DIGITAL HERITAGE (2025)
S. Campana, D. Ferdani, H. Graf, G. Guidi, Z. Hegarty, S. Pescarin, and F. Remondino (Editors)

Determining a General CH Digitization Process Based on a Scoping Review

C. Kröber[https://orcid.org/0009−0001−0494−9913] and S. Münster[https://orcid.org/0000-0001-9344-912X]

J. Professorship for Digital Humanities (Object/Image), FSU Jena, Germany

Abstract

The digitization of cultural heritage (CH) serves a range of critical purposes, including preservation, accessibility, research, and education. It also plays a key role in raising public awareness about the value of cultural assets, while fostering community engagement through thoughtfully developed digital heritage initiatives. This paper presents a qualitative analysis based on a scoping review of recent CH digitization literature, aiming to identify and synthesize key themes and approaches. The findings are clustered and abstracted to outline core steps in the digitization process, forming a structured framework. For each step, associated tasks and guiding questions are defined to support implementation. This process-oriented overview offers practical value for the planning and execution of new digitization projects e.g, strategic guidance on resource allocation and stakeholder engagement, including timing, methods, and rationale for involving communities, tourism sectors, and other relevant actors.

CCS Concepts

• ***General and reference*** → *Surveys and overviews;* • ***Applied computing*** → *Digital libraries and archives;*

1. Introduction

The digitization of cultural heritage (CH) plays a critical role in preserving the collective memory of humanity. By converting physical artefacts, documents, and intangible heritage into digital form, societies can safeguard their cultural assets against loss, decay, and natural disasters, while also enabling broader access for research, education, and public engagement. In an increasingly digital world, the importance of such efforts has grown, offering new opportunities for interaction with heritage and fostering cross-cultural understanding on a global scale.

CH institutions e.g., museums, archives, and libraries often possess expertise, infrastructure, and strategic vision necessary to implement digitization projects, but the rapid technological developments and discoveries can be challenging. A growing interest can be witnessed among smaller communities, local history clubs, and non-profit organizations to undertake similar initiatives. Their motivations range from ensuring the safekeeping of locally significant artefacts to supporting educational initiatives or contributing to the digital reconstruction of lost or damaged heritage. Despite their enthusiasm, these groups often face significant challenges due to limited resources, lack of experience, and the inherently complex nature of the digitization task.

Digitizing CH is far more than a technical operation. It requires thoughtful planning and coordination across multiple domains, including object selection, metadata creation, rights management, technological standards, and long-term preservation strategies. The process must accommodate diverse perspectives and priorities.

To support both experienced and novice stakeholders, there is a clear need for a structure that can deconstruct the digitization process into distinct, manageable stages, phases or steps. Such a breakdown would ideally present tasks with guiding questions or assignments, each addressing specific issues relevant at that point in the workflow. By clarifying what needs to be done, when, and why, this structured guidance could help ensure that digitization efforts are not only technically successful but also meaningful and sustainable in the long term.

This study aims to contribute to that goal by systematically analyzing existing scientific literature on CH digitization. Through this analysis, it seeks to identify common patterns, best practices, and essential components that collectively define the digitization process. The resulting insights will support the development of a generalized, adaptable structure suitable for a wide range of digitization initiatives, from large-scale institutional projects to small community-led efforts.

DOI: 10.2312/dh.20253369

2. Research context

CH digitization is a complex field encompassing various approaches and interpretations. Rapp [Rap2021] outlines three complementary perspectives: the conversion of analog objects into digital surrogates; the transformation of content into structured, machine-readable data; and the methodological spectrum involved, ranging from traditional hermeneutic practices to contemporary artificial intelligence applications.

While digitization projects often document specific technical workflows, these are typically shaped by the project's scope, tools, and domain, offering limited insight into a more abstract, transferable digitization process. In reviewing existing literature, it becomes evident that while many CH digitization efforts are structured around recurring phases like data acquisition, post-processing, and visualization, these reflect project-specific workflows rather than a generalized, high-level process. The distinction is important: a process, in this context, refers to a conceptual sequence of steps necessary to achieve digitization goals, independent of specific technologies or roles. Workflows, by contrast, detail the practical execution, tools, and task allocation within a particular project.

Despite growing interest in digitization across different CH domains, no widely accepted conceptual model exists that abstracts and generalizes these efforts into a cross-cutting process. Educational efforts e.g., often break down digitization into tasks to introduce students to domain-specific challenges [DD2009], but do not offer a full process framework. Likewise, procedural models for document digitization e.g., by Bandi et al. [BAS2015], or Chisita et al. [CDN2021] outline stepwise sequences like selection, scanning, OCR, metadata, and storage, yet remain tailored to document and library contexts. Variations in step ordering, as observed in Koganuramath and Angadi [KA2010], further illustrate the lack of standardization.

More recent contributions have started to expand the scope. Casey [Cas2022] introduces preservation principles for media digitization, some of which align with broader process stages. The International Federation of Library Associations and Institutions (IFLA) [IFL2015] offers a set of guiding questions for manuscript digitization that are useful for structuring a general framework. Importantly, stakeholder involvement and collaboration have also been recognized as critical components. Aparac-Jelušić [Apa2017] emphasizes the need for clearly timed, designed, and communicated engagement strategies especially relevant for inclusive digitization practices.

The present paper is the first to derive a generalized CH digitization process based on a scoping review across diverse domains. By analyzing published workflows, tools, and activities from various projects and initiatives, it aims to abstract common steps and associated questions into a high-level process model. This is useful for future CH digitization efforts by offering a structured, adaptable approach for planning, execution, and stakeholder integration across both tangible and intangible heritage contexts.

The creation of a generalized process model is particularly relevant in the context of the Horizon Europe project DIGICHer, which focuses on digitizing cultural heritage of minority communities to support equity, diversity, and participatory engagement. The project's planning stages have already benefited from a conceptual overview of the digitization process, particularly in determining how and when to involve stakeholders such as communities and tourism institutions.

3. Research questions

This study aims to identify topics in CH digitization papers and use them to recognize certain steps that are relevant for each digitization undertaking. These information after a generalization will help to derive the steps of the digitization process as well as connected tasks and guiding questions.

The research questions are:

- What does a generalized CH digitization process based on topics and connected steps look like?
- What topics are connected to each step?
- What are connected questions relevant to the different steps?

The outcome can help people and projects new to digitization with planning, preparation and execution of the undertaking. It is also helpful in considering when, how and for what stakeholders need to be involved in the digitization process.

4. Methodology

4.1. Scoping review

A scoping review was chosen for this study. This methods is usually used to compile essential information and ideas on a topic and to gain a sense of the academic discussions and developments [TLZ*2016]. It is usually utilized as a groundwork aimed at a fresh understanding to extract others' contributions and opinions together with clarifying definitions and identifying research gaps [PRG*2014]. This broad overview is ideal for identifying and categorizing topics in CH digitization which will help to reconstruct and understand the digitization process and its steps.

A scoping literature review is usually conducted according to a specific protocol to safeguard its reliability and replicability. The procedures used in this analysis were (1) defining the research

questions; (2) searching for appropriate works of literature; (3) screening the papers for relevance; (4) extracting the data, and (5) compiling, summarizing, and presenting the results [AO2005].

An initial search of relevant literature was conducted using the available scientific database Web of Science with the following search strings: ("Cultural Heritage") AND ("Digitization/Digitisation") as part of the title, keywords or abstract. The search results included 586 Journal articles in English language that were published between 2014 and 2024. The publications included a title, authors, abstract and partially full papers as PDFs (for 402 papers, 84,5%). Duplicates were removed automatically while a review of title and abstract resulted in further exclusion of papers that did not hold relevance for the research topic and rather dealt with topics like e.g., climate change and medicine.

A total of 487 papers was considered as a sample for the scoping review that deals with CH in 65 international countries. Thematic analysis was applied to identify and report themes within the articles [BC2012]. The abstracts were analyzed through coding using MaxQDA. Coding helped to identify patterns in each of the abstracts [KR2019] that are connected to 1) the CH aspect, 2) tools and methods used for digitization, 3) actors, stakeholder and disciplines involved, 4) intentions and outcomes of projects, 5) the articles 'main topic and 6) any pitfalls and issues mentioned. Subsequently, the codes were used to build themes and categories. This approach helped in identifying significant topics which could later be clustered according to similar or connected topics. This was essential in recognizing and aligning the different tasks and questions of the digitization process. This inductive method facilitated the eventual deriving of distinct steps of the digitization process. Additionally, guiding questions based on examples of the papers were created.

4.2. Limitations

The examined papers were based only on English-written journal articles that only came from Web of Science. This database already offered such a large and diverse number of relevant articles that it was decided not to include data from other databases. Articles from 2014 to 2024 were chosen to cover a timeframe that allows for older projects, topics and technology to appear because they might still provide valuable insights. Deviating from the usual scoping review process, the selection and analysis of the articles was carried out by a single person which leaves room for bias.

5. Findings

5.1. Insights into the data

The review considers a large number of publications - in total 487 articles from 2014 to 2024. Citations in the findings are kept to a minimum. The number of publications related to CH digitization steadily increased over the 10-year span with an exception in 2018 (see Figure 1). The sample includes articles from 240 different journals. The Journal of Cultural Heritage offered the largest number of desired articles for this scoping review with 29 publications (12%). Figure 2 gives an overview of the 11 journals with the most publications considered for this review.

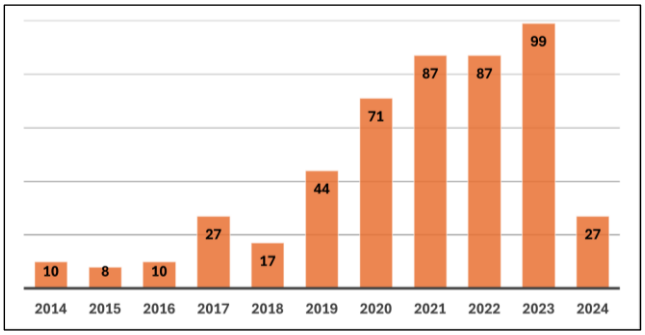

Figure 1: *Number of articles from 2014 – 2024 considered for the sample.*

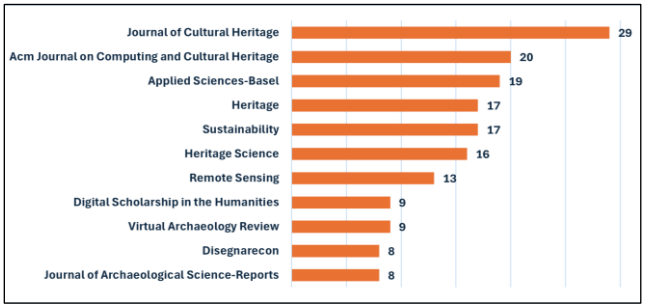

Figure 2: *Top 11 out of 240 journals with articles in the sample.*

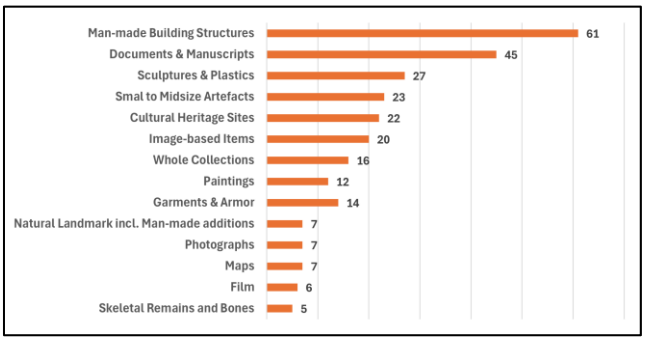

Figure 3: *Breakdown of mentions categorized as tangible CH.*

The samples included 315 (64,7 %) project papers describing specific digitization projects or attempts, whereas the rest (172 papers, 35,3%) of the articles focused on the history of digitization or the developments in specific countries. The project papers deal with the digitization of a large variety of tangible (288 articles,

91,4%) and intangible (27 articles, 8,6%) CH. Specific mentions of intangible CH include music and audio recordings, dance and performance and traditional craftsmanship. Many objects and items of tangible CH are mentioned in the articles with man-made building structures e.g., buildings, bridges, monuments and architectural elements and documents and manuscripts e.g., music sheets, hand- and typewritten documents, books, letters, newspapers being the most frequently named ones. Figure 3 gives an overview of how often certain types of tangible CH were mentioned in the articles.

5.2. Digitization process steps based on topic categorization

After categorizing the diverse topics on CH digitization, it was revealed that the digitization process can be divided into 7 consecutive steps:

(1) Planning
(2) Preparation
(3) Technology selection and data acquisition
(4) Post-processing
(5) Storage and maintenance
(6) Access and dissemination
(7) Documentation

Figure 4 provides an overview of the derived steps and connected tasks and topics.

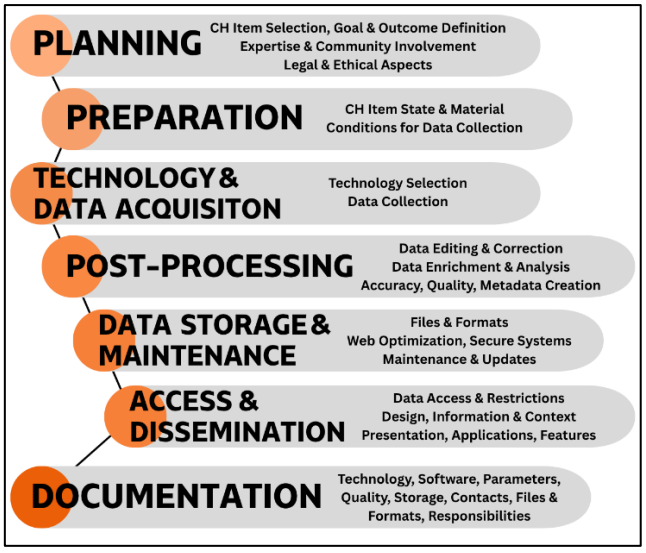

Figure 4: *Steps and connected tasks of a generalized digitization process for CH.*

Planning. The planning phase is a critical foundation of any CH digitization project. It is the time to define a project's direction, set priorities, and ensure the initiative is sustainable, respectful, and impactful. The item selection is a core task and is usually guided by cultural significance, physical condition, and scholarly or public interest. Many of the investigated papers touch on this topic within a dedicated section emphasizing the importance of an initiative. Another aspect is the prioritization of fragile or environmentally threatened materials as well as underrepresented or marginalized perspectives from indigenous or colonized communities for ethical and inclusive practices [Odu2020]. Guiding questions for this task are:

- Is the CH aspect of significance and interest to be digitized?
- Are there any known threats to the CH aspect?

Clearly defining project goals (e.g., preservation, access, education, research, reconstruction, or community engagement), outcomes (e.g., applications, visualizations, virtual reconstructions and experiences, digital archives and repositories) and audiences (e.g., experts, specific communities, specific age group) are also essential. Digital reconstruction can aid restoration, while educational applications enhance public understanding. Identifying the target audience that can range from local communities to researchers and ensures broader societal value. The connected questions are:

- What are the specific goals and planned outcomes of the digitization project?
- Who will be interested in the result or benefit from it?

Legal and ethical considerations form one of the most challenging aspects of planning and are constantly named as a major pitfall. Issues include copyright, intellectual property, and ethical and moral rights, particularly where digital reproductions complicate ownership and attribution [Bor2018], [KS2023]. Further complexities arise from orphan works [MT2019], licensing gaps, and public domain restrictions [VS2021]. Tensions also emerge between open access and copyright, especially with newer technologies like 3D printing and AR/VR. Ethical frameworks must also consider indigenous knowledge, cultural protocols, and community rights [KJ2023], [RN2023], [WC2023]. Guiding questions might be:

- What is the copyright status of the CH item and do intellectual property rights apply?
- Who is the owner of the CH item?
- Have Indigenous knowledge systems, community ownership rights, and cultural protocols been acknowledged and respected?
- What measures are in place to ensure compliance with national and international ethical guidelines?

Complex digitization projects call for the involvement of multidisciplinary teams (consisting of e.g., historians, conservators, technologists, legal experts, and community stakeholders) to meet all demands of a well-rounded project. This of course makes a good communication strategy and infrastructure indispensable. Collaborative frameworks and shared digital environments facilitate collaboration and cross-disciplinary

integration [TVV2019]. Relevant questions to address this topic are:

- What types of expert knowledge are required to ensure a holistic and responsible digitization process?
- How can collaboration across disciplines and institutions be effectively facilitated?
- What should communication with the community look like?
- How can the community be actively involved?
- How can community involvement enhance the CH digitization?

Preparation. Before the data collection can start, careful preparation tailored to the condition and context of the CH items or aspect is necessary which may need physical stabilization or non-invasive, non-damaging technology. Environmental constraints also shape digitization. Underwater sites need specialized techniques [RDSL2018], and objects on public display may require non-disruptive workflows [KZM*2023] or closing down of sites. Predictive digitization fills in missing data and can be factored in when full access is not possible [PSM*2028]. Guiding questions are:

- Is the object stable enough for digitization, or does it require physical recovery or conservation first?
- Can the object be safely handled or must non-contact methods be used?
- Is the object located in a challenging environment (e.g., underwater, on continuous public display)?
- What logistical strategies can minimize disruption in these settings?
- For intangible CH: Does digitization depend on a certain event with specific conditions/situations? Is a specific group of people needed at a specific place?

Technology selection and data acquisition. Selecting appropriate technologies for CH digitization depends on the object's or aspect's characteristics, intended use, and available resources which are all related to information and decisions from the first two steps. 3D technologies such as photogrammetry and laser scanning have become central tools for CH documentation, offering precise, non-invasive data acquisition and are the most prominent technologies mentioned in the sample literature. Comparative studies have evaluated various methods with regard to accuracy, cost, and applicability [RTA2022]. Recent developments emphasize low-cost, portable, and user-friendly solutions, including smartphone-based setups and self-built equipment, enabling broader adoption by smaller institutions and non-experts [BBGG2023]. Key challenges e.g., material reflectivity, lack of surface texture, or internal features require specialized acquisition techniques and lighting. Figure 5 provides an overview of the use of certain technology for the CH digitization within the sample data.

Moreover, participatory models including *crowdsourcing* and public-private partnerships offer scalable approaches to data collection and enrichment, while also fostering public engagement with heritage [AC2017].

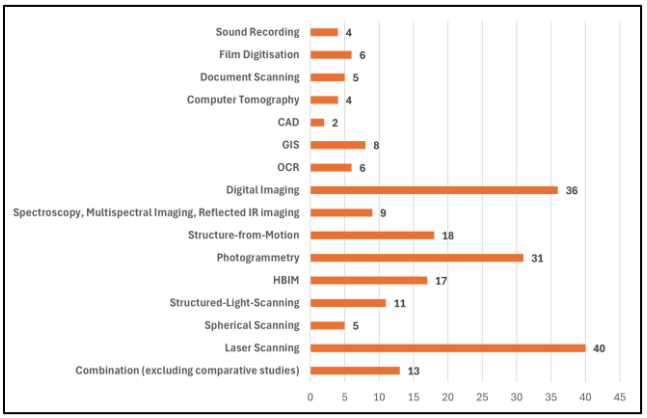

Figure 5: *Overview of the use of certain digitization technologies within the sample data.*

The following questions can help with technology selection, acquisition setup and data collection.

- What acquisition technologies are most suitable given the object's size, fragility, and material or the aspect's characteristics?
- What materials and surface characteristics (e.g., reflectivity, transparency, repetitive patterns) of the objects might pose challenges?
- Are low-cost, portable non-expert alternatives appropriate for the project's goals and available expertise?
- Is lighting setup needed to capture key surface details?
- Is automation (e.g., robotic arms or turntables) required to handle a large number of artifacts efficiently?
- How will lighting, filters, or imaging be optimized for accuracy and color fidelity?
- What level of resolution, accuracy, and metadata are needed for future uses?
- Will the process involve real-time constraints (e.g., tourists, outdoor conditions, immovable objects) that limit scanning conditions?
- Can the project benefit from crowdsourcing? What systems are in place to manage, motivate, and retain volunteers?

Post-processing. Post-processing is essential for enhancing and refining digitized CH data. It involves cleaning, aligning, optimizing, reducing and semantically enriching digital outputs to ensure usability, accuracy, and preservation quality. Data correction, cleaning, optimization and reduction techniques such as contrast adjustment, denoising, grayscale conversion and

morphological operations prepare input data for tasks like OCR and 3D modeling. Image alignment and registration ensure spatial coherence across datasets, with cross-time 3D point cloud registration and photogrammetric workflows improving consistency and accuracy. 3D model optimization includes mesh cleaning and down-sampling. The transfer of annotations, especially with HBIM supports recognition and further analysis of building elements and structure. Texture mapping and color correction enhance visual realism through deblurring, illumination correction, and integration of scanned geometry with photographic textures.

Segmentation and recognition, supported by deep learning, enabled object, text, and structural analysis in 2D/3D environments, facilitating classification, annotation, and searchability [KIP*2021], [MST*2019]. Metadata and annotation practices emphasize structured schemas, controlled vocabularies, and semantic technologies to support discoverability and contextual understanding, with AI increasingly assisting in metadata extraction and linkage [CT2021], [WCT2020]. Data fusion techniques integrate multi-source inputs (e.g., photogrammetry, CT, LiDAR) to construct accurate digital twins and comprehensive models. Digital restoration techniques range from colorimetry and degradation diagnosis to virtual reconstructions and recovery of lost content, aiding both conservation and research [SCH*2019].

The sample articles provide a huge number of very diverse, detailed and specific approaches and descriptions for post-processing of collected data. Guiding questions which may be considered for the individual undertakings are:

- Which approaches for data cleaning, noise reduction, optimizing, color correction, deblurring and illumination correction can help with data improvement and preparation for further processing?
- Which approaches to data aligning, registration and georeferencing can support data combination or fusion?
- Which approaches for mesh cleaning, down-sampling and texture mapping are suitable for 3D model optimization?
- Which approaches concerning annotation, segmentation, recognition and classification are suitable for data enrichment?
- Which approaches connected to controlled vocabulary and automatic metadata extraction among others are beneficial for metadata creation?
- Which approaches from semantic technologies and linking can support contextualization?
- Which approaches are needed for digital restauration and virtual reconstruction?

Storage and maintenance. Effective storage and maintenance of digitized CH data require scalable and sustainable infrastructures and long-term maintenance. Custom digital repositories support secure storage, controlled access, and sharing of materials, while cloud computing enhances scalability and processing power [WS2022]. Large 3D datasets from scanning or photogrammetry demand optimization strategies. These include mesh simplification and model generalization for web access while preserving high-resolution versions for archiving and research. Cultural Big Data Repositories further enable efficient storage and analytics [SB2020], though maintenance of digital infrastructures, particularly for audiovisual content, remains costly and complex. Metadata standards and archival choices also shape CH preservation calling for international standards. A list of topics and connected questions is provided below.

Secure storage.

- What methods are used to ensure the confidentiality and integrity of stored CH data?
- How are backups managed to prevent data loss due to hardware failure or cyberattacks?
- What protocols are in place to secure long-term preservation of digital assets against evolving threats?

Controlled access.

- How are access permissions defined for different user roles (e.g., researchers, curators, public users)?
- What authentication and authorization mechanisms are implemented to manage access to sensitive or restricted CH data?
- How is access to culturally sensitive or proprietary data governed in collaboration with source communities?

Data sharing.

- What platforms or infrastructures are used to share CH data with partners or the public?
- How are licensing and copyright issues handled when sharing digitized cultural materials?
- Are there standardized formats or protocols used to facilitate interoperability in data sharing?

Processing power.

- What computational resources are available for processing large datasets such as 3D models or high-resolution images?

Data optimization for online access and maintenance: mesh simplification, model generalization.

- How does 3D data need to be optimized for web access?
- What is necessary to have improved user experience?

Access and dissemination. CH initiatives emphasize open access, personalization, and participation to improve user engagement and inclusivity. Advanced tools such as content-based image retrieval, semantic search, and facial recognition support navigation through large, diverse collections [MVV*2023]. These systems address challenges tied to distributed data, varying standards, and heterogeneous sources [DKBD2022].

Web platforms enhance access via spatiotemporal maps, enriched metadata, and exploratory interfaces [ABVV2022],

[Haw2022]. Mobile AR and 3D repositories offer in-situ experiences, aided by Semantic Web technologies and Knowledge Graphs [LMS*2022]. Special attention is given to diverse audiences, including the elderly and Indigenous communities, through open, accessible archives and tools for community-driven knowledge sharing [BD2016], [Cap2023], [MPL*2023].

Applications range from VR simulations and interactive storytelling to educational tools for intangible CH transmission and embodied experiences [GL2023]. Digitization efforts often yield 3D models, enabling reconstructions, virtual tours, and 3D-printed replicas for education and accessibility [KSA2023]. These innovations, supported by ontology-based databases, cloud platforms, and XR systems, contribute to both scholarly research and public engagement. Any relevant questions for this step are again highly individual, therefore the focus here is more on listing connected topics.

- What are engaging, useful, usable and supportive features, tools and applications for CH data presentation and dissemination?
- Useful features: content-based data retrieval, semantic search, facial recognition, navigation, ontology-based database, visual access & browsing, linked data
- Enhanced access: spatio-temporal maps, enriched metadata, exploratory interfaces, 3D repositories, AR, semantic web technologies, knowledge graphs, multilingual access, multimodal access
- Audience consideration.
- Applications: VR/AR/XR, simulations, interactive storytelling, education, virtual tours, 3D models, visualizations

Documentation. Despite the technical sophistication of many digitization initiatives, documentation practices are hardly mentioned in the journal articles, with only a small fraction of reviewed studies directly engaging with the topic. The case for rigorous documentation becomes especially critical in 3D modeling and reuse of data, where the utility of digital surrogates depends on the inclusion of detailed information like data on photographic quality, mesh and point cloud resolutions, and quantitative modeling errors, alongside software and processing parameters [LLMD2019]. A promising methodological approach is the Scientific Reference Model, a structured, published framework that consolidates object information, credibility levels, and rights metadata, thereby serving as a reusable knowledge base for further applications and research [KABC2023]. Furthermore, emerging technologies such as machine learning and AI present new documentation challenges, particularly concerning potential biases and data provenance. Scholars advocate for embedding CH datasets within broader socio-technical narratives which are rooted in data archaeology [Lee2021], or create digital object biographies [Zua2023].

Documentation should include information on selected technology, quality, resolutions errors, software, processing parameters, storage and access, contacts, files and formats. Some guiding questions are:

- Which communities or stakeholders have been consulted or involved, and how?
- What acquisition methods and technologies were used?
- What are the known errors, limitations, or uncertainties in the data?
- What file formats are used for raw and processed data and were is the data stored?
- Who is responsible for data preservation and maintenance over time?

6. Discussion

This study sought to develop a generalized process for the digitization of CH based on insights gathered through a comprehensive scoping review with project-based literature. The insights collected through the scoping review helped to identify 7 consecutive steps that belong to the digitization process:

(1) Planning
(2) Preparation
(3) Technology selection and data acquisition
(4) Post-processing
(5) Storage and maintenance
(6) Access and dissemination
(7) Documentation

Previous works (see section 2) often focused narrowly on the stages before, during, and after data acquisition. However, this structure proved insufficient for capturing the full complexity of digitization processes. Our findings suggest the need for a more detailed process that better reflects the variety of decisions and tasks involved. While steps such as planning, data acquisition, and post-processing remain essential, they were expanded upon, and previously overlooked phases like documentation added. Documentation was never part of any of these previous contributions. Post-processing was sometimes split into data cleaning and processing or analysis, not acknowledging how these tasks are dependent on each other and should be considered in the same step. Access and dissemination were kept to a minimum not relying on visualization, or applications with new features. The concept that ideation and brainstorming take place before planning and should be a separate phase are currently discussed in the DIGICHer project.

Another goal was to organize the topics connected to each step and compile suitable questions that help with planning and steering the digitization effort. The identified topics with explanations and descriptions indicated what tasks and questions need to be considered at which stages during the digitization process.

However, the gathered topics and questions are a good starting point but need further expansion and refinement. Although the proposed process integrates all topics found in the reviewed literature, it is likely that not all relevant aspects of digitization were captured. The literature sample may omit emerging or less widely reported practices, which highlights the need for ongoing refinement of the process with new insights and future developments.

It is clear that the data and study were good to derive the desired information but not good enough to result in an all-encompassing process. This is in part due to the fact that topics are either missing or sometimes over- or under-represented. Topics such as data acquisition, post-processing, data analysis, and the creation of engaging presentations were heavily emphasized in the literature, while e.g., documentation received little attention. Figure 5 showing the technology mentioned for digitization indicates that 3D digitization is very much over-represented in the literature while dealing with audio and film is probably under-explored leaving the connected topics and tasks undetected. The limited attention to documentation in academic literature may stem from its perceived lack of novelty, or because it often occurs at the end of projects leaving less opportunity for timely publication.

Notably absent from most studies were practical considerations such as resource calculation, budgeting, staffing, technical infrastructure, and the creation of a sustainable business model including long-term digital preservation and data curation. These topics, which are crucial to the feasibility of any digitization project and belong to the planning stage, were underrepresented in science-based literature perhaps due to the academic nature of the sources. The set up of the literature search routine with a focus on digitization and CH probably contributed to the limited insights. Even though the scientific community is very interdisciplinary, these topics were not considered in the papers. To address these gaps, future research should include project reports, grey literature, and publications from GLAM institutions and industry stakeholders. These sources are more likely to address operational and strategic topics that are underrepresented in academic literature.

7. Conclusion

This paper presents an initial attempt to develop a generalized digitization process for cultural heritage, derived from a comprehensive scoping review of diverse scholarly sources. The resulting process bridges the gap between project-specific workflows and high-level strategic guidance, offering a flexible yet structured foundation adaptable to various digitization contexts. It outlines essential stages, decision points, and thematic areas to support practitioners throughout the digitization effort. Accompanying topics and guiding questions are intended to assist those undertaking CH digitization projects in planning and execution.

Future work will focus on refining this process into a practical tool tailored to the needs of various stakeholders, including decision-makers and project teams. Insights from additional sources—such as project reports, grey literature, and contributions from the GLAM sector—will further enrich the framework. Surveys or interviews with domain experts would also support its continued development and evaluation.

Once the overarching digitization process is firmly established, it can serve as the basis for creating task-specific workflows, tool guidelines, and role-based responsibilities. This structured approach not only facilitates better planning and coordination, but also opens pathways for optimization, automation, and workflow efficiency in CH digitization initiatives. By defining key steps and their interdependencies, the conceptual model enhances the design and implementation of digitization projects, aligning stakeholder involvement and expertise with specific project goals. Its sequential structure ensures that each phase informs the next, ultimately improving planning, effectiveness, and the likelihood of project success.

Acknowledgments

The research upon which this paper is based has received funding from the European Union's research and innovation program Horizon Europe under the grant agreement No. 101132481.

References

[AC2017] ALAM, S. L., CAMPBELL, J.: Temporal Motivations of Volunteers to Participate in Cultural Crowdsourcing Work. Information Systems Research, 28, 4 (Dec 2017), 744-759

[ABVV2022] ALI, D., BLYAU, T., VAN DE WEGHE, N., VERSTOCKT, S.: Context-Aware Querying, Geolocalization, and Rephotography of Historical Newspaper Images. Applied Sciences-Basel, 12, 21 (Nov 2022)

[Apa2017] APARAC-JELUŠIĆ, T.: Digitization and Coordination of Digitization in the European CH Sector. In Digital Libraries for Cultural Heritage: Development, Outcomes, and Challenges from European Perspectives. (2017), pp. 37-71.

[AO2005] ARKSEY, H., O'MALLEY, L.: Scoping studies: towards a methodological framework. International journal of social research methodology, 8, 1 2005), 19-32

[BBGG2023] BACCI, G., BOZZOLA, M., GAIANI, M., GARAGNANI, S.: Novel Paradigms in the Cultural Heritage

Digitization with Self and Custom-Built Equipment. Heritage, 6, 9 (Sep 2023), 6422-6450

[BAS2015] BANDI, S., ANGADI, M., SHIVARAMA, J.: Best practices in digitization: Planning and workflow processes. 2015)

[BD2016] BHOWMIK, S., DÍAZ, L.: Hot Stones and Cool Digitals: Sustainable Contact Zones for Intangible Cultural Heritage in Finland. International Journal of Intangible Heritage, 11, 2016), 162-171

[Bor2018] BORISSOVA, V.: Cultural heritage digitization and related intellectual property issues. Journal of Cultural Heritage, 34, (Nov-Dec 2018), 145-150

[BC2012] BRAUN, V., CLARKE, V.: Thematic analysis. 2012)

[Cap2023] CAPASSO, M. C.: Changes in use and accessibility to Museums as a consequence of the cultural heridage digitalisation. Museologia Scientifica, 17, 2023), 60-66

[Cas2022] CASEY, M.: Media Preservation and Digitization Principles. 2022)

[CDN2021] CHISITA, C. T., DURODOLU, O. O., NGOAKETSI, J.: Evaluating the processes and procedure of digitalization workflow. In Digital Libraries-Advancing Open Science. (2021).

[CT2021] CHRISTOU, D., TSOUMAKAS, G.: Extracting Semantic Relationships in Greek Literary Texts. Sustainability, 13, 16 (Aug 2021)

[DD2009] DAHLSTRÖM, M., DORACIC, A.: Digitization education. D-lib Magazine, (15, 3/4 2009), 1082-9873

[DKBD2022] DE MOOIJ, J., KURTAN, C., BAAS, J., DASTANI, M.: A Computational Framework for Organizing and Querying Cultural Heritage Archives. Acm Journal on Computing and Cultural Heritage, 15, 3 (Sep 2022)

[GL2023] GAVIOLI, F., LIGAS, E.: Leonardo Da Vinci's Last Supper: Enhanced Uses and New Digital Pathways. Archeomatica-Tecnologie Per I Beni Culturali, 14, 3 2023), 16-19

[Haw2022] HAWKINS, A.: Archives, linked data and the digital humanities: increasing access to digitised and born-digital archives via the semantic web. Archival Science, 22, 3 (Sep 2022), 319-344

[IFL2015] IFLA Rare Books: Guidelines for Planning the Digitization of Rare Book and Manuscript Collections. 2015)

[KSA2023] KANTAROS, A., SOULIS, E., ALYSANDRATOU, E.: Digitization of Ancient Artefacts and Fabrication of Sustainable 3D-Printed Replicas for Intended Use by Visitors with Disabilities: The Case of Piraeus Archaeological Museum. Sustainability, 15, 17 (Sep 2023)

[KJ2023] KEENEY, G., JONES, D. S.: The Debauched Commons: A Dark Parable. International Journal for the Semiotics of Law-Revue Internationale De Semiotique Juridique, 36, 5 (Oct 2023), 2115-2132

[KZM*2023] KESIK, J., ZYLA, K., MONTUSIEWICZ, J., MILOSZ, M., NEAMTU, C., JUSZCZYK, M.: A Methodical Approach to 3D Scanning of Heritage Objects Being under Continuous Display. Applied Sciences-Basel, 13, 1 (Jan 2023)

[KS2023] KLINOWSKI, M., SZAFAROWICZ, K.: Digitisation and Sharing of Collections: Museum Practices and Copyright During the COVID-19 Pandemic. International Journal for the Semiotics of Law-Revue Internationale De Semiotique Juridique, 36, 5 (Oct 2023), 1991-2019

[KA2010] KOGANURAMATH, M. M., ANGADI, M.: Digitisation in an academic library: a success story at Tata Institute of Social Sciences. DESIDOC Journal of Library & Information Technology, 30, 1 2010), 38

[KIP*2021] KOUTSOUDIS, A., IOANNAKIS, G., PISTOFIDIS, P., ARNAOUTOGLOU, F., KAZAKIS, N., PAVLIDIS, G., CHAMZAS, C., TSIRLIGANIS, N.: Multispectral aerial imagery-based 3D digitisation, segmentation and annotation of large scale urban areas of significant cultural value. Journal of Cultural Heritage, 49, (May-Jun 2021), 1-9

[KR2019] KUCKARTZ, U., RÄDIKER, S.: Analyzing qualitative data with MAXQDA. Springer. 2019.

[KABC2023] KUROCZYŃSKI, P., APOLLONIO, F. I., BAJENA, I., CAZZARO, I.: Scientific reference model–defining standards, methodology and implementation of serious 3d models in archaeology, art and architectural history. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 48, 2023), 895-902

[Lee2021] LEE, B.: Compounded Mediation: A Data Archaeology of the Newspaper Navigator Dataset. Digital Humanities Quarterly, 15, 4 2021)

[LMS*2022] LEMOS, D. L. D., MARTINS, D. L., SA, A. M. E., MARTINS, L. C., DO CARMO, D.: A Proposal in Creating a Semantic Repository for Digital 3D Replicas: The Case of Modernist Sculptures in Public Spaces of Rio De Janeiro. Knowledge Organization, 49, 3 2022), 151-171

[LLMD2019] LORENZO, G., LÓPEZ, L., MORALEJO, R. A., DEL PAPA, L. M.: Sfm Photogrammetry Applied to Taxonomic Determination of Archaeofauna Remains. Virtual Archaeology Review, 10, 20 (Jan 2019), 70-83

[MPL*2023] MAGNANI, M., PORSANGER, J., LAITI, S., MAGNANI, N., OLLI, A. M., RAUHALA, P., VALKEAPÄÄ, S., HOLLINGER, E.: Small collections remembered: Sámi material culture and community-based digitization at the Smithsonian Institution. Museum Anthropology, 46, 2 (Sep 2023), 92-105

[MT2019] MARTINEZ, M., TERRAS, M.: 'Not Adopted': The UK Orphan Works Licensing Scheme and How the Crisis of

Copyright in the Cultural Heritage Sector Restricts Access to Digital Content. Open Library of Humanities, 5, 1 2019)

[MVV*2023] MILLEVILLE, K., VAN DEN BROECK, A., VANDERPERREN, N., VISSERS, R., PRIEM, M., VAN DE WEGHE, N., VERSTOCKT, S.: Enriching Image Archives via Facial Recognition. Acm Journal on Computing and Cultural Heritage, 16, 4 (Dec 2023)

[MST*2019] MUEHLBERGER, G., SEAWARD, L., TERRAS, M., ARES OLIVEIRA, S., BOSCH, V., BRYAN, M., COLUTTO, S., DÉJEAN, H., DIEM, M., FIEL, S., GATOS, B., GREINOECKER, A., GRÜNING, T., HACKL, G., HAUKKOVAARA, V., HEYER, G., HIRVONEN, L., HODEL, T., JOKINEN, M., KAHLE, P., KALLIO, M., KAPLAN, F., KLEBER, F., LABAHN, R., LANG, E. M., LAUBE, S., LEIFERT, G., LOULOUDIS, G., MCNICHOLL, R., MEUNIER, J. L., MICHAEL, J., MÜHLBAUER, E., PHILIPP, N., PRATIKAKIS, I., PUIGCERVER PÉREZ, J., PUTZ, H., RETSINAS, G., ROMERO, V., SABLATNIG, R., SÁNCHEZ, J. A., SCHOFIELD, P., SFIKAS, G., SIEBER, C., STAMATOPOULOS, N., STRAUSS, T., TERBUL, T., TOSELLI, A. H., ULREICH, B., VILLEGAS, M., VIDAL, E., WALCHER, J., WEIDEMANN, M., WURSTER, H., ZAGORIS, K.: Transforming scholarship in the archives through handwritten text recognition Transkribus as a case study. Journal of Documentation, 75, 5 (Sep 9 2019), 954-976

[Odu2020] ODUMOSU, T.: The Crying Child On Colonial Archives, Digitization, and Ethics of Care in the Cultural Commons. Current Anthropology, 61, (Oct 1 2020), S289-S302

[PRG*2014] PHAM, M. T., RAJIĆ, A., GREIG, J. D., SARGEANT, J. M., PAPADOPOULOS, A., MCEWEN, S. A.: A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Research synthesis methods, 5, 4 2014), 371-385

[PSM*2018] PRATIKAKIS, I., SAVELONAS, M. A., MAVRIDIS, P., PAPAIOANNOU, G., SFIKAS, K., ARNAOUTOGLOU, F., RIEKE-ZAPP, D.: Predictive digitisation of cultural heritage objects. Multimedia Tools and Applications, 77, 10 (May 2018), 12991-13021

[Rap2021] RAPP, A.: Digitisation: Opportunities for Tradition and Research in the Humanities and Cultural Studies. Bibliothek Forschung Und Praxis, 45, 2 (Jul 2021), 255-261

[RN2023] REIJERKERK, D., NYITRAY, K. J.: (Re)Moving Indigenous Knowledge Boundaries: A Review of Library and Archival Collection Literature since the Enactment of the Native American Graves and Repatriation Act (NAGPRA). Collection Management, 48, 1 (Jan 2 2023), 22-47

[RDSL2018] REPOLA, L., DI CARLO, N. S., SIGNORETTI, D., LEIDWANGER, J.: Virtual simulation of a late antique shipwreck at Marzamemi, Sicily: Integrated processes for 3D documentation, analysis and representation of underwater archaeological data. Archaeological Prospection, 25, 2 (Apr-Jun 2018), 99-109

[RTA2022] RUIZ, R. M., TORRES, M. T. M., ALLEGUE, P. S.: Comparative Analysis Between the Main 3D Scanning Techniques: Photogrammetry, Terrestrial Laser Scanner, and Structured Light Scanner in Religious Imagery: The Case of The Holy Christ of the Blood. Acm Journal on Computing and Cultural Heritage, 15, 1 (Feb 2022)

[SB2020] SHARMA, S., BAWA, S.: CBDR: An efficient storage repository for cultural big data. Digital Scholarship in the Humanities, 35, 4 (Dec 2020), 893-903

[SCH*2019] STROMER, D., CHRISTLEIN, V., HUANG, X. L., ZIPPERT, P., HAUSOTTE, T., MAIER, A.: Virtual cleaning and unwrapping of non-invasively digitized soiled bamboo scrolls. Scientific Reports, 9, (Feb 19 2019)

[TVV2019] THIVET, M., VERRIEZ, Q., VURPILLOT, D.: Aspectus: a flexible collaborative tool for 3D data exploitation in the field of archaeology and cultural heritage. In Situ-Revue De Patrimoines, 39 2019)

[TLZ*2016] TRICCO, A. C., LILLIE, E., ZARIN, W., O'BRIEN, K., COLQUHOUN, H., KASTNER, M., LEVAC, D., NG, C., SHARPE, J. P., WILSON, K.: A scoping review on the conduct and reporting of scoping reviews. BMC medical research methodology, 16, 2016), 1-10

[VS2021] VOLKOVA, K. Y., SHRAYBERG, Y. L.: Copyright and information market today: European initiatives and library privileges. Nauchnye I Tekhnicheskie Biblioteki-Scientific and Technical Libraries, 11 2021), 54-64

[WD2023] WAGNER, A., DE CLIPPELE, M. S.: Safeguarding Cultural Heritage in the Digital Era - A Critical Challenge. International Journal for the Semiotics of Law-Revue Internationale De Semiotique Juridique, 36, 5 (Oct 2023), 1915-1923

[WCT2020] WANG, X. G., CHANG, W. L., TAN, X.: Representing and Linking Dunhuang Cultural Heritage Information Resources Using Knowledge Graph. Knowledge Organization, 47, 7 2020), 604-615

[WS2022] WANYAN, D., SHANG, T.: Application of cloud technology in digital cultural heritage: an analysis of public culture cloud platforms in China. Digital Library Perspectives, 38, 2 (Apr 18 2022), 222-236

[Zua2023] ZUANNI, C.: Object biographies in the digital age: documentation, life-histories, and data. International Journal of Heritage Studies, 29, 7 (Jul 3 2023), 695-710