graphs from both Chinese and Latin corpora. While this represents one of the most advanced LLMs currently available, supporting over 100 languages, including Chinese and Latin, and while these historical texts likely appear in the model's pretraining data, generative AI methods inevitably carry risks. Despite our efforts at expert sampling validation and iterative prompt refinement to normalize the graphs, this automated approach may introduce undiscovered errors or reflect inherent model biases that we cannot fully characterize or eliminate.

Semantic similarity limitations. BGE-M3 embeddings capture modern-language descriptions rather than original classical texts. While this design enables cross-lingual comparison, it may miss nuances and connotations apparent only in original formulations. The 0.75 cosine similarity threshold for lineage construction, while empirically motivated through preliminary experiments, remains somewhat arbitrary. Different thresholds would yield different lineage structures and potentially different conclusions.

Limited cultural representativeness. We aim to reveal quantifiable patterns and mechanisms of cultural evolution, yet our analysis encompasses only two cultural traditions. This limits the generalizability of our findings. Including additional cultures, such as Arabic, Indian, Ancient Greek, Nordic, African, Egyptian and other traiditons, would substantially strengthen claims about universal versus culture-specific patterns of intellectual inheritance. Our current binary comparison, while revealing, cannot distinguish whether observed differences reflect fundamental civilizational divergences or merely idiosyncratic features of these two particular traditions.

Need for deeper humanistic integration. Our analysis is fundamentally computational and descriptive. While we identify correlations between historical events and quantitative patterns, and while our findings align with established historiographical narratives, we cannot establish causal relationships between political, religious, or social developments and intellectual dynamics. Richer integration with qualitative humanistic scholarship—including detailed case studies, expert interpretation of specific communities and lineages, and engagement with debates in comparative philosophy and intellectual history—would strengthen both the validity and the interpretive depth of our quantitative findings.

Broader Impact and Ethical Considerations

We believe this work extends beyond methodological contributions to computational humanities. By demonstrating that state-of-the-art NLP tools can effectively extract and analyze intellectual content from classical multilingual texts, we open possibilities for large-scale cultural analytics that were previously infeasible. Our framework could enable humanities scholars to identify cross-cultural patterns, trace intellectual influences across civilizations, and test long-standing historiographical hypotheses with quantitative evidence. The extracted idea graphs and lineages themselves represent a structured knowledge resource that could support downstream applications in education, cultural heritage preservation, and cross-cultural dialogue.

More broadly, as large language models increasingly shape how knowledge is organized and retrieved, understanding the intellectual traditions embedded in their training data becomes crucial. Classical Chinese and Latin texts likely constitute significant portions of LLM pretraining corpora, meaning the cumulative versus substitutional inheritance patterns we identify may subtly influence how these models represent and reason about Eastern versus Western thought. Our quantitative characterization of these traditions could inform more culturally-aware model development and evaluation.

Regarding ethical considerations, we acknowledge several concerns inherent to computational cultural analysis. First, quantitative comparison of intellectual traditions carries risks of oversimplification or implicit value judgments. We emphasize that our findings—such as longer persistence of Chinese idea lineages—should not be interpreted as indicating cultural superiority. Both cumulative and substitutional inheritance mechanisms have produced profound intellectual achievements; they represent different, equally valid modes of cultural transmission shaped by distinct historical circumstances.

Second, our framework necessarily imposes modern categorical schemes onto historical materials. The domain labels (Ethics, Politics, Religion, etc.) and entity types reflect contemporary scholarly conventions that may not align with how historical actors understood their own intellectual work. Ancient Chinese thinkers did not conceptualize their ideas as belonging to discrete domains," nor did medieval Latin authors necessarily distinguish Religion" from "Politics" as we do today. While such categorization enables systematic analysis, readers should recognize these as analytical constructs rather than ontological truths about how ideas were historically organized.

Third, computational methods risk being perceived as providing objective" or definitive" answers to inherently interpretive questions. Cultural analysis involves hermeneutic complexity that no algorithm can fully capture. Our quantitative findings should be understood as one lens among many—complementing rather than replacing close reading, contextual interpretation, and humanistic judgment. We encourage scholars to engage critically with our methods and findings rather than treating them as computational ground truth.

Finally, we commit to releasing our code, prompts, and foundation corpora to enable verification, critique, and extension of this work, while respecting copyright constraints on underlying corpora. We hope this transparency facilitates productive dialogue between computational and humanistic approaches to cultural analysis.

References

The latin library. https://www.thelatinlibrary.com/.

Ryan Becker and Aron Culotta. 2025. A computational analysis of in-text reference networks across philosophical texts. *arXiv preprint arXiv:2504.20065*.

Patrick J. Burns and 1 others. 2021. Latinbert: A contextualized latin embedding. In *Proceedings of NAACL-HLT*, pages 4900–4907.

Jianlv Chen and 1 others. 2024. Bge m3-embedding: Multi-lingual, multi-functionality, multi-granularity text embeddings. *arXiv preprint arXiv:2402.03216*.

Neil Coffee and 1 others. 2012. Intertextuality in the digital age. *Transactions of the American Philological Association*, 142(2):383–422.

Matthijs Douze, Alexandr Guzhva, Chengqi Deng, Jeff Johnson, Gergely Szilvasy, Pierre-Emmanuel Mazaré, Maria Lomeli, Lucas Hosseini, and Hervé Jégou. 2025. The faiss library. *Preprint*, arXiv:2401.08281.

Siyu Duan and 1 others. 2023. Disentangling the cultural evolution of ancient china. *Humanities and Social Sciences Communications*, 10:503.

Shmuel N. Eisenstadt. 1986. *The Origins and Diversity of Axial Age Civilizations*. SUNY Press.

Yifan Feng and 1 others. 2019. Hypergraph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 3558–3565.

D. Henke. 2025. Intertextual analysis as a tool for memory transformation in literature. *Memory Studies*, 17(5):1234–1252.

Matthew L. Jockers. 2013. *Macroanalysis: Digital Methods and Literary History*. University of Illinois Press.

Eric L. Jones. 1981. *The European Miracle: Environments, Economies and Geopolitics in the History of Europe and Asia*, 3rd ed., 2003 edition. Cambridge University Press, Cambridge.

Edward L Kaplan and Paul Meier. 1958. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481.

J. Lee and 1 others. 2025. Gemma embedding: Generalizable embeddings from gemma 2. *arXiv preprint arXiv:2509.20354*.

Victor Lin. 2011. opencc-python: A python wrapper for open chinese converter. `https://pypi.org/project/opencc-python/`.

Arthur O. Lovejoy. 1936. *The Great Chain of Being*. Harvard University Press.

Hong Luo and 1 others. 2025. Hypergraphrag: Retrieval-augmented generation via hypergraph-structured knowledge representation. *arXiv preprint arXiv:2503.21322*.

Joel Mokyr. 2016. *A Culture of Growth: The Origins of the Modern Economy*. Princeton University Press, Princeton, NJ.

OpenAI. 2025a. GPT-5.1 model documentation. `https://platform.openai.com/docs/models/gpt-5.1`.

OpenAI. 2025b. tiktoken: Fast BPE tokeniser for use with OpenAI models. `https://github.com/openai/tiktoken`.

Florian Riemenschneider and Anette Frank. 2023. Detecting latin allusions to ancient greek literature. *arXiv preprint arXiv:2308.12008*.

Andrea Schimmenti, Valentina Pasqual, Fabio Vitali, and Marieke van Erp. 2025. Knowledge graphs generation from cultural heritage texts: Combining llms and ontological engineering for scholarly debates. *Preprint*, arXiv:2511.10354.

Aaron Stalnaker. 2006. *Overcoming Our Evil: Human Nature and Spiritual Exercises in Xunzi and Augustine*. Georgetown University Press.

Vincent A. Traag, Ludo Waltman, and Nees Jan van Eck. 2019. From louvain to leiden: Guaranteeing well-connected communities. *Scientific Reports*, 9:5233.

Liang Wang and 1 others. 2024. Multilingual e5 text embeddings: A technical report. *arXiv preprint arXiv:2402.05672*.

Jason Wei and 1 others. 2023. Zero-shot information extraction via chatting with ChatGPT. *arXiv preprint arXiv:2302.10205*.

Tariq Yousef and 1 others. 2021. Profiling of intertextuality in latin literature using word embeddings. In *Proceedings of the LaTeCH-CLfL Workshop*, pages 43–52.

Yifan Zeng. 2024. Histolens: An llm-powered framework for multi-layered analysis of historical texts – a case application of yantie lun. *Preprint*, arXiv:2411.09978.

Zhongguo Xueshu Mingzhu Tiyao Bianweihui (), editor. 2019. *Zhongguo Xueshu Mingzhu Tiyao (He Ding Ben) [Bibliography of Major Chinese Scholarly Works (Collected Edition);]*. Fudan University Press, Shanghai, China.