

Article

Tradition Meets Innovation: Research Trends in Libraries, Archives, and Museums Convergence From a Digital Humanities PerspectiveGuan Wang¹, Yanti Idaya Aspura Mohd Khalid^{1,*}, Noorhidawati Abdullah¹,
Yi-Shuai Xu¹, Ye Yuan¹¹Department of Library and Information Science, Faculty of Arts and Social Sciences, Universiti Malaya, 50603 Kuala Lumpur, Malaysia*Correspondence: yanti@um.edu.my (Yanti Idaya Aspura Mohd Khalid)

Academic Editor: Natália Tognoli

Submitted: 3 April 2025 Revised: 2 June 2025 Accepted: 13 June 2025 Published: 19 August 2025



Guan Wang is currently a Ph.D. candidate in the Department of Library and Information Science at the Faculty of Arts and Social Sciences, Universiti Malaya. She obtained her Master's degree in Library and Information Science from Zhengzhou University, P.R. China. Her research interests primarily include data librarianship, archive management, data visualization, and digital humanities. Her work has been published in international and national core journals, such as Journal of Academic Librarianship, Electronic library, and Archives Management.



Yanti Idaya Aspura Mohd Khalid is an associate professor in the Department of Library and Information Science and the deputy dean of the Faculty of Arts and Social Sciences at Universiti Malaya. Dr. Yanti Idaya obtained her Ph.D. in 2014 from the National University of Malaysia in System Science and Management. Her research interests include ontology, archive management, digital library, information management, open access, scholarly publication and bibliometrics. Her work has been published in peer-reviewed journals, such as Malaysian Journal of Library & Information Science, Applied Ontology, Electronic library and Learned Publishing.



Noorhidawati Abdullah is a professor and the head of the Department of Library and Information Science, Faculty of Arts and Social Sciences, Universiti Malaya. She obtained her Ph.D. from the University of Strathclyde, Glasgow, in 2008. Dr. Noorhidawati is the Editor-in-Chief of the Malaysian Journal of Library & Information Science, which is indexed in both SSCI and Scopus. Her research interests include digital libraries, digital literacy, informetrics, and the research area of open science. Her work has been published in scholarly journals, such as Information Development, Scientometrics, Information Research and Library & Information Science Research.



Yi-Shuai Xu is a Ph.D. candidate at Universiti Malaya, specializing in Library and Information Science with a research focus on artificial intelligence. He received his Master of Library and Information Science from the same university. His doctoral research delves into advanced text classification techniques and recommendation algorithms, with the goal of enhancing information retrieval processes and develop more effective personalized recommendation systems.



Ye Yuan is currently a Ph.D. candidate in the Department of Library and Information Science at the Faculty of Arts and Social Sciences, Universiti Malaya. He received his Master's degree from Heilongjiang University in 2019, P.R. China. His research interests include open data, data management, and institutional repositories. His previous research has been published in *Journal of Academic Librarianship*, *Information development*.

Abstract

The convergence of Libraries, Archives, and Museums (LAMs) has become a key component of the digital humanities field, with research focal points evolving alongside rapid technological advancements. This study aims to explore the trajectories and development trends of LAMs' convergence within the digital humanities, addressing their critical role in knowledge organization practices and their impact on the ongoing digital transformation process. Drawing on relevant literature from Web of Science and Scopus, BERTopic and Dynamic Topic Modeling were applied to identify topics and analyze thematic evolution over time. The findings identify 32 research topics, which are grouped into four primary research directions: Engaged Knowledge Sharing, Digital Technological Competency, Cultural Digital Preservation, and Research & Sociocultural Impact. The analysis highlights how digital preservation and heritage initiatives enhance public interaction, facilitate knowledge exchange, and promote cultural awareness, reflecting the intersection of technology and culture. By integrating deep learning methodologies, this study addresses the limitations of conventional analytical approaches, offering novel insights into interdisciplinary collaboration and knowledge innovation. Furthermore, it delineates the transition from traditional to emerging technological paradigms within LAMs, exhibiting both homogeneity and complementarity. These insights serve as a valuable reference for institutions, researchers, and policymakers, informing strategies for digital adaptation and fostering more inclusive and efficient public cultural services.

Keywords: LAM institutions; digital humanities; BERTopic; topic modeling; evolutionary analysis

1. Introduction

Amidst the rapid advancements in information technology, digital humanities, as an emerging interdisciplinary field, is catalyzing a global revolution in knowledge exploration and cultural dissemination (Stewart and Simeonov, 2019). This field skillfully integrates the sophistication of digital technology with the profound depth of humanities research, employing innovative digital methods to broaden access to humanities knowledge, deepen analytical dimensions, optimize integration strategies, and enrich modes of dissemination and presentation (Lee, 2025; Su et al, 2023). Within this context, libraries, archives, and museums—collectively known as Libraries, Archives, and Museums institutions (LAMs)—serve as knowledge-organizing infrastructures tasked with preserving historical memory and fostering cultural heritage through structured representation of collective memory (Golub et al, 2021). However, these institutions face challenges of resource fragmentation and technological isolation. In the digital era, advancing the convergence of LAM institutions has emerged as a pivotal strategy for addressing diverse cultural needs, innovating service models, and ensuring the effective transmission of cultural resources (Moid et al, 2024; Zeng, 2019).

Traditionally, LAM institutions have distinct focuses: libraries prioritize knowledge dissemination, archives specialize in record management, and museums concentrate on the collection and exhibition of cultural heritage (McDonald and Levine-Clark, 2017; Ryholt and Barjamovic,

2019). However, the shift toward digitization has rendered this “siloe” model inadequate to meet the public's demand for diversified cultural resources, prompting LAMs to seek resource integration and technological innovation. The digital humanities framework, which emphasizes the digitization of resources, computational research methods, and interactive service models (Manovich, 2020; Schreibman et al, 2016), aligns closely with LAMs' digital practices such as semantic tagging, virtual exhibitions, and linked data. This synergy supports the analysis of thematic evolution within LAMs discussed in this research. By adopting digital convergence, LAM institutions enhance the efficiency of cultural resource sharing and utilization, establishing collaborative mechanisms that break traditional boundaries. For example, the Netherlands' Golden Age Digitization Program utilizes 3D scanning and augmented reality to blend 17th-century art with historical archives, creating new interactive cultural presentations (Crijns and Rademakers, 2009). Similarly, the Europeana Digital Library consolidates metadata from various LAMs to provide a unified digital access platform, enhancing the educational use and dissemination of cultural assets (Macri and Cristofaro, 2021). These initiatives exemplify how digital convergence facilitates broader preservation and sharing pathways, pushing public cultural services toward more sustainable and collaborative models (Qian et al, 2021; Valtysson et al, 2022).

Although existing research has explored LAMs' participation and collaborative models from perspectives such as case studies, technological integration, and policy analysis (Loach and Rowley, 2021; Rasmussen et al, 2022; Zeng, 2019), systematic examinations of how LAM institutions achieve resource sharing, technological complementarity, and management innovation through convergence mechanisms in the context of digital humanities remain insufficient. This study seeks to address this gap by employing advanced text-mining techniques—specifically, the BERTopic model and Dynamic Topic Modeling (DTM)—to systematically examine the development patterns of digital convergence among LAM institutions. The primary aim of this research is to clarify the practical pathways and developmental trajectories of deep convergence among LAM institutions in the digital humanities domain, emphasizing their pivotal roles as information and knowledge organization laboratories and the far-reaching impacts in the digital transformation process. The findings are intended to guide LAM institutions in leveraging technological innovation and strategic adjustments to better address the challenges of digitalization, while offering policymakers and researchers theoretical frameworks and practical insights for constructing inclusive and efficient public cultural service systems. This research is guided by the following three questions:

- RQ1: What key research topics have emerged in the study of LAMs' digital convergence from 2010 to 2024?
- RQ2: How do these topics reflect the primary research directions of LAMs' convergence within the digital humanities?
- RQ3: How have the evolving trends of these topics over the past 15 years shaped the progression and transformation of LAMs' digital convergence?

2. Literature Review

2.1 Development of the Concept of LAMs Convergence

LAMs have existed in various scales and forms throughout history, with their boundaries and definitions often being fluid and not clearly defined (Rasmussen et al, 2022). Rayward (1998) was the first to propose the concept of resource integration among LAMs, which formally initiated convergent research in this area. This pioneering work inspired scholars to explore how information technology could enhance the functional integration and innovation of services in these institutions. Since 2000, the term "LAMs" has been widely adopted, reflecting the growing interest among researchers in approaching these institutions through an integrated lens. The scope of convergent research has progressively expanded, encompassing multiple facets such as "institutional convergence", "physical convergence", and "digital convergence" (Vårheim et al, 2019; Warren and Matthews, 2018a). However, Robinson (2016) pointed out that, although the term "convergence" is widely used, significant differences persist in practice between in-

stitutions in areas such as partnerships, mergers, and reorganizations. These disparities, particularly in how they impact the roles of staff and institutional missions, require further investigation. Warren and Matthews (2018b) outlined three critical dimensions of the convergence framework: the integration of diverse collections, the co-creation of innovative cultural programs and services within traditional domains, and, most critically, the collaborative advancement of information management practices—which is increasingly driven by the adoption of shared metadata standards, linked open data, and semantic interoperability to support cross-institutional data integration.

As the world experienced rapid socio-economic and cultural development, many nations began to explore greater possibilities for the convergence of cultural resources across LAMs. Projects initiated by the Museums, Libraries, and Archives Council (MLA) in the UK and the Institute of Museum and Library Services (IMLS) in the U.S. have fostered LAMs collaboration through institutional restructuring and the application of new technologies (Warren and Matthews, 2018a). The European Commission's introduction of the term "cultural memory institutions" as a convergent concept for LAMs was also widely embraced (Dempsey, 1999). Meanwhile, national governments and international organizations have also actively developed policies and standards to promote digital collaboration among LAM institutions. Notably, the United Nations Education Scientific and Cultural Organization (UNESCO)'s initiatives on open access and digital heritage preservation have played a key role in guiding the construction, sharing, and utilization of digital resources (Von Schorlemer, 2020). Furthermore, acknowledging the differing digitization requirements of tangible and intangible heritage, LAMs have increasingly engaged in developing differentiated preservation frameworks to support the sustainable transmission of cultural heritage (Apaydin, 2022; Lázaro Ortiz and Jiménez de Madariaga, 2021). Historically, LAMs have primarily operated independently, but the contemporary trend toward convergence signals a renewed integration, now within the framework of digitization (Given and McTavish, 2010; Marcum, 2014). As research priorities have shifted from the integration of physical resources to the management of digital resources, service collaboration, and technological convergence, digital resources have increasingly become the central driving force behind LAMs' convergence and sustainable development (Loach and Rowley, 2021; Rasmussen, 2019).

2.2 Engaging LAM Institutions in the Digital Humanities

As cultural memory institutions and core sites of knowledge organization practice, LAMs have evolved from traditional information providers into foundational infrastructures for digital humanities (Golub and Liu, 2023). This engagement accelerates digital convergence through the field's transformative impact on institutional operations and

service models (Wang et al, 2020). Such convergence facilitates interdisciplinary collaboration via digital tools, firmly rooted in LAMs' dual roles as organizers of knowledge and stewards of cultural heritage (Golub et al, 2021; Wang and Ye, 2021). Rayward (1998) predicted that as the digital revolution progresses, the boundaries between LAMs would increasingly blur, rendering traditional institutional distinctions less significant. Within this context, emerging digital technologies such as 3D modeling and augmented/virtual reality (AR/VR) have not only spurred service innovation and resource sharing but have also prompted society to reassess and integrate the functions and roles of these institutions (Su et al, 2023). From the user's perspective, the integration of LAMs demonstrates the immense potential of digital tools in merging resources like artifacts, books, data, and archives, offering the public a rich interdisciplinary cultural experience (Marty, 2014; Rayward, 1998). Fueled by the wave of digitization, the digital convergence of LAMs has become a distinct and prominent topic of academic focus.

The rapid development of digital technologies has driven the formation and evolution of digital projects within LAMs (Dilevko and Gottlieb, 2004; Valtysson et al, 2022). Through theoretical analysis of collection resources and the application of advanced technologies such as metadata standards, linked data, visualization, and cloud computing, LAMs have achieved effective digital integration of their resources. For instance, the Europeana digital library and Library and Archives Canada (LAC) have integrated and linked various digital repositories, enabling users to access digital resources across institutions and retrieve information directly from original websites. These platforms have broken down data silos, significantly enhancing data discoverability and interoperability (Benardou and Dunning, 2017; Kandiuk, 2016). Moreover, initiatives such as the UK's "Resource Discovery Taskforce" and the U.S.'s Digital Public Library of America (DPLA) have significantly improved the efficiency of information retrieval through one-stop search platforms (Hlasten, 2017). Although these digitization projects have greatly fostered interdisciplinary collaboration and the integration of cultural resources, many LAM institutions still operate as independent entities, and their internal processes and operational mechanisms require further optimization (Zorich et al, 2008).

Of course, digitization extends beyond resource integration; its innovative practices in service convergence are also becoming increasingly prominent. For example, virtual museums like the London Science Museum have used digital means to extend physical exhibitions, offering entirely new interactive experiences (Schweibenz, 2019). At the same time, countries are actively exploring the construction of hybrid venues that integrate libraries, archives, and museums to enhance the efficiency of cultural resource utilization and the integration of physical spaces (Warren

and Matthews, 2018a). Practices such as those at the Kyoto Prefectural Comprehensive Archive in Japan illustrate the vast potential of digital convergence in enhancing public services (Pan, 2021). Overall, although the boundaries between LAMs have not been entirely eliminated, the prospects for digital convergence remain vast. As Rayward (1998) foresaw, digital technologies are driving the integration of resources and the innovation of services, making this an increasingly vital issue in the modern era.

2.3 Progress in Exploring Related Research Themes

The digital convergence of LAM institutions has prompted extensive scholarly exploration in areas such as cultural heritage preservation, digital resource management, user engagement, and technological innovation (Botticelli et al, 2019; Chen et al, 2024). Research indicates that digital technologies, particularly 3D modeling, virtual reality (VR), and augmented reality (AR), exhibit significant potential in driving service innovation for LAMs (Sebastián Lozano et al, 2023). Additionally, the concepts of open data and knowledge sharing have greatly advanced inter-institutional resource collaboration and joint research (Zeng, 2019). Within the digital humanities framework, global scholars have analyzed the paths toward LAMs convergence through various case studies. For instance, Chinese researchers have facilitated the digital preservation and dissemination of Lijiang Dongba cultural heritage by integrating LAMs resources (Kong and Hu, 2023). Similarly, Zahila et al. (2021) highlighted the use of event ontology frameworks to extract content from historical manuscripts, improving both the organization and accessibility of digital collections. These studies demonstrate not only practical outcomes of LAMs convergence but also provide essential experiences for guiding future research.

Understanding research themes and trends is vital for capturing the evolving dynamics of the LAMs field. In recent years, diverse analytical methods have been employed to investigate emerging topics and their progression. Candela and Carrasco (2021) applied text analysis tools to extract and visualize themes from large corpora, revealing a shift toward digital inclusivity in cultural preservation and educational practices. This highlights the increasing complexity of digital resource management across LAM institutions. Furferi et al. (2024) conducted a comprehensive review to categorize the applications of information technology in modern LAMs, emphasizing the crucial role that digitization plays in enhancing public engagement and cultural experiences. Rasmussen et al. (2022), through thematic coding, examined how regional LAM institutions are responding to societal transformations, underscoring the significance of collaborative projects, policy support, and digitization in fostering resource sharing and improving public services. Although these studies have deepened understanding of LAMs' collaborative development, existing research remains predominantly case-specific and has not yet

fully articulated the fundamental role of knowledge organization in enabling systematic and scalable digital humanities frameworks (Golub and Liu, 2023). Recent advancements in deep learning, particularly the BERTopic model, have been increasingly applied across disciplines like library and information science and computer science for their outstanding performance in topic extraction (Egger and Yu, 2022; Yang and Wu, 2024). This study aims to integrate the BERTopic model with DTM to extract and analyze themes from LAMs-related literature. Such an approach is intended to provide meaningful insights for advancing the digital convergence and collaboration among LAMs, while also informing future research directions and practical strategies.

3. Materials and Methods

This study integrates the BERTopic model and DTM to systematically analyze the thematic evolution and developmental trends of LAM institutions within the context of digital humanities. Through rigorous data selection and preprocessing, the bibliographic dataset is carefully curated to ensure precise alignment with the research topic. Subsequently, the BERTopic model is employed to identify core topics, while DTM is used to analyze the temporal evolution of these topics. This approach significantly enhances the efficiency and accuracy of topic extraction and dynamic analysis, providing profound insights and robust theoretical support for research on the digital convergence of LAM institutions. The following subsections detail the specific methodological steps.

3.1 Data Collection and Preprocessing

The study utilizes Web of Science (WOS) and Scopus as data sources. These databases are well-regarded for their comprehensive scope, inclusion of high-quality, peer-reviewed literature, and advanced search capabilities, indexing publications from over 5000 publishers to ensure rigorous quality control and diverse disciplinary coverage (Samsuddin et al, 2020). To guarantee the retrieved records align closely with the research focus, a combination of keywords related to “LAM institutions” and “digital humanities” was employed. This search yielded 746 records from WOS and 1346 from Scopus, resulting in a total of 2092 initial records ($N = 2092$). Detailed search queries are outlined in Table 1. During the data screening process, duplicate records were removed, and the following exclusion criteria were implemented: (i) articles published before 2010, (ii) exclusion of conference abstracts, book reviews, and other non-research articles, and (iii) retention of only English-language documents. After applying these criteria, a final dataset of 1159 bibliographic records spanning the period from 2010 to August 2024 was prepared for subsequent analysis. The cutoff date of August 2024 was selected to allow sufficient time for analysis and research preparation. Titles, keywords, and abstracts from the filtered records

were merged into unified text blocks and subsequently pre-processed using the Natural Language Toolkit (NLTK). The preprocessing steps involved tokenization, stop word removal, and other standard procedures to produce a normalized text dataset suitable for topic modeling and dynamic analysis.

The annual publication trends in the dataset reveal key insights into the developmental trajectory of this research field. As illustrated in Fig. 1, the period from 2010 to 2013 marked an initial growth phase, characterized by increasing academic attention and research activities. Although a brief fluctuation in publication volume occurred during 2014–2015, the field subsequently entered a phase of steady growth, reflecting sustained global interest in the convergence of digital humanities and LAM institutions. Despite the global disruptions caused by the COVID-19 pandemic in 2020, the field demonstrated relative stability and maturity over the next three years. Publication volume peaked at 122 papers in 2023, likely driven by rapid advancements in digital technologies and the proactive adaptation of LAM institutions to digitization and networking. Although 2024 data collection only extends to August, it is anticipated that the annual publication total will surpass current figures. Overall, the field has experienced significant growth after initial fluctuations, driven by continuous advancements in digital technologies that enable efficient digitization of cultural resources, alongside policy support and an increase in international collaboration projects. These factors collectively provide a solid foundation for further exploration of the digital convergence of LAM institutions.

3.2 Topic Modeling Methodology

This study employs the BERTopic model to identify topics in the titles, abstracts, and keywords of LAM-related publications. BERTopic, introduced by Grootendorst (2022), is a thematic modeling framework that integrates pre-trained language models deeply, enabling the capture of nuanced semantic relationships and latent topics within text. Compared to traditional methods such as Latent Dirichlet Allocation (LDA), BERTopic excels in key term extraction, depiction of inter-topic relationships, and thematic visualization. It offers higher thematic consistency, greater diversity, and enhanced interpretability (Egger and Yu, 2022). Additionally, BERTopic autonomously determines the optimal number of topics, mitigating issues of vocabulary mismatch and reducing subjective intervention. Its efficiently constructed topic representation framework provides robust support for the structured analysis of complex bibliographic data. This study employs the BERTopic model to uncover core topics and patterns within LAM-related literature. As shown in Fig. 2, the modeling process comprises three key stages:

(a) Embedding Documents: Document embedding is performed using a pre-trained model optimized for semantic similarity tasks (sentence-BERT). The “all-MiniLM-L6-

Table 1. Search strategy.

Databases	Search Query
Web of Science	TS=(((“librar*” AND “archiv*” AND “museum*”) OR (“LAM institutions” OR “cultural heritage institutions” OR “Memory Institutions” OR “LAM Sectors” OR “LAM Labs”)) AND (“digital*” OR “digi*” OR “digital humanities” OR “digital humanistic studies” OR “computational humanities”))
Scopus	TITLE-ABS-KEY(((“librar*” AND “archiv*” AND “museum*”) OR (“LAM institutions” OR “cultural heritage institutions” OR “Memory Institutions” OR “LAM Sectors” OR “LAM Labs”)) AND (“digital*” OR “digi*” OR “digital humanities” OR “digital humanistic studies” OR “computational humanities”))

TS, TOPIC; TITLE-ABS-KEY, TITLE-ABSTRACT-KEYWORDS.

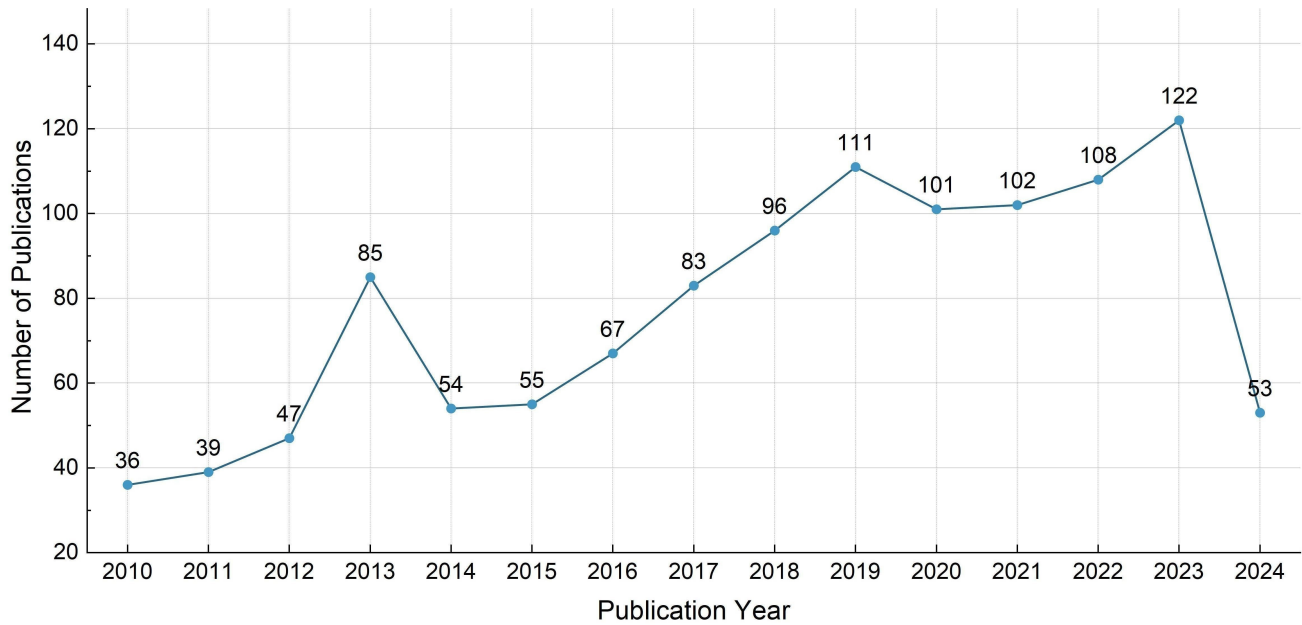


Fig. 1. Annual publication trends.

v2” model is used to vectorize the preprocessed data. After performing document embedding, the documents are transformed into a text corpus, serving as the foundational dataset for topic modeling.

(b) Clustering Documents: Uniform Manifold Approximation and Projection (UMAP) was used to reduce the dimensionality of the embeddings, followed by the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm to further reduce embeddings and create clusters of semantically similar documents. UMAP was initialized with 10 components based on multiple experimental results and literature references (Sánchez-Franco et al, 2022). HDBSCAN was initialized with a minimum cluster size of 25, meaning each cluster must contain at least 25 samples, with a cluster diversity setting of 0.5, to ensure both clustering effectiveness and diversity.

(c) Representing Topics: An enhanced Class-based Term Frequency–Inverse Document Frequency (c-TF-IDF) algorithm is used to calculate term importance, facilitating the identification of topics and the extraction of highly rel-

evant keywords. To enhance the comprehensiveness of the model, topics are further clustered based on inter-topic distances and cosine similarities, enabling the precise summarization of primary research directions and the construction of a static topic model.

To further analyze the temporal evolution of topics, this study introduces DTM, an extended thematic modeling approach that incorporates a temporal dimension to reveal patterns such as the emergence of new topics and the decline of older ones. The modeling process involves two key steps: (i) grouping publications by year to construct a time-series topic model and (ii) applying weighted averaging to smooth topic values over time, ensuring temporal continuity and accurately capturing dynamic thematic developments. By combining the strengths of BERTopic and DTM, this study systematically analyzes the complex topic structures and temporal dynamics within LAM-related literature, providing critical insights into the convergence of LAM institutions within the digital humanities domain.

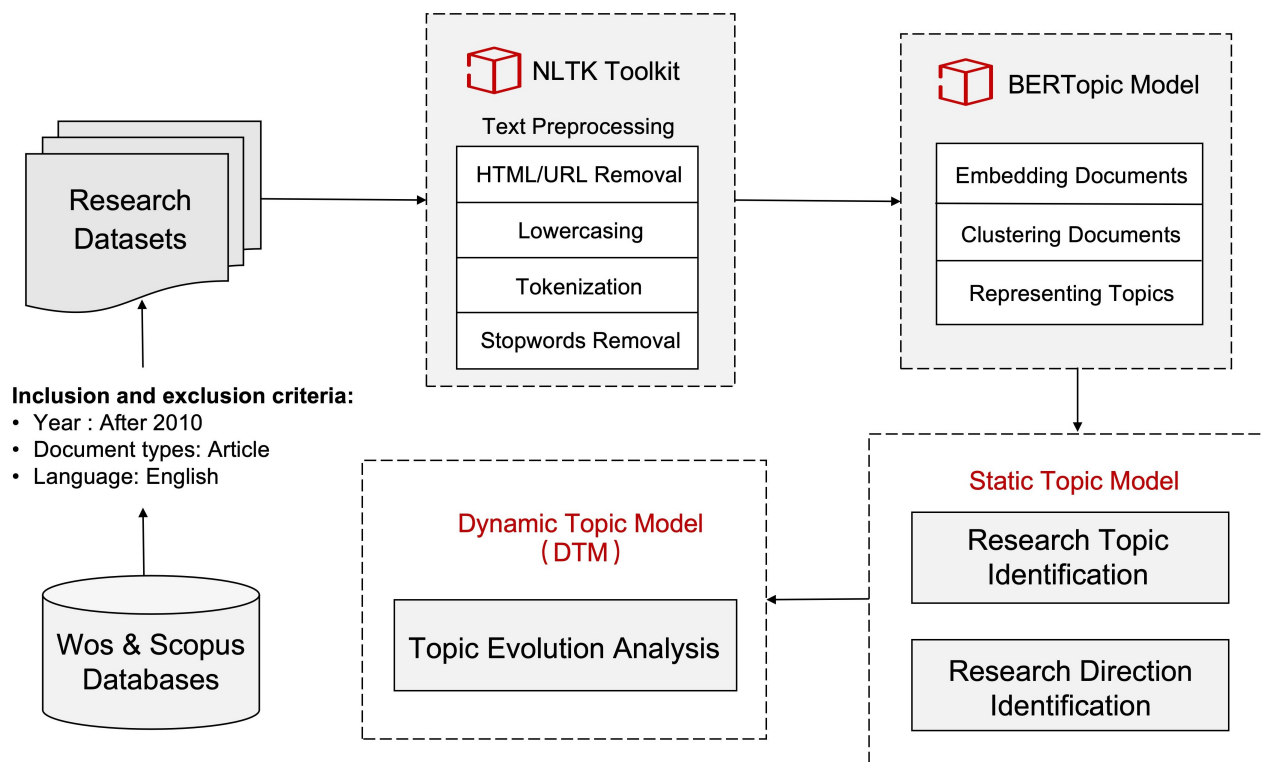


Fig. 2. Research framework. NLTK, Natural Language Toolkit.

4. Results

4.1 Identification of Research Topics

This study utilizes the BERTopic model to perform thematic modeling on bibliographic information from documents collected, including titles, keywords, and abstracts, related to the digital convergence of LAMs. By leveraging the Sentence-BERT framework, the study captures deep semantic relationships within the textual data through sentence vector embeddings. Ultimately, a text corpus of 9809 items was identified from 1159 documents. After extensive experimentation, the model was optimized to automatically identify 32 major research topics (Topic 0 to Topic 31), which encompass a broad range of perspectives within the field of LAMs' digital convergence. These topics provide significant guidance for future LAMs practices and digital humanities research, particularly in identifying growth areas in digital technologies and key academic research priorities. To ensure the representativeness of the thematic analysis, the model excluded topics that did not meet the minimum cluster size requirement (at least 25 textual entries per topic), categorizing them as noise data (5067 entries, labeled as Topic-1). Since these textual instances had minimal impact on the identified research topics, they were excluded from further analysis. Each identified topic is characterized by a set of feature terms, which are weighted based on their importance within the theme. The study found that, although each set typically consists of 10 terms,

most topics can be effectively captured by 3–5 core feature terms, with additional terms offering diminishing returns in thematic expression (Yang and Wu, 2024).

The BERTopic model generates the set of feature terms for each topic using the c-TF-IDF algorithm. Specifically, this method merges all textual entries within a cluster into a single “long text” and calculates term frequency to build a bag-of-words model. The Maximal Marginal Relevance (MMR) algorithm is then applied to eliminate terms that do not contribute to theme representation while enhancing term diversity and reducing synonym redundancy. Compared to traditional document-level bag-of-words approaches, this feature extraction method is highly effective in understanding key concepts and the underlying research context of topics, helping to clarify the structure and content of each topic. Fig. 3 illustrates the feature terms and their corresponding weights for the top 10 topics by representation frequency in the corpus, visually confirming that the BERTopic model outputs feature terms that effectively describe their respective topics. For example, Topic 0 (visitor_augmented_heritage_heritage_vr) highlights research on augmented reality and interactive heritage exhibitions designed for visitors, while Topic 2 (linked_metadata_interoperability_ontology_rdf) exemplifies core knowledge organization practices, demonstrating how semantic metadata frameworks enable cross-institutional resource integration in digital contexts. Additionally, feature term weights help identify core concepts



Fig. 3. Feature terms and weights of the top 10 research topics.

within topics, such as Topic 1 (preservation), Topic 3 (evaluation), and Topic 7 (multilingual), which indicate a focus on digital preservation, project evaluation, and multilingual integration as key areas of study.

Based on the detailed presentation of topic information in Table 2, the top ten topics collectively account for approximately 69% of the text corpus, indicating their dominant position in the field of LAMs' convergence. For example, Topic 0 alone covers 17.44%, making it the most prominent research theme in the field of LAMs collaborative development over the past 15 years, reflecting researchers' intense focus on augmented reality technologies and cultural heritage preservation. Similarly, Topic 1 and Topic 2, which account for 13.56% and 9.05%, respectively, emphasize the importance of digital archive preservation and metadata linking in LAM institutions. As digital technologies continue to advance, these themes are expected to remain highly relevant in future research. Although other topics hold relatively more minor shares, they remain crucial. For instance, Topic 12, which discusses artificial intelligence-generated content (AIGC), and Topic 20, which highlights data curation, have emerged as research hotspots in various academic fields and are likely to represent new trends and potential growth areas in the convergence of LAMs.

Therefore, a more comprehensive analysis of all 32 topics is necessary to explore the broad landscape of research on LAMs' digital convergence. In the LAMs field,

many feature terms, such as "orphan", "lutsk", "aboriginal" and "flickr" while specific and precise, may not fully convey the overall intent of the topic when considered in isolation. For instance, "orphan" refers to efforts by LAM institutions, encouraged by the government, to determine the copyright status of orphan works and to digitize and exhibit them. Similarly, "aboriginal" refers to the representation and digital preservation of Indigenous cultural heritage, enabling these cultural resources to be disseminated and utilized for educational purposes in the digital age. Thus, a more comprehensive analysis of these feature term sets, combined with their specific application in the context of the literature, allows for the assignment of more intuitive and descriptive names to each topic. These themes span multiple key research areas, from resource coordination and management to the application of digital technologies. Such labeling helps define each topic's hierarchical role within the research landscape of LAMs' digital convergence, while also reflecting knowledge organization practices that support the classification and interpretation of research themes in broader academic discourse.

4.2 Clustering Analysis of Research Directions

Building on the 32 previously identified research topics, this study further applies the BERTopic model to conduct hierarchical clustering based on the distance and cosine similarity between topics. By analyzing the shared characteristics and commonalities of the clustered topics,

Table 2. Topic identification results and proportion.

Topic	Topic name	Top keywords	Text	Prop.
0	Interactive Heritage Exhibitions	visitor, augmented, heritage...	827	17.44%
1	Digital Preservation and Access	preservation, archivist, archival...	643	13.56%
2	Metadata Interoperability	linked, metadata, interoperability...	429	9.05%
3	Collaborative Evaluation Strategies	qualitative, evaluation, collaboration...	384	8.10%
4	Copyright Integration in Digital Technology	orphan, directive, exception...	213	4.49%
5	Digital Conservation and Reproduction of Heritage	moving, photographic, rephotography...	175	3.69%
6	Digital Memory and Multi-source Evidence	intelligence, cloud, forgotten...	160	3.37%
7	Multilingual and Cross-Cultural Integration	multilingual, europe, aggregator...	148	3.12%
8	Academic Exchanges and Cultural Exhibitions	luts, petersburg, academy...	145	3.06%
9	Indigenous Culture and Digital Knowledge Sharing	aboriginal, mukurtu, sovereignty...	144	3.04%
10	Audiovisual Archiving and Recording	audiovisual, musicology, recording...	143	3.02%
11	Crowdsourcing and Volunteer Engagement	crowdsourcing, motivation, participation...	138	2.91%
12	AIGC and Technological Innovations	aigc, sector, bias...	117	2.47%
13	Digital Printing and Environmental Stability	humidity, inkjet, gloss...	113	2.38%
14	Digital Infrastructure and Knowledge Popularization	infrastructure, cdssk, cross...	86	1.81%
...	...	...	...	...
31	Literary Criticism in Digital Humanities	estonian, literary, criticism...	29	0.61%

The complete topic name can be found in the Table 3. AIGC, artificial intelligence-generated content.

four primary research directions in the field of LAMs' digital convergence have been revealed: Engaged Knowledge Sharing, Digital Technological Competency, Cultural Digital Preservation, and Research & Sociocultural Impact. The results of the topic clustering are presented in Fig. 4.

The clustering analysis of research directions provides important insights into the relationships between topics and their collaborative development. Table 3 presents the research direction identifiers, the number of associated texts, and their respective proportions. Overall, Direction C (Cultural Digital Preservation) appears most frequently, accounting for nearly half of the total volume and encompassing 12 research topics. This result not only reflects the scholarly community's strong focus on this area but also demonstrates that this direction, as a mainstream domain in LAMs research, is continually expanding and deepening, showing a trend toward greater specialization and refinement. Similarly, Direction A (Engaged Knowledge Sharing) also stands out, representing a significant proportion of 32.35% of the total texts. Although the topics within this direction are more narrowly focused, their academic prominence indicates significant potential for further development and practical application. Notably, while Direction B (Digital Technological Competency) and Direction D (Research & Sociocultural Impact) exhibit relatively lower text frequencies, their thematic diversity indicates that scholars have conducted multi-dimensional explorations in these fields. Against the backdrop of future societal transformation and institutional advancement, these areas are likely to become new research hotspots, attracting increasing academic attention and discussion. In the following sections, this study will analyze each research direction and its corresponding topics in detail, aiming to gain deeper insights.

(a) Direction A: Engaged Knowledge Sharing

This direction highlights seven research topics focused on leveraging digital technologies to enhance cultural resource sharing, foster meaningful public engagement, and advance knowledge exchange within the digital humanities. These topics emphasize how digital tools such as augmented reality and virtual exhibitions (T0: Interactive Heritage Exhibitions), as well as crowdsourcing platforms (T11: Crowdsourcing and Volunteer Engagement), disrupt traditional dissemination models by attracting global audiences and increasing interactivity. Furthermore, by leveraging multilingual support and cross-cultural integration (T7), digital platforms facilitate knowledge sharing and the seamless transmission of cultural resources worldwide. Notably, topics like T8 (Academic Exchanges and Cultural Exhibitions) showcase how international conferences and exhibitions streamline academic exchange, while T23 (Intercultural Integration and Digital Consumerism) explores the evolution of cultural consumption patterns in a digital environment and its impact on global cultural dissemination and consumer behavior. Topics such as Indigenous cultural preservation (T9) and the democratization of cultural education (T15) underscore the inclusive nature of knowledge sharing and the promotion of cultural diversity to ensure equal access to knowledge. Collectively, these themes reveal how public engagement in heritage protection and knowledge sharing can be enhanced, and how cultural institutions can foster the global dissemination of cultural resources and educational outreach. Overall, this direction emphasizes the use of innovative digital tools and participatory models to expand the reach of cultural institutions and strengthen public engagement and identity.

Hierarchical Clustering

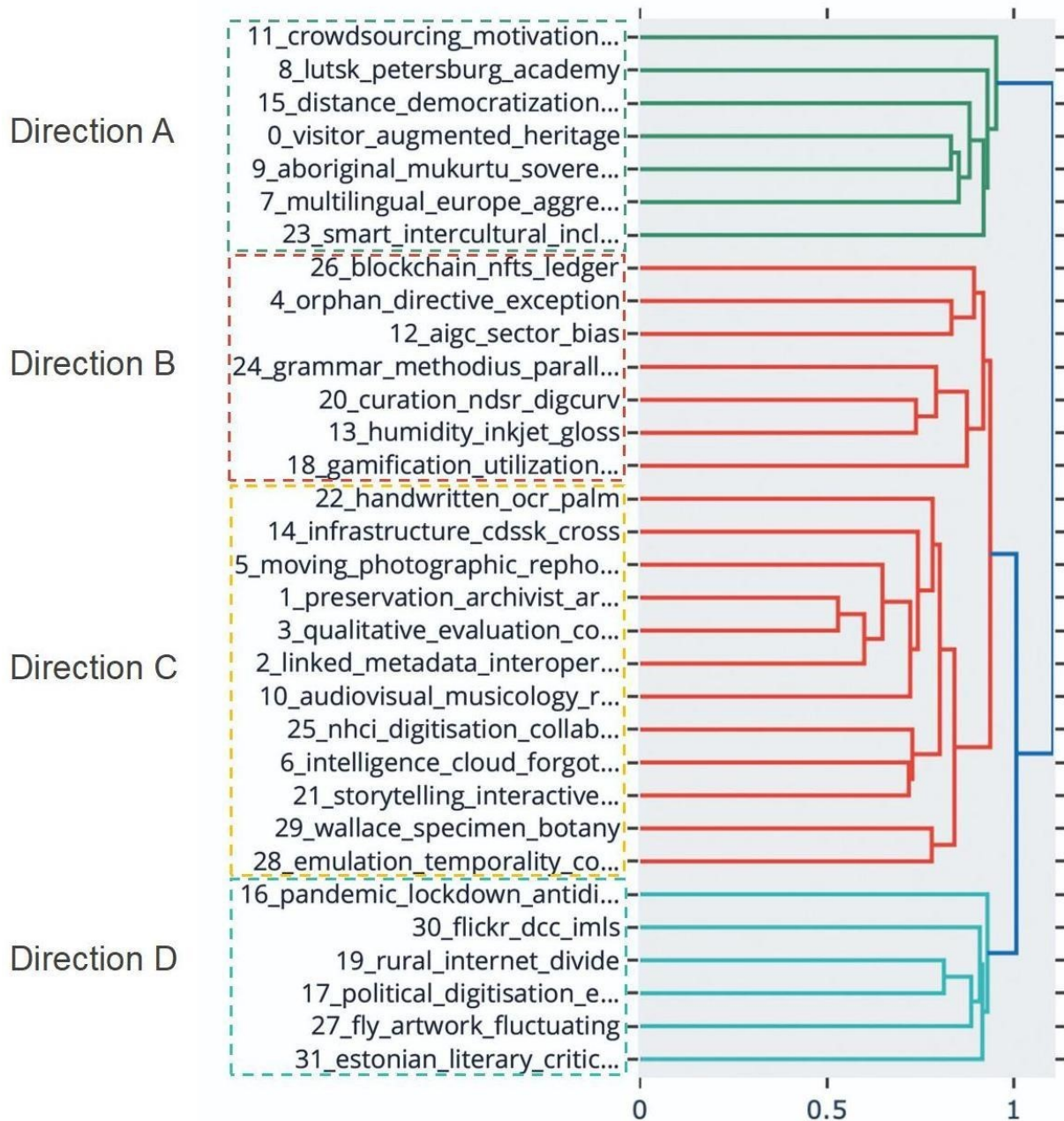


Fig. 4. Latent hierarchy of topic distribution.

(b) Direction B: Digital Technological Competency

The direction B covers seven research topics that focus on the role of emerging technologies and innovation in driving the digital transformation of cultural institutions, as well as the collaborative progress of LAMs in technological applications and competency development. The research can be divided into two primary branches: the first branch focuses on the application of emerging technologies, such as blockchain and Non-Fungible Tokens (NFTs)

(T26) and artificial intelligence (T12), which are redefining management and dissemination models in the cultural heritage domain. These technologies create new revenue streams (e.g., NFTs providing additional income for artists and museums), but also pose challenges related to digital copyright (T4), especially concerning cross-border usage, orphan works, and policy adaptation. Additionally, this branch critically explores the environmental and legal issues associated with technological opportunities. The sec-

Table 3. Research direction and topic distribution.

No.	Research directions	Associated topic numbers & names	Text	Prop.
A	Engaged Knowledge Sharing	T0: Interactive Heritage Exhibitions	1534	32.35%
		T7: Multilingual and Cross-Cultural Integration		
		T8: Academic Exchanges and Cultural Exhibitions		
		T9: Indigenous Culture and Digital Knowledge Sharing		
		T11: Crowdsourcing and Volunteer Engagement		
		T15: Digital Democratization of Cultural Education		
B	Digital Technological Competency	T23: Intercultural Integration and Digital Consumerism	660	13.92%
		T4: Copyright Integration in Digital Technology		
		T12: AIGC and Technological Innovations		
		T13: Printing Technology and Environmental Stability		
		T18: Gamification Techniques in Data Management		
		T20: Digital Curation and Competency Development		
C	Cultural Digital Preservation	T24: Linguistic Standards in Parallel Digital Management	2243	47.30%
		T26: Blockchain and NFTs in Cultural Heritage		
		T1: Digital Preservation and Access		
		T2: Metadata Interoperability		
		T3: Collaborative Evaluation Strategies		
		T5: Digital Conservation and Reproduction of Heritage		
		T6: Digital Memory and Multi-source Evidence		
		T10: Audiovisual Archiving and Recording		
		T14: Digital Infrastructure and Knowledge Popularization		
D	Research & Sociocultural Impact	T21: Interactive Storytelling in Heritage	305	6.43%
		T22: Digital Recognition for Ancient Manuscripts		
		T25: Digital Collaboration and Institutions Integration		
		T28: Emulation Technology for Digital Preservation		
		T29: Specimen Management and Biodiversity		
		T16: Digital Research Adaptations During the Pandemic		
		T17: Political Digitization and Historical Reconstruction		
		T19: Digital Divide and Social Impact		
		T27: Symbolism and Value Fluctuations in Artworks		
		T30: Feasibility of Digital Collection Platforms		
		T31: Literary Criticism in Digital Humanities		

NFTs, Non-Fungible Tokens.

ond branch focuses on establishing technical standards and enhancing digital competencies, showcasing how technology can address environmental challenges, enhance user engagement, and maintain accuracy in global management practices. Topics such as Linguistic Standards in Digital Management (T24), Digital Curation and Competency Development (T20), and Gamification Techniques in Data Management (T18) examine diverse aspects of managing multilingual content, enhancing curatorial skills, and boosting data management engagement through gamification. Overall, this direction highlights the optimization of operational efficiency and precision within cultural institutions, enhancing LAMs' technological proficiency and providing practical guidance for the sustainable management of cultural assets.

(c) Direction C: Cultural Digital Preservation

The direction C focuses on how LAMs play a pivotal role in safeguarding cultural memory and re-enacting

heritage, with a particular emphasis on their involvement in digital convergence and preservation efforts. Through a comprehensive exploration of core trends in cultural preservation, topics such as Digital Preservation and Access (T1) and Collaborative Evaluation Strategies (T3) underscore the importance of knowledge organization-driven standardization in digital preservation, emphasizing that effective heritage preservation requires not only technical support but also legal safeguards, project evaluation, and multi-stakeholder collaboration to enhance efficiency. Topics such as Digital Conservation and Reproduction of Heritage (T5) and Audiovisual Archiving and Recording (T10) address the complex challenges multimedia resources face in digitalization. For example, tangible heritage (such as monuments and artifacts) and intangible heritage (such as oral traditions and craftsmanship) present diverse digital preservation needs: the former relies on precision replication and 3D modeling technologies, while the latter requires multi-

modal recording and community participation to effectively capture contextual knowledge. Meanwhile, T2 (Metadata Interoperability) emphasizes the critical role of metadata in ensuring the long-term preservation and accessibility of digital assets across platforms and institutions. Additionally, topics like Digital Collaboration and Institutional Integration (T25), Digital Memory and Multi-source Evidence (T6), and Interactive Storytelling in Heritage (T21) explore cross-institutional collaboration, the role of multi-source evidence in digital memory preservation, and the use of interactive storytelling to engage the public with heritage. Unique approaches, such as Emulation Technology for Digital Preservation (T28) and Specimen Management and Biodiversity (T29), offer solutions for preserving software-dependent cultural objects, further enhancing the sustainability of digital assets. Collectively, these themes highlight the central role of digital technologies in cultural memory preservation, particularly through standardization, multi-source evidence, and cross-institutional collaboration, thus driving the global effort toward sustainable digital preservation and management of cultural heritage.

(d) Direction D: Research & Sociocultural Impact

This direction gathers six research topics, each examining the multifaceted role of digitization in reshaping sociocultural structures, offering new perspectives on understanding social dynamics and academic development in the digital age. Topics such as Digital Research Adaptations during the Pandemic (T16) and Political Digitization and Historical Reconstruction (T17) emphasize the role of digitization in enhancing research adaptability during global crises and reshaping historical narratives, advocating for transparency through remote collaboration, data sharing, and public participation. Digital Divide and Social Impact (T19) and Symbolism and Value Fluctuations in Artworks (T27) address the sociocultural implications of digitization, with the former focusing on social inequality exacerbated by the digital divide, and the latter exploring how digital technologies reshape the symbolic meaning and cultural value of artworks. Meanwhile, Feasibility of Digital Collection Platforms (T30) highlights the potential of digital platforms in enhancing resource sharing and user engagement, while Literary Criticism in Digital Humanities (T31) demonstrates how digital humanities offer new avenues for literary criticism and expanded influence. Together, these topics reflect the dual-edged nature of digital technologies—as tools that advance academic research and cultural preservation but also potentially exacerbate existing social inequalities.

Through a holistic examination of these four research directions, this study not only showcases LAMs' active advancements in digitalization and technological innovation but also reveals how these technologies deepen public engagement, strengthen cultural preservation, and promote knowledge sharing and academic innovation. It is important to note that these directions are not isolated; instead,

they demonstrate the interconnected and collaborative development of themes, reflecting the overall progress and coordinated evolution within this field. The thematic analysis presented in this study highlights the multidimensional role and broad impact of digitalization in modern society, offering valuable perspectives and insights for future research and practice.

4.3 Trends in Research Directions and Topic Evolution

The DTM method employed by BERTopic provides a clear visualization of the shifts in research popularity and scholarly attention within a given field, making it easier to analyze the evolutionary trends of topics across different research directions in the digital collaborative development of LAMs. In this study, we use one-year intervals as time stamps to represent the evolution of topics. The number of relevant texts is regarded as an indicator of the research intensity on each topic. Figs. 5,6, respectively, illustrate the evolution of four primary research directions in the field of LAMs convergence development over the past fifteen years, along with the temporal changes in their associated research topics.

From the data trends in Fig. 5, it is evident that the four main research directions within the field of LAMs convergence development exhibit distinct evolutionary characteristics. These trends provide clear insights into the reorganization of research knowledge structures, shifts in scholarly interest, and the emergence of key thematic domains. In Fig. 6, each colored line (automatically assigned by the system) represents a different research topic. Through an analysis of both macro and micro data features, combined with a retrospective review of relevant literature, four critical evolutionary trends in LAMs convergence research within the digital humanities framework have been identified:

(a) Core-Driven Growth with Collaborative Diversification

Amid the flourishing development of research topics, the overall trend is characterized by the multidimensional evolution of “Cultural Digital Preservation”, fostering deeper public engagement, promoting knowledge sharing, and enhancing social awareness. Direction C, “Cultural Digital Preservation”, has maintained a dominant position, with steadily increasing research intensity, leading to the emergence of more specialized sub-topics. The promotion of cultural preservation and heritage through digital technologies has long been a central goal of LAMs convergence. Sub-topics within this direction, such as cultural memory preservation and heritage (T1, T2, T3), have remained at the forefront of research, while secondary themes such as digital preservation, evidence integration, and infrastructure development (T5, T6, T14) have also gained prominence, reflecting the sustained and expanding nature of research in this area. Additionally, the other three directions exhibit unique growth patterns. Direction A has experienced significant fluctuations, peaking in 2013 and 2019,

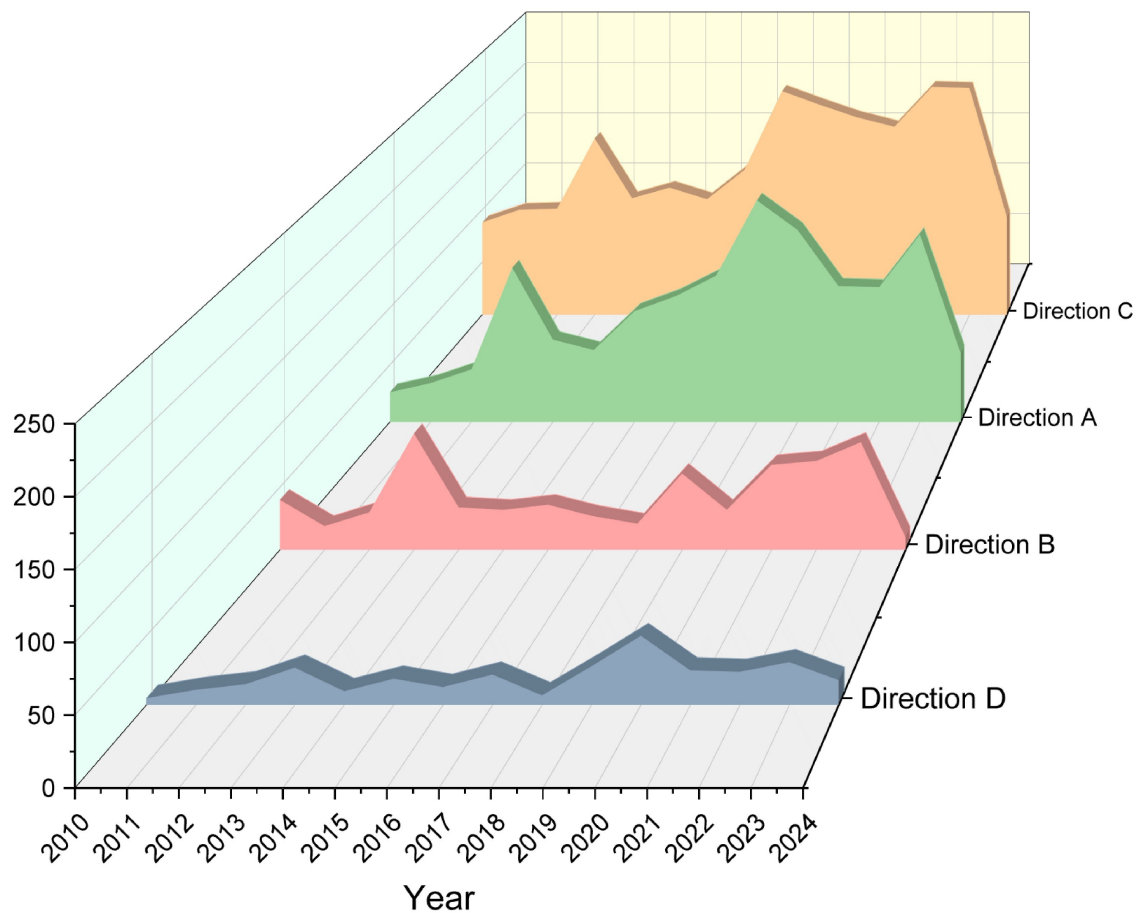


Fig. 5. Evolutionary trends of research directions (3D).

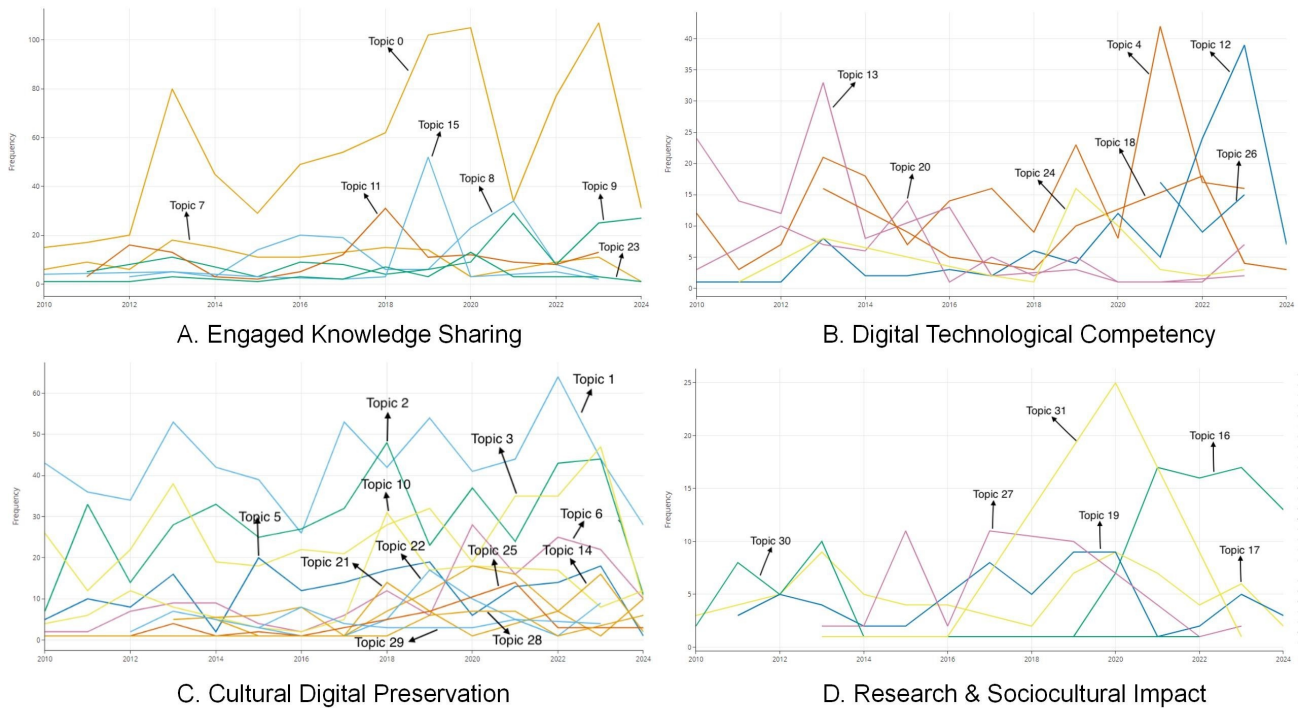


Fig. 6. Evolutionary trends of topics within directions A–D (2D). Each sub-figure illustrates the evolution of the topic in each direction.

followed by a decline in 2021–2022 due to external factors such as the pandemic. The trend of T0 (Interactive Cultural Heritage Exhibition) aligns with this overall pattern, underscoring its prominence and the influence of external factors, such as public participation and available spaces. In contrast, Direction B shows more segmented growth, with several technological topics (e.g., T13, T4, T12) peaking at different times, reflecting the academic community's strong focus on technological innovation. Although Direction D has a relatively lower research intensity, it shows a noticeable upward trend, with its diverse topics demonstrating unique research value in the context of global societal changes. For instance, topics like T16 (Digital Research Adaptation During the Pandemic) and T19 (Digital Divide and Social Impact) emphasize the application of digital technology in research adaptability and social inequality, highlighting the growing influence of sociocultural contexts on these topics.

(b) Coexistence of Homogeneity and Complementarity

As LAMs digital convergence research continues to thrive, the evolution of research hotspots over time reveals a dynamic pattern of both synchronicity and complementarity. During specific periods (e.g., 2013 and 2023), multiple research directions reached their peaks simultaneously, reflecting the vitality and robust development of the entire field. However, in other periods (e.g., 2018 and 2020), the research intensity exhibited a pattern of reciprocal shifts. Notably, in 2018, Directions C and A reached their peak intensity, while Directions B and D were relatively subdued. Conversely, 2020 saw an increase in Direction D's research intensity, accompanied by a decline in the others. These fluctuations may be attributed to the natural progression of research topics and shifts in societal focus. Moreover, the evolution of individual research topics reflects a dynamic balance and mutual adjustment, showcasing the regulatory interplay among different research themes.

(c) Shift from Traditional to Emerging Technologies

Technological transformation and the evolution of research are inevitable processes in the course of historical development. In the field of digital humanities, the close integration of technology and culture has significantly influenced research directions, particularly due to the high reliance on digital technologies. Over time, the focus of technological research has shifted from the early stages of digital preservation to the use of emerging technologies to realize cultural value, incorporating innovations such as AIGC, blockchain, and digital interaction technologies. This shift has significantly facilitated the preservation of cultural heritage and innovation in digital art. In particular, Direction B demonstrates clear stages of technological transformation. Between 2012 and 2014, the rise of topics T4 and T18 marked in-depth discussions on the modernization of traditional copyright integration and printing technologies through digital advancements. With ongoing technological

progress, recent years have seen a notable increase in interest in emerging technologies such as Generative AI, gamification technology, blockchain, and NFTs (e.g., T12, T18, T26). The interaction between these technologies and cultural resource management has become especially prominent post-2021. Similar trends are observed in other research directions, where traditional topics have been revitalized through the application of emerging technologies. For instance, topics like T17, T16, and T0 highlight new developments in technological adaptability and innovation. The application of these technologies not only advances cultural heritage preservation and digital art but also facilitates the strategic repositioning of LAMs within the digital age, achieving deeper integration between culture and technology.

(d) Interdisciplinary Integration with Expanding Boundaries

Within the digital humanities framework, LAMs convergence research is characterized by significant cross-cultural and interdisciplinary integration, reflected in the diversification of research topics and the increasing depth of specialization. Technological advancements and evolving societal needs have driven LAMs research to continually expand its disciplinary boundaries, integrating knowledge and technologies from various fields to address broader cultural and academic challenges. Particularly in Directions C and D, multidisciplinary research—ranging from audiovisual to specimen-based studies (e.g., T10, T22, T28, T29)—has enhanced the social and cultural value of research outcomes through interdisciplinary methods and technologies. Meanwhile, topics such as politics, artistic value, and humanistic critique (T17, T27, T31) have drawn widespread scholarly attention, driven by shifts in social relevance. Additionally, topics in Directions A and B, such as cross-cultural integration and digital consumerism (T23), as well as blockchain and NFTs (T26), underscore the fusion of LAMs institutions with emerging digital technologies. This fusion has not only transformed the operational dynamics of the art market but also facilitated the formation of new economic models within the cultural industries. This interdisciplinary research enriches the academic layers within the LAMs field, breaking traditional disciplinary boundaries and offering new possibilities for cultural education and public engagement.

5. Discussion

This study employed BERTopic and DTM to analyze literature related to the digital convergence of LAM institutions, systematically identifying and tracking the evolution of research topics. Over the past fifteen years, 32 research topics have been identified and categorized into four primary directions: Engaged Knowledge Sharing, Digital Technological Competency, Cultural Digital Preservation, and Research & Sociocultural Impact. The findings high-

light the continued deepening of LAMs convergence within the digital humanities, with research hotspots evolving in response to societal needs and technological advancements. The robustness of the BERTopic model in identifying key topics within the literature on LAMs' digital convergence underscores its significance for future research and the digital convergence practices of LAM institutions.

The identified topics and directions align closely with existing research, especially in the areas of digital preservation of cultural memory, technological development, and knowledge sharing. In engaged knowledge sharing direction, this study emphasizes the technology-driven nature of engagement experiences, covering interactive heritage exhibitions, multilingual and cross-cultural integration, and crowd-sourcing. These findings resonate with the conclusions of Furferi et al. (2024) and Valtysson et al. (2022) regarding the use of artificial intelligence and augmented reality in public engagement. In digital technological competency, the study echoes the consensus of Candela and Carrasco (2021) on digital technologies in such cultural and memory institutions, with a focus on AI-generated content, blockchain, and NFTs. Regarding cultural digital preservation, this study's analysis of metadata interoperability and digital infrastructure aligns with Zeng's (2019) discussions on semantic enhancement and interoperability. The importance of knowledge organization and digital collaboration, as proposed by Rasmussen et al. (2022), is further affirmed through this study's focus on digital collaboration and multi-source evidence preservation.

While this study corroborates much of the existing literature, it also introduces new insights in key areas. In terms of public engagement, the study goes beyond Furferi et al.'s (2024) focus on technological interaction by shifting attention to the democratization of digital cultural education, emphasizing the crucial role of LAM institutions in promoting cultural education and social equity. This perspective extends the social justice framework proposed by Apaydin (2022) and Groenewoud (2023), highlighting how inclusive digital practices and equitable access to cultural knowledge can help address long-standing structural imbalances within heritage institutions. The integration of cross-cultural and digital consumerism complements Candela and Carrasco's (2021) analysis, promoting the sharing and consumption of cultural knowledge. In technological innovation, the study introduces discussions on gamified data management and digital curation skills, expanding the scope of LAMs technologies. The most notable contributions lie in sociocultural impact, where the analysis of the digital divide and social impact, as well as symbol and value fluctuation, reveals LAMs' pivotal role in addressing global cultural challenges, particularly in adaptation during the pandemic. This extends Valtysson et al.'s (2022) research on platform transformations. While this study aligns broadly with previous research, it offers new insights into social responsibility, cross-cultural integration, and techno-

logical optimization, providing a richer framework for future LAMs development and digital heritage integration.

Using dynamic topic analysis, this study reveals the evolving trends in LAMs convergence over the past 15 years, reflecting both prior research and new insights. Overall, the findings reinforce the view that interdisciplinary and cross-sector collaboration—key to overcoming institutional silos—is essential for LAM institutions to drive digital humanities development (Golub and Liu, 2023). Consistent with Rasmussen (2019) and Loach and Rowley (2021), cultural digital preservation remains a central issue, underscoring the ongoing scholarly attention to digital resource management and long-term preservation. Technological advancements, cultural sharing, and public engagement have become critical in driving LAMs integration and sustainable development, accelerating the transition of cultural resources to digital formats. This supports Rayward's (1998) prediction of blurred boundaries between LAM institutions as digital humanities practices evolve. The study also reveals underexplored topics, such as cross-cultural integration, political digitalization, historical reconstruction, and pandemic adaptability, all of which show significant growth, demonstrating that LAMs are active participants in addressing global cultural and societal challenges.

As societal needs evolve and artificial intelligence technologies advance rapidly, future research should focus on the following four directions to address emerging challenges and seize development opportunities: (a) In-depth Exploration of Emerging Technologies: Investigate the application of large language models (LLMs), AIGC, intelligent platforms, and the metaverse in LAMs, particularly in addressing ethical and legal challenges in cultural preservation; (b) Diversified Public Engagement Models: Explore public engagement across different cultural contexts, assessing the role of virtual and augmented reality in enhancing public interaction; (c) Ongoing Research on Sociocultural Impact: Study the profound societal and cultural impact of digital processes, especially in global events, and explore how digital technologies can aid in sustainable management of cultural resources and social responsibility; (d) Innovative Interdisciplinary Integration: Promote cross-disciplinary collaboration between LAMs and fields such as data science, law, and sociology to develop innovative solutions for the complex challenges of the digital age. These directions will showcase LAMs' potential in technological innovation and cross-cultural exchange, providing strategic guidance for future convergence and enhancing global cultural influence. Ultimately, they will drive cultural heritage preservation and dissemination into a more open, interconnected digital era.

6. Conclusions

This study, based on deep learning techniques, employs BERTopic topic modeling and DTM to systemati-

cally analyze the role of LAMs convergence in the process of digital transformation. It uncovers 32 key topics and four core directions, along with their evolution trends, presenting a framework in which cultural digital preservation is at the core, with knowledge sharing, technological innovation, and sociocultural impact flourishing together. The research enriches the theoretical foundation of LAMs within digital humanities and offers reflective insights into the implications of digital preservation for cross-cultural dialogue, with a particular emphasis on how innovative technologies enhance public engagement, advance information and knowledge organization, and reframe cultural resource management.

Furthermore, the study clarifies the critical role of digital technologies in advancing the preservation and dissemination of cultural heritage within LAMs, addressing gaps in existing research regarding sociocultural impact and cross-cultural interaction. However, there are certain limitations to the study. First, the data are drawn solely from publicly available literature, potentially overlooking regional or unpublished research, which could affect the breadth of the analysis. Second, the topic categorization in the BERTopic and DTM models may be influenced by the structure of the corpus and model parameters, resulting in some limitations in the accuracy of identifying emerging topics. Future research should delve deeper into the mechanisms of each topic, integrating qualitative analysis to develop a more comprehensive understanding of LAMs' digital transformation. Additionally, the potential of emerging technologies such as artificial intelligence and blockchain in cultural resource management warrants further exploration. Overall, this study offers new insights into the role of LAMs in the digital era, envisioning a more open and interconnected global cultural future and demonstrating how technological innovation can effectively support the protection and dissemination of cultural heritage.

Availability of Data and Materials

All data reported in this paper will be shared by the correspondence author upon reasonable request.

Author Contributions

GW and YIAMK designed and drafted the research study. YIAMK and NA supervised the study, providing overall guidance and critical advice throughout the research process. GW performed the data processing and analysis, and wrote the manuscript. YX and YY contributed to parts of the initial drafting, participated in data interpretation, and provided critical revisions to enhance the manuscript's intellectual content. All authors have reviewed and approved the final manuscript, and contributed to its editorial revisions. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Acknowledgment

We would like to express our sincere gratitude to all individuals who provided valuable suggestions and constructive feedback, which significantly improved the quality of this article.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

Declaration of AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work the authors used InstaText in order to check spell and grammar. After using this tool, the authors reviewed and edited the content as needed and takes full responsibility for the content of the publication.

References

- Apaydin V. Heritage, education and social justice. Cambridge University Press: Cambridge. 2022.
- Benardou A, Dunning A. From Europeana Cloud to Europeana Research: Tools, users and methods. In Benardou A, Champion E, Dallas C, Hughes L (eds.) Cultural heritage infrastructures in digital humanities (pp. 136–152). Routledge: London. 2017. <https://doi.org/10.4324/9781315575278-10>.
- Botticelli P, Mahard MR, Cloonan MV. Libraries, archives, and museums today: Insights from the Field. Rowman & Littlefield Publishers: Lanham. 2019.
- Candela G, Carrasco RC. Discovering emerging topics in textual corpora of galleries, libraries, archives, and museums institutions. *Journal of the Association for Information Science and Technology*. 2021; 73: 820–833. <https://doi.org/10.1002/asi.24583>
- Chen H, Kim J, Chen J, Sakata A. Demystifying oral history with natural language processing and data analytics: a case study of the Densho digital collection. *The Electronic Library*. 2024; 42: 643–663. <https://doi.org/10.1108/el-12-2023-0303>
- Crijns H, Rademakers A. Memory of the Netherlands: Past and Future of the Dutch National Database on Cultural Heritage. Alexandria. 2009. 21: 1–17. <https://doi.org/10.7227/alx.21.2.2>
- Dempsey L. Scientific, industrial, and cultural heritage: a shared approach. *New Review of Information and Library Research*. 1999; 5: 3–29.
- Dilevko J, Gottlieb L. The evolution of library and museum partnerships: Historical Antecedents, Contemporary Manifestations, and Future Directions. Bloomsbury Publishing USA: Westport. 2004.
- Egger R, Yu J. A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twit-

- ter Posts. *Frontiers in Sociology*. 2022; 7: 886498. <https://doi.org/10.3389/fsoc.2022.886498>
- Furferi R, Di Angelo L, Bertini M, Mazzanti P, De Vecchis K, Biffi M. Enhancing traditional museum fruition: current state and emerging tendencies. *Heritage Science*. 2024; 12: 20. <https://doi.org/10.1186/s40494-024-01139-y>
- Given LM, McTavish L. What's old is new again: the re-convergence of libraries, archives, and museums in the digital age. *The Library Quarterly*. 2010; 80: 7–32. <https://doi.org/10.1086/648461>
- Golub K, Kamal AM, Vekselius J. Knowledge organisation for digital humanities: An introduction. *Information and Knowledge Organisation in Digital Humanities* (pp. 1–22). Routledge: New York. 2021. <https://doi.org/10.4324/9781003131816-1>
- Golub K, Liu YH. *Information and knowledge organisation in digital humanities: Global perspectives*. Routledge: New York. 2023.
- Groenewoud M. CHAPTER 13 Digital humanities, social justice and the Pluricultural realities of Dutch Caribbean heritage archives. In Stipriaan A, Alofs L, Guadeloupe F (eds.) *Caribbean Cultural Heritage and the Nation: Aruba, Bonaire and Curaçao in a regional context* (pp. 235–250). Leiden University Press: Leiden, NL. 2023. <https://doi.org/10.24415/9789400604278-015>
- Grootendorst, M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv*. 2022. <https://arxiv.org/abs/2203.05794> (preprint)
- Hlasten Y. *Library Linked Data: Early Activity & Development*, edited by Erik T. Mitchell. *Medical Reference Services Quarterly*. 2017; 36: 207–209. <https://doi.org/10.1080/02763869.2017.1294004>
- Kandiuk M. The rhetoric of digitization and the politicization of Canadian heritage. *Library Trends*. 2016; 65: 165–179. <https://doi.org/10.1353/lib.2016.0029>
- Kong YX, Hu Y. Research on the construction of Dongba cultural heritage resources in Lijiang from the perspective of LAM cross-museum collaboration. *Shanxi Archives*. 2023; 2: 158–167 (In Chinese)
- Lázaro Ortiz S, Jiménez de Madariaga C. The UNESCO convention for the safeguarding of the intangible cultural heritage: a critical analysis. *International Journal of Cultural Policy*. 2021; 28: 327–341. <https://doi.org/10.1080/10286632.2021.1941914>
- Lee BCG. The “Collections as ML Data” checklist for machine learning and cultural heritage. *Journal of the Association for Information Science and Technology*. 2025; 76: 375–396. <https://doi.org/10.1002/asi.24765>
- Loach K, Rowley J. Cultural sustainability: A perspective from independent libraries in the United Kingdom and the United States. *Journal of Librarianship and Information Science*. 2021; 54: 80–94. <https://doi.org/10.1177/0961000621992824>
- Macri E, Cristofaro CL. The digitalisation of cultural heritage for sustainable Development: The Impact of Europeana. In Demartini P, Marchegiani L, Marchiori M, Schiuma G (eds.) *Cultural Initiatives for Sustainable Development. Contributions to Management Science* (pp. 373–400). Springer: Cham. 2021. https://doi.org/10.1007/978-3-030-65687-4_17
- Manovich L. *Cultural Analytics*. MIT Press: Cambridge. 2020.
- Marcum D. Archives, libraries, museums: Coming back together? *Information & Culture*. 2014; 49: 74–89. <https://doi.org/10.7560/ic49105>
- Marty PF. Digital convergence and the information profession in cultural heritage organizations: Reconciling internal and external demands. *Library Trends*. 2014; 62: 613–627. <https://doi.org/10.1353/lib.2014.0007>
- McDonald JD, Levine-Clark M. *Encyclopedia of Library and Information Sciences*. CRC Press: Boca Raton. 2017.
- Moid A, Raza MM, Javed M, Jahan K. Global research mapping in records and archives management: an investigation. *Global Knowledge Memory and Communication*. 2024. <https://doi.org/10.1108/gkmc-08-2023-0307> (online ahead of print)
- Pan M. The museum that leaps over the physical wall: Museum practices in Japan under the COVID-19 pandemic. *Journal of Urban Culture Research*. 2021; 23: 107–122.
- Qian Y, Xing Z, Shi X. From collection resources to intelligent data: Construction of intelligent digital humanities platform for local historical documents of Shanghai Jiao Tong University. *Digital Scholarship in the Humanities*. 2021; 36: 439–448. <https://doi.org/10.1093/lc/fqaa027>
- Rasmussen CH. Is digitalization the only driver of convergence? Theorizing relations between libraries, archives, and museums. *Journal of Documentation*. 2019; 75: 1258–1273. <https://doi.org/10.1108/jd-02-2019-0025>
- Rasmussen CH, Rydbeck K, Larsen H. Libraries, archives, and museums in transition: Changes, Challenges, and Convergence in a Scandinavian Perspective. Taylor & Francis: New York. 2022.
- Rayward WB. Electronic information and the functional integration of libraries, museums, and archives. In Higgs E (ed.) *History and electronic artefacts*. Routledge: London. 1998.
- Robinson H. Curating convergence: interpreting museum objects in integrated collecting institutions. *International Journal of Cultural Policy*. 2016; 24: 520–538. <https://doi.org/10.1080/10286632.2016.1218859>
- Ryholt K, Barjamovic G. *Libraries before Alexandria: Ancient Near Eastern Traditions*. Oxford University Press: Oxford, UK. 2019. <https://doi.org/10.1093/oso/9780199655359.001.0001>
- Samsuddin SF, Shaffril HA, Fauzi A. Heigh-ho, heigh-ho, to the rural libraries we go! - a systematic literature review. *Library & Information Science Research*. 2020; 42: 100997. <https://doi.org/10.1016/j.lisr.2019.100997>
- Sánchez-Franco MJ, Calvo-Mora A, Perriñez-Cristobal R. Clustering abstracts from the literature on Quality Management (1980–2020). *Total Quality Management & Business Excellence*. 2022; 34: 959–989. <https://doi.org/10.1080/14783363.2022.2139674>

- Schreibman S, Siemens R, Unsworth J. A new companion to digital humanities. John Wiley & Sons: Chichester. 2016.
- Schweibenz W. The virtual museum: An overview of its origins, concepts, and terminology. *The Museum Review*. 2019; 4: 1–29.
- Sebastián Lozano J, Alba Pagán E, Martínez Roig E, Gaitán Salvatella M, León Muñoz A, Sevilla Peris J, et al. Open access to data about silk heritage: A case study in digital information sustainability. *Sustainability*. 2023; 15: 14340. <https://doi.org/10.3390/su151914340>
- Stewart R, Simeonov S. Transformed into digital resources - a new life for the cultural heritage related research and scientific archives. *Digital Presentation and Preservation of Cultural and Scientific Heritage*. 2019; 9: 205–212. <https://doi.org/10.55630/dipp.2019.9.17>
- Su F, Li S, Liu Q, Zhang Y. A bibliometric study of digital humanities research in China from 2012 to 2021. *Information Research an International Electronic Journal*. 2023; 28: 62–82. <https://doi.org/10.47989/ir284336>
- Valtysson B, Kjellman U, Audunson R. The impact of digitalization on LAMs. In Rasmussen CH, Rydbeck K, Larsen H (eds.) *Libraries, archives, and museums in transition* (pp. 117–129). Taylor & Francis: New York. 2022. <https://doi.org/10.4324/9781003188834-12>
- Vårheim A, Skare R, Lenstra N. Institutional convergence in the LAM sector: a contribution towards a conceptual framework. *Information Research*. 2019; 24.
- Von Schorlemer S. UNESCO and the challenge of preserving the digital cultural heritage. *Santander Art and Culture Law Review*. 2020; 6: 33–64. <https://doi.org/10.4467/2450050xsnr.20.010.13013>
- Wang JJ, Ye Y. An exploratory analysis on the interdisciplinary knowledge diffusion in the international studies for digital humanities. *Journal of Academic Libraries*. 2021; 39: 45–51. <https://doi.org/10.16603/j.issn1002-1027.2021.02.006> (In Chinese)
- Wang X, Tan X, Li H. The evolution of digital humanities in China. *Library Trends*. 2020; 69: 7–29. <https://doi.org/10.1353/lib.2020.0029>
- Warren E, Matthews G. Public libraries, museums and physical convergence: Context, issues, opportunities: A literature review Part 1. *Journal of Librarianship and Information Science*. 2018a; 51: 1120–1133. <https://doi.org/10.1177/0961000618769720>
- Warren E, Matthews G. Public libraries, museums and physical convergence: Context, issues, opportunities: A literature review Part 2. *Journal of Librarianship and Information Science*. 2018b; 52: 54–66. <https://doi.org/10.1177/0961000618769721>
- Yang SL, Wu LJ. Research progress of foreign information resource management: An analysis based on the BERTopic model. *Information Studies: Theory & Application*. 2024; 42: 189–197. <https://doi.org/10.16353/j.cnki.1000-7490.2024.02.025> (In Chinese)
- Zahila MN, Noorhidawati A, Yanti Idaya Aspura MK. Content extraction of historical Malay manuscripts based on Event Ontology Framework. *Applied Ontology*. 2021; 16: 249–275. <https://doi.org/10.3233/ao-210247>
- Zeng ML. Semantic enrichment for enhancing LAM data and supporting digital humanities. Review article. *Profesional De La información*. 2019; 28. <https://doi.org/10.3145/epi.2019.ene.03>
- Zorich D, Waibel G, Erway R. Beyond the silos of the LAMs: Collaboration among libraries, archives and museums. *OCLC Programs and Research*. 2008. Available at: <https://www.oclc.org/content/dam/research/publications/library/2008/2008-05.pdf> (Accessed: 1 June 2025).