The Evolution of Digital Archive Preservation Topics from 2015 – 2025: A Bibliometric Analysis

[1]Putri Fajar Ayu Hendrayani, Tamara Adriani Salim[2], Muhamad Prabu Wibowo[3], Mad Khir Johari Abdullah Sani[4]

[1,2,3]Library and Information Sciences, Faculty of Humanities, Universitas Indonesia, Indonesia,

[4]Universiti Teknologi Mara, Malaysia

[1]putrifay30@gmail.com; [2]tamara.susetyo@gmail.com; [3]mprabuw@ui.ac.id; [4]madkhir@yahoo.com

ABSTRACT

Purpose Research. This study analyzes the evolution of scholarly discourse on digital archive preservation from 2015 to 2025 using a bibliometric approach. The objective is to identify thematic trends, conceptual linkages, and evolving dimensions in academic literature. **Research Method.** This study adopts a quantitative approach through bibliometric analysis. This approach is employed to identify and map the development of topics, research trends, and the interrelationships between relevant concepts within the field of digital archive preservation. **Analysis Data**. Fifteen peer-reviewed articles from the Scopus database were analyzed using VOSViewer software. The analysis was structured through a preservation framework comprising three core dimensions: organizational, technological, and functional, to classify and interpret each contribution systematically. **Results.** Results indicate a clear thematic shift from a primarily technical focus to a broader, interdisciplinary perspective incorporating collaborative, participatory, and contextual strategies. Co-occurrence analysis revealed three major thematic clusters: digital archives, digital preservation & cultural heritage, and digital humanities indicating increasing convergence across disciplines. **Conclusions.** Digital archive preservation has shifted from a technical focus to a socio-technical practice shaped by governance, community participation, and technological adaptability. Bibliometric findings reveal fragmented theoretical development and uneven practical implementation, especially in resource-constrained contexts. Addressing these gaps requires context-sensitive national frameworks, institutional roadmaps, and collaborative strategies to build resilient, inclusive preservation systems that safeguard cultural memory and strengthen long-term archival stewardship.

Keywords: Digital Preservation; Digital Archives; Bibliometric Analysis; Cultural Heritage.

A. INTRODUCTION

The rapid advancement of digital technology has fundamentally transformed how humans create, store, and access information. In this context, the issue of digital archive preservation is no longer merely a technical matter, but also encompasses complex technological, social, and institutional policy dimensions (Corujo et al., 2025). Digital archives possess characteristics that differ significantly from conventional archives. Unlike traditional formats, digital archives exist not

114

only as data that can be accessed at any time. However, they also depend on the continued functionality of specific hardware, software, and digital code structures (Barons et al., 2022). In other words, the existence of digital archives is highly vulnerable to technological changes, risks of data loss, as well as challenges related to authentication and sustainable access (Gainza et al., 2022; Johri et al., 2022).

Gainza et al. (2022) highlight that digital literature, as one form of digitally-based creative work, can only be accessed through the correct execution of digital code. Consequently, the continued existence of such works is entirely dependent on the supporting digital infrastructure. This situation underscores that digital preservation involves more than merely safeguarding bits and bytes; it also entails preserving the context, technological structures, and the performativity of code that constitute the digital entity itself (Shah et al., 2021). Thus, digital preservation is not merely a conservation practice, but a comprehensive effort to maintain the meaning and accessibility of information over the long term (El-Fakdi & de la Rosa, 2022).

Nevertheless, despite their widespread adoption, dominant approaches to digital preservation, such as the Open Archival Information System (OAIS) framework, exhibit both conceptual and operational limitations. Moles (2022) critiques the OAIS's foundational concept of the Designated Community, arguing that it is not always suitable for cultural and archival institutions that serve highly heterogeneous publics. When user communities cannot be narrowly and homogeneously defined, digital preservation practices risk becoming problematic, as they may exclude broader groups of users (Jaillant & Caputo, 2022).

Furthermore, Condotta (2015) emphasizes that digital preservation is fundamentally a leadership issue, requiring the integration of managerial, technical, and content-related dimensions. They critique the dominance of overly technocratic approaches to digital preservation and advocate for more reflective and strategic frameworks in designing long-term digital projects. Institutions that fail to account for organizational continuity, technical flexibility, and content validity from the outset significantly increase the risk of long-term preservation failure (Ahmad et al., 2025; Seemann et al., 2025).

Amid the evolving challenges and ongoing debates, it is crucial to understand how the dynamics of digital archive preservation topics are developing within contemporary scholarly literature (Jiang et al., 2025; Corrado, 2022). This study formally examines the evolution of digital archive preservation research from 2015 to 2025 using bibliometric analysis, with particular attention to thematic developments, interdisciplinary linkages into library, and the positioning of organizational, technological, and functional dimensions within the scholarly discourse. The findings are intended to provide a structured mapping of research trends and conceptual progressions that reflect the growing complexity of digital archive preservation in contemporary contexts.

In the digital era, archives have shifted from static repositories of physical records to dynamic traces of human activity embedded in networked environments. Rather than serving merely as electronic copies of documents, digital archives now include personal and institutional data circulating across platforms such as cloud storage, email, and social media (Modiba & Shekgola, 2024; Cannelli & Musso, 2022; Ernst, 2013). Academic discussions on digital archives focus on authenticity, sustainability, technological obsolescence, and the tension

between archives as technical systems and cultural constructs (Caswell, 2016). Key challenges include metadata standardization, interoperability, and curating volatile or short-lived digital records (Johnston, 2020).

The organizational dimension refers to institutional responsibility for establishing policies and procedures to manage large volumes of digital data, including storage planning, staff allocation, and risk governance (Bettington et al., 2008). Digitization efforts produce substantial data that demand dedicated storage systems and structured management. To ensure long-term access and prevent data loss, institutions must separate storage from image management systems and maintain systematic backup protocols, including off-site copies (Shekgola & Ngoepe, 2025; Makanjuola, 2020).

The technological environment includes hardware, software, and systems such as Collection Management Systems (CMS) that enable the sustainable storage and access of digital archives (Friedewald et al., 2024). Digitization produces images and metadata that require standardized, reliable systems to remain accessible (Mkadmi et al., 2025). Because storage media are prone to degradation and obsolescence, strategies like media refreshing, format standardization, and system migration are critical (Bettington et al., 2008). Functional requirements define the system capabilities needed to support the digitization workflow. This includes preparing archival materials through functions that record physical condition, conservation needs, and technical features like fold-outs or double-sided pages (Bettington et al., 2008). Systems must also accommodate managerial tasks such as scheduling, task allocation, and drafting technical guidelines. Whether digitization is conducted in-house, collaboratively, or outsourced, adaptable functionalities are required for documenting specifications and tracking quality control (Ahmad et al., 2025; Khan et al., 2023).
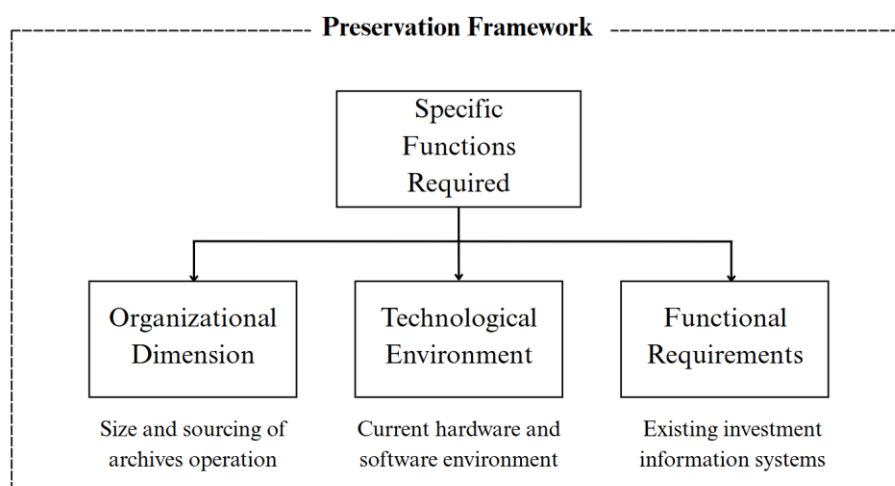
Figure 1. Preservation Framework
Source: Adaptation from (Roberts, D., 2008: 420)

B. METHODS

This study adopts a quantitative approach through bibliometric analysis. This approach is employed to identify and map the development of topics, research trends, and the interrelationships between relevant concepts within the field of digital archive preservation. VOSViewer is utilized as an analytical tool to support the visualization of the mapped bibliometric results. In addition, the researcher conducted a close reading and interpretation of the collected research articles to provide a more exploratory explanation of the evolving discourse on digital archive preservation within information institutions.

B.1 Research Questions

This study aims to address the central issue outlined in the title by exploring the urgency of the topic through a bibliometric analysis approach. The results of the mapping are then used to formulate three research questions that represent key dimensions in the evolution of digital archive preservation studies.

RQ1. How have the trends and thematic focuses in digital archive preservation research evolved during the period 2015–2025?

RQ2. How are the organizational dimension, technological environment, and functional requirements within the preservation framework used to map thematic focuses and trends in contributions to the digital archive preservation literature?

RQ3. How has the evolution of thematic focus and interdisciplinary connections in digital archive preservation research progressed from 2015–2025, based on bibliometric and co-occurrence network analysis?

RQ4. How do theoretical fragmentation and practical gaps in digital archive preservation emerge in the literature, and what limitations are revealed through bibliometric findings from 2015–2025?

B.2 Literature Source

The literature sources for this bibliometric analysis were drawn from the SCOPUS database. SCOPUS was selected due to its comprehensive coverage of high-quality research publications from around the world, thereby ensuring the credibility and reliability of the content for scholarly analysis. The selected studies are those published between 2015–2025, representing the most recent ten-year period.

B.3 Search Terms

The data retrieval process for the bibliometric analysis was conducted using relevant keywords combined with Boolean logic operators to systematically broaden or narrow the search results. In this study, the search strategy employed a combination of keywords and Boolean logic in the following query: ("preservation" AND "digital archive*" AND "organi*ation"). The asterisk (*)

117

serves as a wildcard to capture variations of terms, such as "organization" and "organisational," thereby expanding the scope of documents relevant to the research topic. This approach is intended to ensure that the retrieved articles align with the focus of the study and support a comprehensive bibliometric analysis.
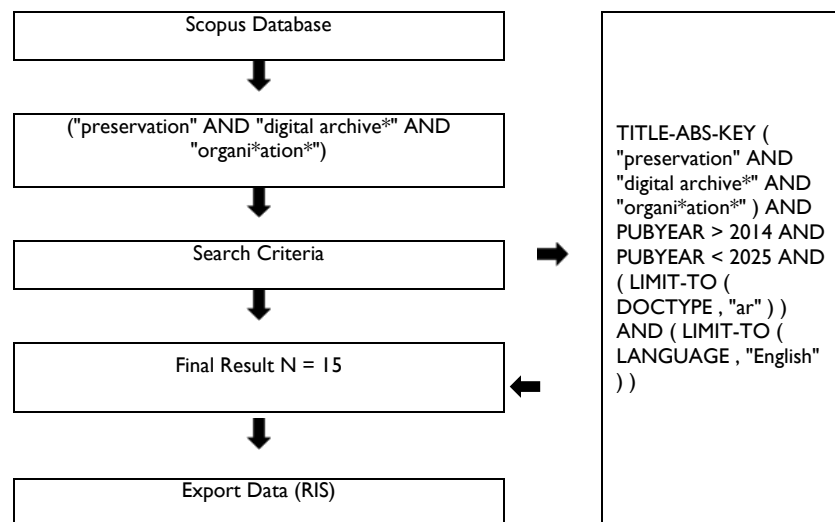
Figure 2. Search Strategy in Scopus Database

Source: Researcher Interpretation (2025)

B.4 Tools and Techniques

The analysis materials, consisting of a collection of articles retrieved from the SCOPUS database, were analyzed using VOSViewer software to visualize the distribution map of research developments. The settings applied in VOSViewer utilized co-occurrence analysis to illustrate the frequency of keywords that construct the research topic. Visualizations in the form of network and overlay maps were also generated.

B.5 Inclusion and Exclusion Criteria

The inclusion and exclusion criteria for selecting research articles were determined based on the relevance of content to the research questions addressed in this study. Articles that met the inclusion criteria were those published within the last ten years (2015–2025), written in English, article document type, and peer-reviewed scholarly articles. Conversely, articles that did not meet these criteria such as non-research articles, publications in languages other than English, or those published outside the defined time range were excluded from the analysis.

118

C. RESULTS AND DISCUSSION

C.1 RQ1: How have the trends and thematic focus in digital archive preservation research evolved during the period 2015–2025?

Between 2015 and 2025, the research theme on digital archive preservation, as seen in Table 1 from the most cited article, has undergone a marked thematic and methodological evolution, shifting from foundational infrastructural concerns to more community-rooted, technologically adaptive, and globally contextualized approaches. Early works emphasized institutional mandates and systemic transformations led by state bodies. For example, in the U.S. context, the Archivist's Directive served as a critical catalyst for rethinking the management, preservation, and access of public-sector email and electronic records, highlighting the urgency of addressing long-term digital stewardship and the structural gaps in lesser-resourced regions (Baron & Thurston, 2016). At the same time, theoretical interventions began interrogating the transformation of archival temporality and practice under digitization, with studies examining how digital archives reshape memory, activism, and historicity through forms of self-archiving and non-institutional documentation (Kaun, 2016).

Table 1. The Most Cited Articles

Code	Title	Author	Source	Year	Cite
A1	What lessons can be learned from the US archivist's digital mandate for 2019 and is there potential for applying them in lower resource countries?	Baron, J. R., & Thurston, A.	Records Management Journal, 26(2), 206–217	2016	19
A2	Archiving protest digitally: The temporal regime of immediation	Kaun, A.	International Journal of Communication, 10, 5395–5408	2016	11
A3	Digital storytelling and participatory local heritage through the creation of an online moving image archive: A case-study of Fraserburgh on film	Davidson, A., & Reid, P. H.	Journal of Documentation, 78(2), 389–415	2022	9
A4	Developing a sustainable cultural heritage information system	Thekkum Kara, G. K.	Library Hi Tech News, 38(6), 17–20	2021	5
A5	Facilitating collaborative metadata creation for faculty-initiated digital projects	Knight, R. C., Rodrigues, E., & Ciota, R.	Journal of Library Metadata, 20(1), 51–64	2020	5
A6	Identification, documentation and promotion of cultural heritage: Problems and prospects in the Indian context	Gireesh Kumar, T. K.	Journal of Cultural Heritage Management and Sustainable Development, 14(6), 950–962	2024	4
A7	Semantic relation extraction from cultural heritage archives	Buranasing, W., & Lilakiataskun, W.	Journal of Web Engineering, 21(4), 1081–1102	2022	3

A8	The Vatican School of Librarianship in a time of change: Current organization and future perspectives	Weston, P. G.	Journal of Education for Library and Information Science, 61(3), 319–331	2020	2
A9	Conceptualising methodological diversity among born-digital users: Insights from the garbage can model	Nix, A., Decker, S., & Kirsch, D. A.	AI and Society, 40(6), 4499–4511	2025	0
A10	Digital preservation of Ukrainian audiovisual heritage during wartime: Challenges and institutional practices	Kovalenko, Y.	Culture Crossroads, 27, 73–86	2025	0
A11	Digitising cultural heritage for community development: Lessons from Nigeria, South Africa and Kenya	Adewumi, A. A., & Sani, N. A.	International Review of Law, Computers and Technology. Advance online publication.	2025	0
A12	'A digital archive of Ireland's Ordnance Survey': Connecting collections for nineteenth-century Ireland	Griffith, L., Garrett, Z. S., Kenny, S., Lilley, K. D., & Porter, C.	Journal of Open Humanities Data, 11, 1–14	2025	0
A13	Digital preservation of Zhujiayu Village's cultural landscape: A new approach to sustainable development	Wang, T., Ghani, I., Ahmad, S., Rahman, N. A., Zhou, S., & Wahid, H.	Built Environment Journal, 21(S1), 245–259	2024	0
A14	A day for preserving things: Activism, technology and the politics of preservation in the Sarah Records archive	Evans, A.	Punk and Post-Punk, 13(3), 319–331	2024	0
A15	Personal digital archives: State of the art guidelines in North America, Australasia and Europe	Gorini, A.	Atlanti+, 33(2), 59–82	2024	0

Source: SCOPUS and VOSViewer (2025)

As the decade progressed, attention expanded to include digital humanities infrastructures, particularly questions of metadata quality, collaborative authorship, and tool-based workflows that ensure both usability and sustained preservation of project outputs (Knight et al., 2020). Educational institutions similarly adapted their training models to accommodate emerging preservation competencies, integrating practical modules on metadata, data management, digital libraries, and archival digitization alongside field-based internships to align curricular design with professional demands (Weston, 2020).

From 2019 onward, the research focus widened to include regional disparities and heritage-oriented digital systems, with research topics stressing the absence of unified information frameworks in countries like India and the need for integrated digital archives that support global visibility and systematic site management using ICT and visualization technologies (Thekkum Kara, 2021). Thematic trends during this period also reflected the growing influence of participatory and community-based frameworks. Studies explored the role of digital storytelling platforms, remastering practices, metadata standards, and dissemination strategies in preserving audiovisual heritage while embedding local voices and fostering grassroots engagement (Davidson & Reid, 2022).

120

Parallel work investigated technological innovations in data extraction and knowledge graph development to address linguistic complexity in digital cultural heritage texts, demonstrating how computational techniques such as unseen word mapping and multiple instance learning enhance information accuracy and retrieval (Buranasing & Lilakiataskun, 2022). By the early 2020s, calls for collaboration, documentation standards, and inclusive digital gateways intensified, particularly for organizations and communities lacking archival infrastructure or clarity in preservation practices; the research emphasized the necessity of international partnerships, standardized documentation, and web-based platforms for collective knowledge sharing (Gireesh Kumar, 2024).

Research in 2024 further illuminated persistent challenges around intangible cultural heritage, noting the fragility of practitioner-based memory systems and advocating for digitization to offset gaps in documentation and ensure long-term transmission (Wang et al., 2024). Other contributions examined how digital archival frameworks preserve political or cultural identities, as in the case of independent music archives that leverage digitization to sustain activist legacies, connect past and present, and subvert dominant narratives (Evans, 2024). Global inequities in access, literacy, and guidance for personal digital archiving also became more visible, prompting calls for stronger institutional support to disseminate knowledge and improve digital preservation competencies (Gorini, 2024).

From 2024 to 2025, thematic developments increasingly reflected geopolitical realities and the integration of AI. Studies on Ukraine underscore how digital technologies safeguard cultural identity during conflict, preserve historical memory, and facilitate international cooperation through standards-based frameworks (Kovalenko, 2025). Concurrently, African analyses critiqued existing heritage governance and promoted digitization, virtual museums, and policy reforms as socio-economic and cultural sustainability engines, stressing multi-stakeholder engagement and alignment with successful continental precedents (Adewumi & Sani, 2025). Large-scale collaborative projects, such as the Irish digital remapping of Ordnance Survey materials, illustrated the maturation of cross-border initiatives and the consolidation of archival materials into unified, publicly accessible repositories (Griffith et al., 2025). Finally, emerging discussions on AI-based archival tools acknowledged the organizational complexity of born-digital access. They introduced user typologies to reimagine tool development and stakeholder involvement amid decentralized decision-making (Nix et al., 2025).

C.2 RQ2: How are the organizational dimension, technological environment, and functional requirements within a preservation framework used to map thematic focuses and contribution trends in the literature on digital archive preservation?

To gain a comprehensive and structured understanding of the evolving dynamics within the literature on digital archive preservation, this analysis refers to a preservation framework developed in prior studies. As illustrated in Table 2, the framework consists of three key dimensions: organizational dimension, technological environment, and functional requirement. These three dimensions represent fundamental aspects of digital preservation practice and theory and

serve as the basis for classifying, evaluating, and mapping the contributions of each article analyzed. The classification results based on these three dimensions are systematically presented in Table 2, offering a more comprehensive overview of thematic focuses and emerging trends in the relevant literature.

Table 2. Literature Result of Preservation Framework

Specific Functions Required	Activity Description	Article Code
1) Organizational Dimension	Digital archive preservation involves selecting valuable records, securing them with metadata, converting formats, checking integrity, and ensuring long-term access	A1
	Community participation in the development of digital repositories and the preservation of local audiovisual heritage	A3
	Development of a national cultural information system based on collaboration and public access	A4
	AI tools are integrated into archival processes, tailored to different user needs, and shaped by organizational decision-making models	A9
2) Technological Environment	Records are preserved through appraisal, secure storage, format migration, integrity checks, and policy-driven management	A2
	Use of Google Forms for collaborative metadata creation in digital archives	A5
	Application of machine learning for Thai language data extraction in cultural heritage archives	A7
	Digital technologies and standards in preserving audiovisual heritage during wartime, focusing on digitization practices, innovative tools, and international collaboration	A10
	Digitisation as the central strategy for reforming heritage management	A11
	Building a digital archive by integrating diverse historic maps and texts, managing metadata, copyright, licensing, and spatial data challenges	A12
	The application of digital technologies (3D scanning, GIS, VR) for preserving architecture, landscapes, and intangible heritage	A13
3) Functional Requirement	Participatory digital documentation for Indian cultural heritage and the development of a community-based digital reporting system	A6
	Professional education in digital preservation, including technical training and soft skills development in library settings	A8
	Use of digital archives as a medium for reconstructing historical narratives and political activism in independent music communities	A4
	Analysis of global disparities in access to personal digital archive management guidelines and the critical role of public institutions	A15

Soruce: Research Data Interpretation (2025)

1) Organizational Dimension in the Functional Development of Digital Archive Preservation Literature

The organizational dimension within digital preservation frameworks provides a crucial lens for tracing thematic emphases and contribution trends in the literature, particularly through its focus on governance, institutional mandates, collaboration models, and technological decision-making. Early research highlights how top-down directives act as structural engines for transformation. In the

122

United States, for example, the Archivist's Directive demonstrates how federal policy compels agencies to address long-term preservation and access for email and other electronic records, while simultaneously exposing disparities in implementation capacity in less-resourced regions (Baron & Thurston, 2016). Such research positions institutional readiness, regulatory enforcement, and archival infrastructure as central thematic concerns. As the field evolves, research moves beyond institutional compliance to include frameworks grounded in participatory engagement. Davidson and Reid (2022) illustrate how preservation initiatives involving audiovisual cultural heritage rely on close collaboration with community stakeholders, where storytelling platforms, metadata practices, and accessible dissemination strategies reshape organizational roles and responsibilities in digital archiving ecosystems.

Parallel developments in emerging contexts further reveal how organizational fragmentation influences preservation outcomes. Thekkum Kara (2021) underscores the absence of integrated information systems for UNESCO heritage sites in India, demonstrating how gaps in governance and coordination limit the potential of digital archives to achieve global visibility and sustainable management. The research emphasizes that organizational dimensions extend beyond bureaucratic design, including strategic integration of ICT, mobile technologies, and visualization systems. Toward the late 2020s, thematic attention shifts toward decentralized and adaptive models shaped by artificial intelligence and multi-stakeholder participation. Nix et al. (2025) apply the Garbage Can Model to conceptualize the iterative and often non-linear decision-making in developing AI-based tools for born-digital archives, identifying differentiated user types and methodological needs as organizational drivers.

2) *Technological Environment in the Functional Development of Digital Archive Preservation Literature*

The technological environment functions as a central axis in shaping thematic priorities and research contributions within digital archive preservation frameworks. Early discussions signal a conceptual shift as digitization reconfigures archival temporality, authorship, and control. Kaun (2016) traces how protest movements transition from being mere subjects of institutional archiving to active agents in self-archiving, reflecting a thematic movement toward user-driven digital repositories and dynamic temporal regimes. This reorientation highlights technology as a preservation tool and an enabler of new archival politics, where immediacy, access, and autonomy emerge as defining concerns. As digital archives expand in scope, the literature identifies metadata as a cornerstone of long-term preservation, particularly in collaborative and interdisciplinary contexts. Knight et al. (2020) illustrate how tool configuration and shared digital workflows, such as those using Google Forms, become crucial to sustaining accessibility and reliability in digital humanities projects.

Advancements in computational methods mark a second wave of thematic focus, emphasizing the challenges of linguistic complexity, data extraction, and knowledge structuring in cultural heritage archives. Buranasing and Lilakiataskun (2022) demonstrate how word-feature modeling, unseen word mapping, and multiple instance learning improve the extraction of digital text in languages with

complex grammar, such as Thai. Their findings indicate a broader trend: preservation frameworks increasingly depend on intelligent technologies to bridge gaps in data interpretation, interoperability, and multilingual accessibility. Concurrently, the literature shows how digital preservation becomes a strategic instrument in volatile geopolitical environments. Kovalenko (2025) highlights Ukraine's reliance on digital technologies to safeguard cultural identity, document historical events, and maintain global connectivity during conflict. This reflects an emergent thematic linkage between technological resilience, humanitarian awareness, and international collaboration in preservation research.

Thematic trajectories evolve through large-scale, multi-institutional projects and call for digitally integrated policy reform. Adewumi and Sani (2025) advocate for embedding virtual museums, digital archives, and educational platforms into Nigeria's heritage governance, emphasizing the role of digitization in inclusive development and socio-economic transformation. Similarly, the Digital Archive of Ireland's Ordnance Survey project (Griffith et al., 2025) demonstrates how preservation technologies facilitate cross-border collaboration, standardization, and public access to dispersed historical materials. However, technological adoption remains uneven, as Wang et al. (2024) observe in the context of intangible heritage preservation, where knowledge persists primarily through oral transmission and personal memory despite rising awareness of digital solutions.

Functional Requirements in the Functional Development of Digital Archive Preservation Literature

Functional requirements in digital preservation frameworks serve as a key lens for mapping thematic directions and research contributions in the literature, particularly through their emphasis on documentation standards, accessibility, and community engagement. Gireesh Kumar (2024) identifies a major thematic concern: the lack of standardized documentation practices and functional clarity among organizations and individuals working with cultural heritage. This gap produces uneven preservation outcomes and underscores the need for frameworks that specify not just tools, but functional protocols for documentation, sustainability, and stakeholder participation. The call for a centralized, web-based gateway reflects a growing recognition that digital preservation requires functional systems that integrate community knowledge, technical resources, and international collaboration. These requirements signal a thematic shift toward participatory, technology-supported environments where both local actors and institutional stakeholders contribute to long-term preservation mechanisms.

A complementary thematic development emerges around functional competencies, workforce readiness, and technological literacy. Weston (2020) illustrates how preservation-focused curricula increasingly incorporate functional requirements such as cataloguing, metadata management, digital repositories, and information literacy, linking academic training to evolving sector demands. Practical workshops and internships further demonstrate how functional adaptability and soft skills are being integrated into preservation education to align with emerging professional standards. Parallel research extends these themes into questions of access, politics, and equity. Evans (2024) shows how functional practices allow digital archives to embody political values and cultural advocacy. Meanwhile, Gorini (2024) highlights disparities in digital preservation literacy and frames public

124

institutions as essential to enabling functional access and standardized guidance. Together, these contributions reveal how functional requirements not only define preservation workflows but also shape thematic trajectories around inclusivity, political agency, professionalization, and the social role of digital archives.

C.3 RQ3: How has the thematic focus and interdisciplinary linkage in digital archive preservation research evolved from 2015–2025 based on bibliometric and co-occurrence network analysis?

In this study, a bibliometric analysis was conducted using a co-occurrence approach to identify the main emerging themes in the evolving discourse on digital archive preservation. To enhance the interpretation of findings, network and overlay visualizations were employed, utilizing VOSViewer software as the primary tool for bibliometric mapping.

Figure 3. Network Visualization of Digital Archive Preservation
Source: Research Data Generated by VOSViewer (2025)

The bibliometric map reveals a fragmented thematic structure in the literature, with two distinct clusters emerging from the visualization presented in Figure 3 and explained in Table 3. Cluster 1 consists of the closely linked terms *"cultural heritage"* and *"digital preservation,"* indicating a strong conceptual and research-based association between safeguarding heritage assets and developing digital strategies to maintain them over time. The proximity and linkage suggest frequent co-occurrence, implying that scholarly discourse in this area prioritizes the preservation of cultural identity, memory, and heritage through digital means. This cluster reflects an interdisciplinary convergence between heritage studies, conservation practices, and digital technologies. Its prominence also indicates that digital preservation is not being treated as a purely technical exercise but as a critical tool for sustaining tangible and intangible heritage across generations.

Table 3. Co-Occurance Analysis of the Digital Archive Preservation Research Topic

Cluster Number	Color Indication	Keyword	Connection	Interpretation
1	Red-orange	Cultural Heritage	Linked to "digital preservation"	Most recent/ highly active term
1	Red-orange	Digital Preservation	Linked to "cultural heritage"	Emering/ content focus area
2	Green	Digital Archives	No connection to other keywords	Earlier/ separate thematic focus

Source: Research Data Interpretation (2025)

In contrast, Cluster 2, represented solely by the term "digital archives," is positioned separately with no visible linkage to the other terms. This isolation suggests a conceptual or topical disconnect between research addressing digital archives and work focused on digital preservation of cultural heritage. The separation may indicate that research on digital archives is evolving as an independent research trajectory, potentially emphasizing areas such as metadata management, platform design, digital access, or institutional records management. The spatial distance between the clusters also highlights an opportunity for integration. Despite the inherent relationship between digital archives and digital preservation practices, the current literature treats them with limited overlap. Strengthening cross-disciplinary connections between archival science, heritage preservation, and digital documentation frameworks may help bridge this gap and foster a more cohesive research landscape in the coming years.
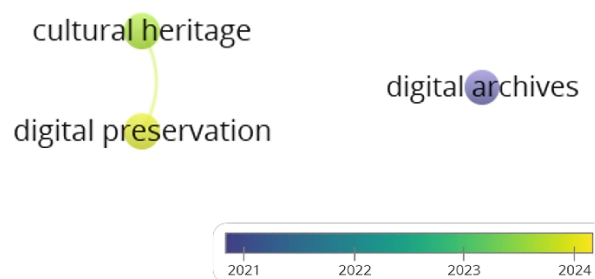
Figure 4. Overlay Visualization of Digital Archive Preservation
Source: Data Result by VOSViewer (2025)

The temporal overlay further clarifies the developmental trajectory of research themes reflected in the network. The terms *"cultural heritage"* and *"digital preservation"* appear in lighter yellow-green tones, indicating that these themes have gained prominence in more recent years, particularly around 2023–2024. This suggests an increasing focus on integrating preservation strategies with heritage-based contexts, likely driven by global cultural sustainability agendas, digitization initiatives, and the growing urgency to protect both tangible and intangible heritage materials. In contrast, *"digital archives"* appears in a darker blue-purple tone, suggesting it represents an earlier or less contemporarily active line of inquiry, with prominence closer to 2021–2022. The temporal gap between clusters reinforces the notion that earlier research on digital archives may not have fully transitioned into or connected with the more current discourse on preservation and cultural heritage.

C.4 RQ4: How do theoretical fragmentation and practical gaps in digital archive preservation emerge in the literature, and what limitations are revealed through bibliometric findings from 2015–2025?

The bibliometric findings reveal a fragmented thematic landscape that carries important theoretical and practical implications for the library and information science field. From a theoretical standpoint, the strong conceptual pairing between digital preservation and cultural heritage suggests that current research treats preservation as a technological problem and an epistemic and socio-cultural

126

construct. This aligns with post-custodial and participatory archival theories, yet the lack of visible integration with "digital archives" as a distinct concept points to an unresolved theoretical gap. Digital archives, as a foundational domain of archival science, appear to be treated as an earlier or isolated discourse rather than being embedded in evolving debates on heritage-centered preservation. This conceptual disjunction risks reinforcing siloed understandings of preservation theory, where infrastructure, access, and memory politics are detached instead of articulated within a cohesive framework.

The findings highlight progress and challenges for libraries as key stakeholders in digital preservation ecosystems. The clustering of "digital preservation" with "cultural heritage" implies that libraries are increasingly expected to fulfill roles beyond custodianship, facilitating community engagement, integrating storytelling practices, and preserving tangible and intangible cultural assets. However, the separate positioning of "digital archives" indicates that many library practices may still rely on legacy models focused on storage, cataloging, and institutional mandates, rather than adaptive, integrative, and collaborative approaches. Libraries must therefore reevaluate their digital strategies to bridge heritage-based preservation with archival infrastructures, particularly in metadata standardization, AI-based retrieval, education, and user-centered access. Without this alignment, libraries risk operating in outdated paradigms that do not fully address contemporary preservation needs or the expectations of diverse knowledge communities.

An explicit limitation of the current bibliometric mapping lies in the narrow cluster composition and minimal keyword diversity. With only three co-occurring terms and two clusters, the analysis likely underrepresents the thematic complexity of research conducted across regions, methodologies, or interdisciplinary collaborations. The visualization also does not capture sub-themes such as AI integration, crisis archiving, legal frameworks, personal digital collections, or digital humanities, all of which emerged in the literature review. Additionally, the lack of connection between clusters may partly result from dataset constraints, indexing limitations, or VOSViewer's sensitivity to keyword frequency and threshold settings. Future analyses should incorporate extended keyword sets, multi-database sources, and longitudinal co-word mapping to provide a more accurate and interconnected representation of the field.

D. CONCLUSION

The bibliometric results clearly demonstrate a fragmented theoretical landscape in digital archive preservation, where concepts such as cultural heritage, digital preservation, and digital archives appear in isolated clusters with limited conceptual and practical integration. This fragmentation reflects a broader disconnect between scholarly discourse and the operational realities of archival institutions, particularly in regions with limited resources. The lack of cohesive theoretical development contributes to practical gaps in implementation, revealing that preservation is often perceived narrowly as a technical function rather than a socio-technical and institutional responsibility.

The extended analysis further confirms that digital archive preservation is undergoing a paradigmatic shift, from a storage-oriented paradigm to one that recognizes preservation as a dynamic system shaped by governance structures,

participatory models, and technological adaptability. However, this shift is not unfolding evenly across contexts. Many institutions, especially in developing settings, struggle with environmental standards, infrastructural limitations, and shortages of skilled personnel. These challenges exacerbate the distance between global best practices and local capabilities, reinforcing the limitations identified in the bibliometric mapping.

Addressing these gaps requires a strategic reorientation that bridges theory and practice. National preservation frameworks must be grounded in context-sensitive policies that support sustainable technologies, environmental monitoring, and workforce development. Institutional roadmaps should prioritize cross-sector collaboration, participatory governance, and evidence-based planning rooted in bibliometric insights. Without such coordinated efforts, archives' material integrity and role in safeguarding cultural memory will remain at risk. Ultimately, overcoming theoretical fragmentation and practical constraints is essential for constructing resilient, inclusive, and future-ready digital preservation ecosystems capable of sustaining national identity and institutional continuity across generations.

REFERENCES

Adewumi, A. A., & Sani, N. A. (2025). Digitising cultural heritage for community development: Lessons from Nigeria, South Africa and Kenya. *International Review of Law, Computers and Technology.* Advance online publication. https://doi.org/10.1080/13600869.2025.2506917

Ahmad, R., Rafiq, M., Khalil, M. F., & Khalil, M. H. (2025). *Bridging managerial and technical competencies in digital preservation: A systematic assessment of essential skills and training needs. Journal of Librarianship and Information Science.* Advance online publication. https://doi.org/10.1177/09610006251349831

Baron, J. R., & Thurston, A. (2016). What lessons can be learned from the US archivist's digital mandate for 2019 and is there potential for applying them in lower resource countries? *Records Management Journal, 26*(2), 206–217. https://doi.org/10.1108/RMJ-04-2015-0015

Barons, M. J., Fonseca, T. C. O., Merwood, H., & Underdown, D. H. (2022). Safeguarding the nation's digital memory: Bayesian network modelling of digital preservation risks. In *Mathematics in Industry* (Vol. 39, pp. 501–508). Springer Medizin. https://doi.org/10.1007/978-3-031-11818-0_65

Bettington, J., Eberhard, K., Loo, R., & Smith, C. (2008). *Keeping archives* (3rd ed.). Australian Society of Archivists.

Buranasing, W., & Lilakiataskun, W. (2022). Semantic relation extraction from cultural heritage archives. *Journal of Web Engineering, 21*(4). https://doi.org/10.13052/jwe1540-9589.2145

Cannelli, B., & Musso, M. (2022). Social media as part of personal digital archives: Exploring users' practices and service providers' policies regarding the preservation of digital memories. *Archival Science, 22*(2), 259–283. https://doi.org/10.1007/s10502-021-09379-8

Caswell, M. (2016). "The archive" is not an archives: Acknowledging the intellectual contributions of archival studies. *Reconstruction: Studies in Contemporary Culture, 16*(1). https://escholarship.org/uc/item/7bn4v1fk

128

Condotta, K. (2015). Digital preservation for libraries, archives, & museums. *Collections: A Journal for Museum and Archives Professionals, 11*(3), 257–258. https://doi.org/10.1177/155019061501100310

Corujo, L., Revez, J., da Silva, C. G., & de Macedo, L. S. A. (2025). Preservation and digital repositories: Connections, possibilities, and needs. In *Handbook of trends and innovations concerning library and information science: A multidisciplinary approach* (pp. 111–134). De Gruyter. https://doi.org/10.1515/9783111443003-005

Corrado, E. M. (2022). Digital preservation is not just a technology problem. *Technical Services Quarterly, 39*(2), 143–151. https://doi.org/10.1080/07317131.2022.2045432

Davidson, A., & Reid, P. H. (2022). Digital storytelling and participatory local heritage through the creation of an online moving image archive: A case-study of Fraserburgh on film. *Journal of Documentation, 78*(2). https://doi.org/10.1108/JD-09-2020-0157

El-Fakdi, A., & de la Rosa, J. L. (2022). Evaluating auction mechanisms for the preservation of cost-aware digital objects under constrained digital preservation budgets. *Mathematics, 10*(1), Article 92. https://doi.org/10.3390/math10010092

Ernst, W. (2013). *Digital memory and the archive* (J. Parikka, Ed.).

Evans, A. (2024). A day for preserving things: Activism, technology and the politics of preservation in the Sarah Records archive. *Punk and Post-Punk, 13*(3), 319–331. https://doi.org/10.1386/punk_00252_1

Friedewald, M., Székely, I., & Karaboga, M. (2024). Preserving the past, enabling the future: Assessing the European policy on access to archives in the digital age. *Preservation, Digital Technology and Culture, 53*(2), 61–71. https://doi.org/10.1515/pdtc-2024-0003

Gainza, C., Zúñiga, C., & González, J. (2022). Digital archive and preservation against technological obsolescence: Building a cartography of Latin American digital literature. *Journal of Latin American Cultural Studies, 31*(2), 257–273. https://doi.org/10.1080/13569325.2022.2101439

Gireesh Kumar, T. K. (2022). Identification, documentation and promotion of cultural heritage: Problems and prospects in the Indian context. *Journal of Cultural Heritage Management and Sustainable Development.* https://doi.org/10.1108/JCHMSD-03-2022-0043

Gorini, A. (2024). Personal digital archives: State of the art guidelines in North America, Australasia and Europe. *Atlanti+, 33*(2), 59–92.

Griffith, L., Garrett, Z. S., Kenny, S., Lilley, K. D., & Porter, C. (2025). 'A digital archive of Ireland's Ordnance Survey': Connecting collections for nineteenth-century Ireland. *Journal of Open Humanities Data, 11,* 1–14. https://doi.org/10.5334/johd.288

Jaillant, L., & Caputo, A. (2022). Unlocking digital archives: Cross-disciplinary perspectives on AI and born-digital data. *AI & Society, 37*(3), 823–835. https://doi.org/10.1007/s00146-021-01367-x

Jiang, L., Li, J., Wider, W., Tanucan, J. C. M., Lobo, J., Fauzi, M. A., Hidayat, H., & Zou, R. (2025). A bibliometric insight into immersive technologies for cultural heritage preservation. *npj Heritage Science, 13*(1), Article 126. https://doi.org/10.1038/s40494-025-01704-z

Johri, S., Mehta, K., Suhashini, J., Shukla, P. K., Podile, V., & Singh, D. P. (2022). The impact of blockchain in accounting and auditing domain: A critical approach for enhanced efficiency and transparency. In *Proceedings of the 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE 2022)* (pp. 1628–1632). IEEE. https://doi.org/10.1109/ICACITE53722.2022.9823420

Johnston, L. (2020). Challenges in preservation and archiving digital materials. *Information Services & Use, 40*(3), 193–199. https://doi.org/10.3233/ISU-200090

Kaun, A. (2016). Archiving protest digitally: The temporal regime of immediation. *International Journal of Communication, 10*, 5395–5408.

Khan, Z., Lee, C., Girinathan, S., Fareed, M. M. U., Ferreira, S., Al Saeedi, F., Bishnoi, S., Al Maareef, A., Khwaja, N., & Bamakhramah, A. (2023). Operational excellence through implementation of data governance policies and procedures covering data retention, disposal and archival. In *Proceedings of the 2023 Gas and Oil Technology Showcase and Conference (GOTS 2023).* Society of Petroleum Engineers. https://doi.org/10.2118/214226-MS

Knight, R. C., Rodrigues, E., & Ciota, R. (2020). Facilitating collaborative metadata creation for faculty-initiated digital projects. *Journal of Library Metadata, 20*(1). https://doi.org/10.1080/19386389.2020.1728479

Kovalenko, Y. (2025). Digital preservation of Ukrainian audiovisual heritage during wartime: Challenges and institutional practices. *Culture Crossroads, 27,* 73–86. https://doi.org/10.55877/cc.vol27.534

Makanjuola, O., Meroney, P., & Gray, M. (2020). The digital transformation journey from digitization to opportunity generation. In *Society of Petroleum Engineers – Abu Dhabi International Petroleum Exhibition and Conference 2020 (ADIP 2020).* Society of Petroleum Engineers.

Mkadmi, A., Hamad, F., & Al-Yaarabi, S. (2025). Accessibility of archives for people with disabilities in Oman: Current state and challenges. *Archival Science, 25*(2), Article 19. https://doi.org/10.1007/s10502-025-09487-9

Moles, N. (2022). Preservation for diverse users: Digital preservation and the "designated community" at the Ontario Jewish Archives. *Journal of Documentation, 78*(3), 613–630. https://doi.org/10.1108/JD-02-2021-0041

Modiba, M., & Shekgola, M. (2024). Artificial intelligence embedded cloud computing technology for the management of digital archives in the Fifth Industrial Revolution in South Africa. *African Journal of Library Archives and Information Science, 34*(2), 229–239. https://doi.org/10.4314/ajlais.v34i2.6

Nix, A., Decker, S., & Kirsch, D. A. (2025). Conceptualising methodological diversity among born-digital users: Insights from the garbage can model. *AI and Society, 40*(6), 4499–4511. https://doi.org/10.1007/s00146-025-02229-6

Seemann, T., Fritzsche, F., Junginger, I., Buck, M., Flik, T., Laufer, D., & Seemann, M. (2025). The impact of leadership mindset on employee behavior in the context of digital transformation. *Change Management, 25*(1), 45–64. https://doi.org/10.18848/2327-798X/CGP/v25i01/45-64

Shah, U. A., Husain, M., Saddiqa, S., Yar, M. S., Hussain, M., & Saddiqa, M. (2021). Problems and challenges in the preservation of digital contents: An analytical study. *Library Philosophy and Practice, 2021*, 1–12.

130

Shekgola, M. M., & Ngoepe, M. (2025). Ingesting digital archives into long-term storage systems through free open-source software in South Africa. *Collection and Curation, 44*(1), 25–33. https://doi.org/10.1108/CC-02-2024-0003

Thekkum Kara, G. K. (2021). Developing a sustainable cultural heritage information system. *Library Hi Tech News, 38*(6). https://doi.org/10.1108/LHTN-08-2021-0053

Wang, T., Ghani, I., Ahmad, S., Rahman, N. A., Zhou, S., & Wahid, H. (2024). Digital preservation of Zhujiayu Village's cultural landscape: A new approach to sustainable development. *Built Environment Journal, 21*(S1), 245–259. https://doi.org/10.24191/bej.v21iS1.1560

Weston, P. G. (2020). The Vatican School of Librarianship in a time of change: Current organization and future perspectives. *Journal of Education for Library and Information Science, 61*(3). https://doi.org/10.3138/JELIS.61.3.2020-0016