The Information Science Analysis of Historical Geography Documents from the Perspective of Digital Humanities

Bingxuan Wang
Henan Normal University, Xinxiang, Henan, China

Abstract: This study explores the application of information science methodologies for systematic analysis of historical geographical documents within the framework of digital humanities. The research aims to reveal the multidimensional informational attributes of these documents and establish an interdisciplinary analytical framework integrating text mining, spatiotemporal analysis, and knowledge organization techniques. Through the integration of theoretical and methodological approaches, the study demonstrates that the value of historical geographical documents can be deeply excavated and restructured, transforming their knowledge form from static texts into dynamic knowledge systems. Ultimately, this analytical pathway provides new methodological support for academic research and opens feasible practical directions for the transformation of cultural resources and innovation in knowledge services.

Keywords: Digital Humanities; Historical Geography; Information Science; Knowledge Service; Analytical Paradigm

1. Introduction

In the current era of deep integration between digitalization and humanities research, digital humanities has not only revolutionized traditional historiographical paradigms but also pioneered new approaches for organizing, interpreting, and knowledge discovery in historical geographical documents. As crucial carriers of spatiotemporal information and records of human activities, these documents inherently form a complex information system. This study aims to systematically analyze historical geographical documents by integrating information science theories with digital humanities methodologies. Through examining their informational attributes, analytical frameworks, and knowledge transformation mechanisms, we seek to establish an interdisciplinary analytical framework. This framework will deepen our understanding of the intrinsic informational value of these documents while promoting their revitalization and innovative knowledge services in the digital environment.

2. The Intelligence Attribute and Value Reconstruction of Historical Geography Documents

2.1 Analysis of the Space-Time-Event Dimensions as a Composite Intelligence Source

Historical geographical documents are not mere textual records, but complex systems containing multidimensional intelligence. Their core characteristics lie in the intricate interplay and mutual construction of three dimensions: space, time, and events. Spatially, these documents establish concrete geographical frameworks and spatial networks through descriptions of place names, directions, territories, distances, and even topographical features-information that inherently reflects environmental cognition and territorial order. Temporally, chronological records, historical evolution, and administrative changes place spatial frameworks within dynamic historical transformations, endowing geographical information with temporal sequence and process. Eventually, whether military campaigns, economic activities, population migrations, or cultural dissemination, these serve as driving forces and concrete content that reshape spatiotemporal patterns, filling the framework of time and space to become a stage for human activities. These three dimensions do not exist in isolation but are mutually locked and interpretive. The evolution of a place name often correlates with pivotal moments of regime changes and social events of conquest and integration; the rise and fall of a canal intertwines with temporal sequences of natural environmental changes, spatial practices

of engineering technology, and major economic-political decisions. It is this multidimensional complexity that makes historical geographical documents a unique intelligence source. Their value lies not only in individual dimensions of information but also in the deep logic and holistic picture revealed through interdimensional interactions, requiring intelligence analysis to adopt a relational and systematic perspective for deconstruction [1].

2.2 The Manifestation and Expansion of the Value of Literature and Information in the Digital Environment

In traditional research paradigms, the composite intelligence embedded in historical geographical documents often remains latent. Constrained by manual reading and linear retrieval methods, their inherent connections and macro patterns are difficult to fully capture. The digital environment, however, has significantly revealed and expanded these documents' inherent intelligence value through technological intermediaries. Text digitization and tagging technologies transform unstructured natural language texts into computable, searchable structured or semi-structured data, enabling batch identification and extraction of key elements like people, places, times, and events from vast document collections. Building on this foundation, Geographic Information System (GIS) technology provides precise positioning, visualization, and spatial analysis platforms for geographic intelligence. It converts abstract textual descriptions into intuitive map layers and supports operations like distance calculation, buffer zone analysis, and overlay comparison, thereby revealing hidden spatial distribution patterns and relational networks within documents. More crucially, linked data and knowledge graph technologies can semantically connect and integrate these discrete entities and relationships, constructing machine-readable, reasoning-capable knowledge networks. This not only breaks the boundaries of individual documents and achieves cross-document, cross-type intelligence integration, but also enables the discovery of implicit connections through reasoning-such as automatically identifying different historical names for the same location or aggregating scattered records of the same event. Thus, the digital environment does not generate new intelligence. Instead, it activates and amplifies the latent, fragmented, and hard-to-access deep intelligence value in literature by providing powerful tools for processing, correlation, and presentation. This transforms what scholars once considered a treasure trove-accessible only through painstaking study-into a dynamic knowledge resource that can be systematically mined and interactively explored [2].

2.3 Attribute Transformation from Static Classics to Dynamic Knowledge Body

Through the combined influence of digital humanities and information science methodologies, historical geographical documents are undergoing a profound transformation from closed, static canonical texts to open, dynamic knowledge systems. Traditional documents, once finalized in their physical form, were characterized by fixed, authoritative knowledge where readers primarily engaged in passive reception and interpretation. However, when digitized and analyzed within computational frameworks, these static boundaries begin to dissolve. Document content is deconstructed into granular metadata elements-individuals, locations, timelines, events, and objects-which transcend their original linear textual form to become independently identifiable, interconnected, and reconfigurable knowledge units. This evolution transforms documents from mere "readings" into "data sources" and "knowledge repositories." Consequently, knowledge systems dynamically expand: new research findings can be seamlessly integrated as contextual data to refine existing knowledge nodes; diverse analytical models and perspectives can reorganize and visualize datasets to generate varied knowledge products; even public participation contributes fresh insights through annotations and collaborative tagging. Document knowledge is no longer a fixed endpoint but a continuously editable, supplemented, verified, and expanded starting point. The core of this transformation lies in the significant enhancement of intelligence's "operability" and "reproducibility," enabling historical geographical knowledge to transcend the constraints of single physical carriers. It facilitates the continuous flow, evolution, and value-added of such knowledge within digital spaces, thereby better meeting contemporary demands for interdisciplinary research, cultural heritage innovation, and public knowledge services [3].

3. Transformation of Information Analysis Paradigm Driven by Digital Humanities Technology

3.1 Content Depth Revealing Based on Text Mining and Geographic Information System

The integration of digital humanities technology has fundamentally transformed our approach to historical geographical documentation through methodological innovation. The breakthrough lies in creatively combining text mining with Geographic Information Systems (GIS). While traditional close reading remains meticulous, it struggles to systematically extract macro patterns and hidden connections from voluminous archives. Text mining technology serves as the initial filter and structuralizer, automatically extracting key historical figures, geographical names, temporal markers, and core event sequences from unstructured textual data through techniques like named entity recognition, topic modeling, and co-occurrence analysis. However, these data remain discrete information points. The introduction of GIS infuses spatial vitality into these points, enabling researchers to precisely map extracted geographical names and transform abstract textual descriptions into visual representations. More importantly, GIS's spatial analysis capabilities-such as buffer zone analysis, overlay analysis, path analysis, and spatial statistics-reveal patterns invisible to text-based studies. These include macro trends in population migration, spatiotemporal trajectories of military operations, patterns of administrative boundary changes, and spatial clustering characteristics of economic activities. This integration is not merely a technical overlay, but establishes a transformation pathway from "textual semantics" to "spatial semantics." It elevates the geographical intelligence embedded in documents from obscure contextual narratives to quantifiable, analyzable, and verifiable research subjects, thereby achieving a profound revelation and interpretation of document content that penetrates from surface to essence and from isolated points to comprehensive patterns [4].

3.2 Relationship Intelligence Discovery Based on Social Network Analysis and Association Data Technology

The paradigm shift in intelligence analysis transcends extraction of entity and attribute information, with its profound transformation now centered on discovering and reconstructing "relationships"-a transformation enabled by the mature application of social network analysis and relational data technologies. The world documented in historical geographical literature is essentially a complex network of nodes (individuals, locations, institutions, events) interconnected through various relationships. While traditional research can detect some explicit connections, it often struggles with large-scale, implicit network structures. Social network analysis provides a robust framework for this purpose. It treats entities extracted from documents as network nodes, quantifies their co-occurrence, dependency, interaction, and influence as edges, and employs centrality analysis, community detection, and path exploration algorithms to visually reveal power cores, key information hubs, community boundaries, and the strategic positions of individuals or locations within historical contexts. This approach enables quantitative representation of historical relationships' "influence," "proximity," and "structural holes." Meanwhile, relational data technology lays the foundation for machine-readable and cross-domain interconnection of relational intelligence at the semantic level. By adopting a unified resource description framework, it assigns globally unique URIs to each entity and defines relationship attributes in the "subject-predicate-object" triplet format. This not only formalizes internal document relationships but, more crucially, facilitates cross-database linking and integration of related entities from diverse sources and data types, thereby constructing a grand knowledge graph that transcends document boundaries. This atlas automatically links a historical figure's activities to all associated locations and events, while integrating evidence chains from diverse sources like local chronicles, memorials, and literary collections to trace the evolution of place names. This transformation elevates relational context from background clues to core analytical objects, driving research to shift from entity-centric to relationship-centric approaches [5].

3.3 Empowerment of the Analysis Process by Visual Narrative and Spatiotemporal Modeling

The paradigm-shifting impact of digital humanities technologies extends beyond

breakthroughs in analytical tools, fundamentally transforming research cognition through visual storytelling and spatiotemporal modeling. These powerful tools create intuitive interfaces between scholars and complex data ecosystems. Visual narratives transcend static charts by transforming spatiotemporal patterns, relational networks, and event sequences into interactive "stories" through dynamic maps, timeline animations, and interactive network diagrams. This process itself becomes an essential analytical component rather than just presenting results. Through interactive exploration, researchers can identify anomalies or hidden patterns by adjusting parameters, filtering layers, and tracing paths, achieving "thinking through visuals." Spatiotemporal modeling takes this further by abstracting and reconstructing historical geographical processes. Whether simulating human migration through agent-based models or reconstructing ancient urban landscapes with 3D technologies, the goal isn't absolute historical accuracy but creating a controllable "digital sandbox" for experimentation. Within this sandbox, researchers test hypotheses and observe macro-level phenomena under specific rules, deepening understanding of the underlying dynamics driving historical geographical processes. Thus, visualization and spatiotemporal modeling transform intelligence analysis from a linear, descriptive process into a cyclical, exploratory, and even experimental cognitive practice, significantly enhancing researchers' ability to interpret complex spatiotemporal information [6].

4. Intelligence Transformation and Practical Path for Knowledge Service

4.1 The Logic from Intelligence Extraction to Structured Knowledge Base Construction

The transition from intelligence extraction to constructing structured knowledge bases signifies a new phase of systematic and sustainable value mining in historical geographical literature. This process begins with systematically cleaning, correlating, semantizing, and persistently storing fragmented metadata extracted through digital humanities technologies. Its core mission is to transform scattered "data" into organized "knowledge" rich in semantic connections, forming a digital knowledge infrastructure for continuous utilization and expansion. This is not mere information accumulation but a systematic endeavor adhering to rigorous knowledge engineering principles. The construction logic first requires establishing a domain ontology or core data model capable of accommodating complex spatiotemporal and semantic relationships, clarifying definitions and hierarchies of entities and their connections within the knowledge base to ensure internal logical consistency. Subsequently, extracted entities and relationships are standardized and deeply correlated according to this model. For instance, a historical figure is not only linked to life events but also requires machine-readable connections with nodes such as ancient and modern place names involved in their activities, corresponding official systems, and relevant literature sources. The entire process involves continuous quality control and knowledge integration to resolve issues like homonymy and spatiotemporal contradictions across different documents. The fully developed knowledge base essentially constitutes a dynamic, scalable digital twin with rich semantic connections. It transforms narratives originally preserved in traditional bound books into a computable and traversable knowledge network, thereby establishing a robust data foundation for advanced knowledge discovery and services that transcends the physical form of original documents [7].

4.2 Knowledge Organization and Intelligent Retrieval Strategies for Supporting In-depth Research

To effectively support in-depth academic research, structured knowledge bases require the design and implementation of knowledge organization and intelligent retrieval strategies that transcend traditional keyword matching. Modern knowledge systems have evolved beyond simple cataloging and classification, aiming to construct semantic networks that reveal multidimensional connections between concepts. The key lies in utilizing technologies like ontologies and knowledge graphs to explicitly express complex semantic relationships-including geographical evolution, official career transitions, causal event chains, and familial connections. These relationships function like the gridlines of a knowledge network, weaving discrete entities into an organic whole rich with contextual context. Only

through such deeply organized knowledge structures can intelligent retrieval strategies be effectively implemented. At this stage, searching no longer resembles submitting isolated keywords to a static database, but rather resembles engaging in a dialogue with domain experts. The system can support complex relational queries such as "Identify all literati who served as Hangzhou Prefects during the Song Dynasty and trace their peer relationships." Furthermore, it provides semantic extensions and recommendations for entities. When users search for a specific canal, the system can proactively link to other engineering projects by its excavators, related water systems within the same hydrological network, or historical records of economic and demographic changes along the canal. The core of this retrieval model lies in the shift from "finding the known" to "exploring the unknown." Through path discovery, associative reasoning, and visualization, it guides researchers to uncover previously unanticipated knowledge connections, stimulates new problem awareness, and transforms the knowledge base from a passive information repository into an intelligent research environment that actively supports problem-driven studies, facilitates complex analysis, and assists hypothesis verification [8].

4.3 Innovative Perspectives of Public Cultural Services Based on Intelligence Analysis

The application of intelligence analysis from historical geographical literature to public cultural services not only expands the boundaries of academic research but also opens innovative perspectives for revitalizing cultural heritage and public historical education. The core of this transformation lies in creatively adapting and redesigning academic achievements that have undergone in-depth excavation, structural reorganization, and visual interpretation. These adaptations align with the public's knowledge background, cognitive habits, and emotional needs, transforming them from purely academic contexts into experiential, participatory, and communicable cultural products and services. This innovative approach manifests in multiple dimensions: In content presentation, it moves away from didactic instruction by utilizing visual tools like narrative maps, spatiotemporal storylines, and dynamic yearbooks to transform complex historical geographical changes into intuitive narratives or explorable spatiotemporal scenarios, immersing the public. In interactive experience, it encourages participatory cultural creation-allowing users to overlay past-present comparisons on maps, annotate family migration routes, or digitally reconstruct historical scenes through crowdsourcing, shifting services from one-way information transmission to two-way cultural co-creation. In communication channels, it deeply integrates online and offline platforms, serving as the core of museum interactive exhibitions and urban historical geographic information systems, while embedding regional historical culture into daily spaces and life trajectories through mobile apps and social media mini-programs. Thus, geographical place names, city rise and fall, and trade route changes once buried in ancient texts are no longer exclusive knowledge of scholars but become vivid materials for telling local stories, shaping cultural identities, and enhancing historical spatial recognition. This perspective finally makes the academic value of intelligence analysis achieve social spillover, and builds a two-way nourishing bridge between academic research and public culture.

5. Conclusion

In conclusion, the core of conducting information science analysis on historical geographical documents through the lens of digital humanities lies in employing interdisciplinary methodologies to systematically deconstruct and reconstruct their inherent multidimensional information elements. This process not only achieves in-depth exploration and reconstruction of document intelligence value, but also fosters new analytical paradigms, facilitating the transformation from original documents to systematic knowledge services. Future research in this interdisciplinary field should focus more on refining and integrating methodologies, strengthening theoretical frameworks, and actively exploring its broad application prospects in enhancing academic research efficiency, empowering cultural heritage revitalization, and supporting macro-level decision-making. This will transform the spatiotemporal information buried in historical archives into truly usable dynamic knowledge resources.

References

[1] Gong Jiaoteng, Zhang Kang, Wang Fengjiao.

The Evolution of Methodological Paradigms in Library and Information Science and Its Academic Impact: An Empirical Analysis Based on 18 CSSCI-Source Journals in the Field [J]. Library and Information Work, 2025,69(21):52-69.

[2] Ji Xiaoling, Zhang Mingping, Liu Guifeng, et al. Evolution and divergence analysis of research themes in library and information science in China based on STM[J]. Library Science Research, 2025, (05):2-15+52.

[3] Xie Linhong. Analysis of Information Resource Integration in Digital Information Environment: Recommended Reading from "Library and Information Science Theory Research" [J]. Information Theory and Practice, 2025,48(01):216.

[4] Huang Yun. Analysis of the Synergistic Development of Library Information Science Education and Library Information Careers [J]. China Management Informatization, 2024,27(11):156-159.

[5]Gao N, Zhou Q. Identification and Evolution Trend Analysis of Frontier Fields in Information Science Based on Co-citation Method[J]. Modern Intelligence, 2024,44(05):3-19.

[6] Dai Hongbo. Application of Baidu Encyclopedia in Academic Research Based on Citation Analysis: A Case Study in Library and Information Science [J]. Science and Technology Communication, 2024,16(02):37-39.

[7] Xie Mengqing. Research on the Development and Utilization of the Yunnan Series from the Perspective of Information Science [D]. Yunnan University, 2019.

[8] Guo Huaxiang. Cross-disciplinary bibliometric analysis of library and information science in China [D]. Zhengzhou University, 2018.

