AL-HAYAT Research Journal

HJRS HEC Journal Recognition System Category Y

Volume 4 Issue 1 (Oct-Dec 2026)

ISSN (Online): 3006-4767
ISSN (Print):: 3006-4759

Artificial Intelligence, Digital Cultural Heritage Rights, and International Law: A Digital Humanities Study of Pakistan and India

Wagar Ali Solangi
Advocate High Court
MPhil Research Scholar
Department of History, Arts and Cultural Heritage
University of Education, Lahore, Pakistan
Sidra Soomro
PEC-Registered Software Engineer
M.E. Software Engineering, Department of Software Engineering
Mehran University of Engineering & Technology, Jamshoro, Pakistan

Corresponding author: Waqar Ali Solangi
Email: waqaralisolangi2012@gmail.com

Abstract

This study examines the effects of the use of Artificial Intelligence in the digitalization of cultural heritage on the assumption of ownership, rights and governance of digital cultural memory. The digitization of collections brings new copyright and ethical issues of control and use to the forefront as AI allows the documentation, interpretation and reproduction of heritage materials, and their dissemination in new forms. It seeks answers to the following questions: What does ownership, access and governance of AI mediated cultural heritage mean in the context of international cultural heritage law and what does a comparison of the two countries reveal about the approaches and the regulatory needs in the two countries? It explores the connection between intellectual property, cultural rights, community interests, and digital sovereignty and the difference between ownership of the physical heritage and rights and responsibilities for its digital representation and for uses of it generated by AI. The methodology adopts a multi-paradigmatic approach that involves a doctrinal analysis of various relevant instruments and national legal frameworks of the international community and document and metadata analysis of selected initiatives in digital heritage in Pakistan, compared with the Indian reference point. The analysis will examine the appropriation of contemporary legal frameworks and norms, and institutional approaches, for the purposes of placing the concept of attribution, consent, access, reuse and accountability in digital systems, including through the visibility and representation of heritage. This abstract does not contain empirical findings as the materials have not yet been analysed. Conclusions will be drawn on the basis of the completed legal and documentary analysis. The study is an interdisciplinary approach to exploring the concepts of ownership and governance of AI-based heritage, and seeks to provide guidance for culturally sustainable policies, institutional stewardship and protection of communities whose heritage is captured digitally or for use in AI systems.

Keywords: Digital memory, Cultural heritage law, Artificial intelligence, Digital sovereignty, Intellectual property, Community rights

1. Introduction

When artificial intelligence systems classify, link with other records, or interpret a more general meaning, a digitized manuscript, photograph, or museum object is not just a surrogate, but rather a new one. In the field of cultural heritage practice, AI can be involved in the discovery, description, reconstruction, preservation, and access of cultural records and relics, and can provide new methods for computational involvement with collections (Gîrbacia, 2024; Rouhani, 2023). The growing capacities raise new issues about who can license the reuse of heritage data, who will determine how it should

CC BY

be represented, and who will benefit from AI-generated outputs beyond the original data, which cannot be solved with the power of digitization. The increasing use of artificial intelligence in digitising, classifying, reconstructing, and reproducing cultural heritage raises important questions concerning copyright protection, attribution, consent, and the lawful reuse of cultural materials. Although digitisation may improve preservation and accessibility, the legal status of digital copies, AI-generated reconstructions, and datasets derived from historical collections remains complex. Copyright protection does not necessarily extend to historical facts or public-domain materials themselves, while original digitisation, metadata, photographs, and creative reconstructions may raise separate rights-related questions. Therefore, AI-based heritage projects should distinguish between ownership of physical objects, rights in digital reproductions, and the interests of communities connected to the heritage. This distinction is particularly important in Pakistan and India, where historical collections may contain culturally significant materials with complex institutional and cross-border histories.

The Partition of British India in 1947 provides an important historical context for understanding fragmented cultural heritage in Pakistan and India. The division of territories and institutions affected the custody, accessibility, and documentation of museums, libraries, manuscripts, and archival records. In some cases, historically connected collections became separated across institutions and national boundaries, complicating their interpretation and scholarly study. Digital technologies may help researchers reconnect dispersed records through shared catalogues, digitised manuscripts, and interoperable databases. However, digital reconstruction does not automatically resolve questions of ownership, provenance, institutional authority, or cultural rights. Accordingly, AI-assisted heritage initiatives should preserve the historical context of collections and transparently document the sources, uncertainties, and institutional relationships underlying digital representations.

The present study explores these questions in the light of Pakistan's legal and institutional landscape with a comparative discussion of India and analytical framework of International cultural heritage law.The trouble is that "ownership" of AI-mediated cultural memory merges different claims into a single. Rights in the digital representation of an object do not necessarily exist; the digitizing process may result in a new object, while any rights in the representation depend on the applicable law and originality (Palladino & Bodard, 2023; Wagner & Clippele, 2023).

Second, there may be protected expression or reproductions without underlying heritage or communal meanings which may be subject to copyright. Third, The database protections and contracts can control compilation, extraction or institutional reuse. Fourth, access conditions such as licensing and platform rules may be used to limit access from the moment goods are purchased regardless of ownership. Fifth, community interests in attribution, cultural protocols, representation, and benefit may not be embodied in formal property rules particularly in contested heritage (Rouhani, 2023). Sixth, AI's output creates inquiries related to the input training data, authorial claims, source, and accountability for interpretations.

The research question for Pakistan is how the international norms and the national law can differentiate and integrate such claims without taking digitisation as a transfer of cultural authority.Digital sovereignty is no longer just about where heritage files are stored, but also about who determines how heritage files are extracted, processed, interpreted and circulated. AI infrastructures are not equally equipped and their training data and language can dominate or marginalise other knowledge systems and linguistic communities. The uneven performance of language technologies and the epistemic harms that can be generated by the assumption of universal transferability of models built in dominant contexts have been documented in the research literature (Helm et al., 2024). These differences are what renders sovereignty not only a question of managerial control in international cultural governance, but meaningful participation and control over cultural data for digital heritage (Ghosh, 2024). These dynamics are relevant to Pakistan and South Asia because digitization and AI can expand access to one's regional heritage and can dictate what languages, stories, and forms of knowledge are machine-readable and discoverable. The intersection of legal protection, institutional capacity and community participation is important in Pakistan's Urdu and Punjabi language context to be explored (Hussain et al. 2025).

Comparative focus on India, especially on resources for the Sindhi language, and on the digitization of libraries can shed light on common problem areas, but it should not presume any similarities in legal contexts or in the needs of the same communities on either side of a border (Wadhwa & Chowdhary, 2024).The main research question is: What does Pakistan's legal and institutional landscape show, and how can international cultural heritage law and governance tackle the issues of ownership, rights and accountability in the context of digitization, processing and reinterpretation of cultural heritage through AI? Secondary questions include how current rules differentiate between physical properties and digital representations and uses created by artificial intelligence, how consent, attribution, access, and community authority are being dealt with, and how a comparison with India might help shed light on regulatory requirements. The aim is to identify relevant international norms and national frameworks; analyze some case studies of digital heritage projects and its documentation and accessibility criteria; explore the allocation of authority and responsibility in the institutional

practices; and propose an interdisciplinary framework for culturally responsive governance. The analysis also acknowledges and discusses community authority and ethical responsibility as governance issues that can be beyond traditional property claims (Rouhani, 2023).

The article ties together the notions of intellectual property, cultural rights, institutional stewardship, and community interests to argue that these are essential and must be coordinated if AI-mediated heritage governance is to be addressed. It does not make any claim about uniformity of rules or outcomes in the two jurisdictions, and it only considers Pakistan as the main case and India as a comparative reference; it considers only some legal instruments and digital heritage initiatives. The literature review maps these issues within the field of heritage law, AI ethics and digital sovereignty. The methodology then outlines the doctrinal and documentary analysis, the results of legal rules and institutional practices, the implications of those rules and practices for governance, and conclusions dealing with the priorities for policy and research.

The article's particular and unique contribution is the integrated analytical framework which differentiates six non-identical interests: (1) the title to the physical object, (2) copyright in qualifying expression or reproductions, (3) database-related protection as may be applicable under law, (4) contractual and institutional access conditions, (5) the community and cultural interest which may not be reduced to an interest in property, and (6) the interest of AI-related questions with respect to training, generated output, provenance, and accountability. In practice there may be overlap between these categories, but they are not synonymous: a right or a permission in one does not mean a right or permission in the other. The study must interrogate this split, not assume that digitisation or accessibility equals consent for AI reuse.

2. Literature Review

The study of AI-assisted heritage involves disentangling legal interests that are frequently smudged by the act of digitization. Ownership of a physical object is title and possession, subject to relevant heritage and restitution law, and does not in itself give the copyright holder the copyright to a photograph or scan. Qualifying original expression is covered by copyright; originality and term limit the scope. Database rights, if they exist, cover qualifying investment or extraction, whereas contracts can do so without copyright. Use is governed by different access rights and terms of use for the platforms and institutions. Community needs regarding attribution, cultural protocols and sensitive knowledge can be stronger than property law. Questions of authorship, training, provenance and liability arise with regard to AI outputs. No one of these categories is predetermined to lead to another one (Palladino & Bodard, 2023; Wagner & Clippele, 2023).

This separation is reflected in the nascent literature of AI-spanning the arts. The scientometric analysis of Gîrbacia (2024) covers Web of Science and Scopus publications from 2019 to 2023 and categorizes the applications used and not legal and social effects of specific deployments, as it maps the field growth, without which no deployment can truly be considered. Tiribelli and colleagues' 2024 ethics analysis makes an aggregation of AI ethics principles for cultural heritage and outlines the following concerns: authenticity, interpretive bias, and authorship, and does not represent a jurisdiction-specific legal test (Tiribelli et al., 2024). Oruç's comparative book examines and contrasts the EU, UK and U.S. approaches to digitization and copyright, drawing on doctrinal analysis and interviews with heritage professionals. It presents a conflict between the stakeholders but does not solve the statutory position of Pakistan in its account (Oruç, 2025).

A documentary dimension is added by archival and/or preservation scholarship. If users are to understand what was captured, changed, or created, then digital files must include contextual metadata, provenance, and records of the transformations that may have been made. In 2023, Bushey's literature review takes the image as an emerging form of record and classifies the themes of research as authenticity, misinformation, bias, attribution, transparency and ethics (Bushey, 2023), without being able to measure the prevalence of these questions in the study of images. Israel and Amer's (2022) conceptual conversation on generated content suggests a framework for attributing and holding accountable for generated content through the use of supported infrastructure for provenance.

The difference leaves documentation at the heart of the issue: a faithful scan is evidence; a reconstructed picture is evidence; a created illustration is evidence. Even the notion that the Internet can fix dispossession is thrown into doubt by discussions of colonial collections. Digital copies can help make visible what is not, without returning objects to their original owners, without shifting the interpretative authority, and without fixing problematic acquisition histories. In Rouhani's 2023 remarks on contested digital heritage, the notions of extraction, unequal access to technology, copyright, interpretation and digital benefit are discussed, suggesting that the technological use of digital heritage can generate new inequalities and new narratives (Rouhani, 2023). While not dealing with AI law, Chakraborty's 2023 analysis of digitised

Partition histories in India looks critically at the way digitized records can reinforce colonial and institutional hierarchies. But it is work that can usefully redirect the question from the question of neutrality in the preservation of the digitized to the question of selection, description, circulation and benefit, without claiming that digital restitution is an alternative to physical restitution.

These concerns are taken to the international level with the concept of digital sovereignty.Thus, the issue of digital sovereignty is extended to the international domain. Ghosh (2024) has written about the history of knowledge commons in the Global South, which is a largely print-based phenomenon that has obscured the existence of other forms of knowledge, such as those in performance and embodied in manuscripts. Ghosh is concerned about how AI tools imported from Anglo-American contexts will become the norm, obscuring the need for other Indigenous knowledge types. Similarly, Helm and colleagues link language-modeling bias to epistemic injustice, and warn not to export systems constructed in dominant contexts without considering the linguistic context (Helm et al., 2024). These analyses could be used for examination of who is controlling the standards and infrastructure and who is interpreting it but they should not reduce sovereignty to data localization. Meaningful sovereignty is also an issue in the governance of heritage, in matters related to participation in decision making in terms of extraction, processing, access and reuse, across borders.

Community governance is a remedy, but with care. The CARE principles of Indigenous data governance complement, not supplant, FAIR data principles, and were developed to account for Indigenous data governance realities (Carroll et al., 2020). Rouhani claims that such concerns can get to be used for broader communities where cultural heritage turns into digital data (Rouhani, 2023). But ethical norms are not necessarily translated as rights, and Indigenous systems do not simply fit into every community due to the identical nature of their power systems. Oguamanam (2025) looks at how memory institutions have been thinking about the digitization of TK and how the 2024 genetic-resources treaty has been oriented towards the patent system through an analysis of strategic partnership with communities. This entails that questions of consent, protocol, attribution and benefit are distinct from copyright questions. Comparative reference is India and not Pakistan.

In 2024, Wadhwa and Chowdhary have developed a framework for the preservation of minor language literature and development of digital capacity in Sindhi libraries of Pakistan (Wadhwa & Chowdhary, 2024). The qualitative case study by Taj and Gala on Digitization of Handwritten Manuscripts focuses on the digitization activities in India and reports on a wide range of projects, conducted through interviews, observation, and analysis of library records, which demonstrate a significant level of work in this area (Taj & Gala, 2024). These works show institutional and linguistic aspects but do not make comparisons between statutory rights and the governance of AI in the two states. Are significant when they highlight the ways that collection practice and language politics influence discoverability and when they remind us that common regional history does not necessarily lead to similar policies or community claims.

Last but not least, the use of AI-generated historical data makes archival integrity and responsibility more challenging. In the absence of source attribution, version histories, and human review, users may assume that the generated interpretation is archival evidence.Users may confuse generated interpretation with archival evidence in the absence of source attribution, version histories, and human review. Authenticity and interpretive bias are described in the current heritage ethics literature, while misinformation and verifiability are described in the archival literature (Bushey, 2023; Tiribelli et al., 2024). These do not only relate to copyright issues, but the responsibility might be of the developer, custodial institutions, researchers and publishers in accordance with control, disclosure and applicable law.

The literature generally justifies a layered approach to the protection of provenance records, disclosure of synthetic material, verifiable citation, and procedures for corrections, with little attention given to the legal allocation of the protections. As a body, scholarship highlights the differences between rights, risks of extraction and epistemic bias, and provides institutional case studies from Pakistan and India. But the examined literature does not often bring together the processes of AI processing, the six categories of legal interest, the legal processes of documents and community authority and finally, Pakistan's role in international governance. The defensible gap is not an absence of relevant research, it is integrative: this study can test how these strands interact through doctrinal analysis and documented institutional practice, and can do so without assuming legal equivalence in a comparative fashion without India.

Judicial decisions concerning copyright, moral rights, and heritage conservation provide useful legal perspectives for analysing the protection of cultural materials in digital environments. In India, Eastern Book Company v. D.B. Modak, (2008) I SCC I, considered originality in copyright law, while Amar Nath Sehgal v. Union of India, 117 (2005) DLT 717, addressed moral rights and the integrity of an artwork. In Pakistan, heritage-conservation disputes may provide a separate legal context for understanding the responsibilities associated with protecting cultural property. These authorities

should, however, be applied within the limits of their actual holdings: they do not by themselves establish a comprehensive legal framework for AI-generated cultural heritage or cross-border digital archives. Their relevance lies in helping identify the distinct legal interests involved, including originality, attribution, integrity, preservation, and institutional responsibility. Their application to AI-mediated heritage therefore requires careful analysis of the relevant facts, statutory provisions, and jurisdiction.

Clear policies on access can help translate broader international principles into practical safeguards for digital cultural heritage in Pakistan and India.

Rights category	Core question	Evidence to inspect	Important limitation
Physical ownership / custody	Who owns or holds the original object, and under what authority?	Acquisition registers, transfer records, inventories, custody agreements	Custody does not itself prove copyright or permission for AI training.
Copyright	Is the underlying work or digital reproduction protected, and who holds relevant rights?	Applicable statute, rights statements, creator/photographer details, licence	Protection depends on law, originality, term, and the specific work.
Database rights	Are there legally recognized restrictions on extraction or reuse of a collection/database?	Database terms, collection structure, investment evidence, applicable law	Do not assume a separate database right exists in every jurisdiction.
Contractual / access rights	What conditions govern viewing, downloading, copying, or reuse?	Terms of use, access policies, licence text, platform conditions	Contractual restrictions and copyright are distinct legal bases.
Community / cultural interests	Are attribution, consultation, sensitivity protocols, or benefit-sharing relevant?	Community protocols, consultation records, institutional policy, provenance notes	Ethical frameworks do not automatically create enforceable legal rights.
AI-related rights / accountability	Who authorizes processing, documents transformations, and corrects harmful or inaccurate outputs?	AI policy, training documentation, provenance logs, human-review and correction rules	AI use must be evidenced; digitization alone does not prove AI deployment.

Table 1. Six rights categories and documentary indicators

This publication's original contribution is an integrated analytical framework that distinguishes between six different and non-identical interests: (1) interest in possession of the physical object; (2) copyright interest in qualifying expression or reproductions; (3) database-related protection (if applicable); (4) contractual and institutional interests; (5) community and cultural interests that cannot be reduced to property; and (6) AI-related interest in training, outputs generated by AI, provenance, and accountability. These categories can overlap, but don't necessarily mean the same thing: having a right or permission in one category does not imply authority in another. The study should examine this difference, rather than take it for granted that consent to AI reuse has been granted by either digitizing or making the document public.

3. Methodology

The research in this study is a mixed doctrinal design, which integrates doctrinal legal analysis and structured examination of institutional documents and the metadata of the digital collection. The doctrinal part aims to identify and interpret the key legal sources, such as laws, regulations, judicial precedent, treaties and key institutional instruments, and to examine the implications for ownership, copyright, custodianship, access, privacy, cultural participation and responsibilities for AI for digitized heritage. The sources will be examined with regard to their authority, legal standing, jurisdiction, date, scope and relevance, but potential conflicts between the national rule and international obligations will be explored beyond relying solely on institutional policy and ethical guidance. The analysis will involve the same six questions for each international instrument: What is the text and legal nature of the instrument? What is it about? Who are the subject matter and actors of the instrument? Is it binding? How does it apply in Pakistan? How is it relevant to AI mediated heritage? What are its limits and gaps? The corpus will include the 2003 UNESCO Charter on the Preservation of the Digital Heritage, the 2015 Recommendation concerning the Preservation of, and Access to, Documentary Heritage Including in Digital Form, the 2021 Recommendation on the Ethics of Artificial Intelligence, the 1954 Hague Convention and its protocols, the 1970 UNESCO Convention, relevant WIPO and copyright instruments, human-rights instruments concerning cultural participation, expression, privacy and access to information, and AI and digital-governance frameworks, including the EU AI Act where its jurisdictional or comparative relevance can be established.

The validity of treaty status and domestic applicability will be assessed with respect to authoritative records, and recommendations that are not treaty obligations and are not required under domestic law will not be binding.

The documentary aspect will challenge legal propositions with what the institutions reveal through practice. The major ones are the Punjab Archives, Lahore, and the JATAN digital repository in India. I selected Punjab Archives because it is directly relevant to the context of Pakistan and published studies have been written about its digitization initiative and an initial set of documents to be digitised exists in the archives (Akhtar et al., 2024). There is an Indian institutional parallel that is relevant to JATAN: published research has identified it as a government-established digital repository and there are ongoing issues with documentation and access (Sharma, 2021). The cases are purposive; they have been selected for their relevance to digitised cultural memory; they are well-suited to be examined by the public in an institution or in the literature; they are suitable for an MPhil project; they are not intended as representative cases for comparison, but rather as a way to demonstrate that there is a variety of cases that can be compared. The researcher will be conducting a survey of the case's official webpages before analysis, and catalogue the collection records, terms of use, and policies for digitisation or AI, and public reports, as well as any procurement or project documents that are available. Inclusion must have a clearly identifiable institutional source; there must be a clear connection to the selected case or the governing law; there must be material for digitization, metadata, rights, access, or AI governance; and scholarly studies may provide context but should not supplant needed primary evidence. Unsupported media claims, inaccessible documents with content that cannot be verified, duplicate documents and materials of a period or scope not included in the definition will be excluded. If the material available is in electronic format, the researcher will note a fixed retrieval date and URL, and will review up to 30 records per case, stratified by material type and language; for those in which fewer than 30 material records are eligible for review, all material records will be reviewed. The number of samples, keywords used, dates searched and any deviations from the protocol will be documented.

Provenance and copyright notice, access conditions, licence, language, community attribution, AI policy, correction mechanisms, and transparency with the disclosure of transformations, source links, will be captured on a coding sheet. Codes will differentiate between explicit evidence, contrary evidence and silence and information unavailable so that the lack of evidence in a sample record is not interpreted as proof of a lack of a policy. All definitions and collection will be very similar for both cases, with a focus on identifying similarities and differences and keeping institutional and jurisdictional context. Coding will be done first on a small subset; ambiguous cases will be revisited against the source text; definitions will be recorded; an audit trail will be maintained that will allow reviews at a later time. To enable the systematic implementation of legal concepts with expert coding, one needs clear conceptualisation, operationalisation and documentation selection (Ovádek et al., 2024). No interviews are assumed or reported, interviews can only be added upon institutional ethics approval, informed consent, safe data handling and a revised protocol. Analysis will be based on publicly available resources, where possible, and take into account the terms of site access and restrictions on access, and will not repeat content or information which are sensitive or culturally restricted, nor be content with making legal conclusions rather than ethical recommendations. The limitations are that web records are incomplete or may change over time, institutional disclosure is not equally complete and transparent, language is a barrier, and the two cases selected for public scrutiny are not representative of the general population; where public metadata cannot be backed up by the others, it can only inform, not establish internal workflows, consent histories or actual use of AI. The researcher will document access dates and file dated copies of coding sheets, a source log, and will take dated screenshots or allowable copies of relevant pages (redacting personal or sensitive information if necessary) to facilitate the process of replication. This design thus does not suggest that undocumented practices do not occur, but rather implies a careful consideration of what is documented and the applicable law.

The six rights categories should be coded separately and not be grouped together under one ‘ownership' variable for transparency. These records/policies should be marked in the following way: explicit evidence, contrary evidence, not stated in the source being inspected, or unavailable/not accessible. Not stated should NOT mean that there is no policy. An evidence log should include the following information for every record: source URL, date retrieved, identifier of the item, language, wording, and the researcher's coding rationale.

4. Results

4.1 Pakistani case

The Pakistan case is Punjab Archives, Lahore, due to the fact that it is directly related to the jurisdiction of the study and there is a published case study to support an institutional digitisation programme. That study mentions that they collected

documents, manuscripts, government papers, photographs, maps and other materials, and that they identified 114,592 documents that were selected for the initial phase of digitisation, and that some of these documents were damaged, forged and torn (Akhtar et al., 2024). These findings do not create an activity of digitization, but the ownership of the copyright of the digitized surrogates, nor the terms on which they may be reused.

The documentary material available for this analysis does not contain a dated inventory of the Punjab Archives webpages, catalogues, terms of use, collection records and a coded sample of metadata. For the purposes of this report, there are no record level findings for copyright notices, licences, access restrictions, provenance fields, language, community attribution, AI disclosure or correction procedures. Neither should silence in individual records be taken as evidence of the absence of institutional policy. Published case studies do not confirm, nor do they disprove, the existence of these specific public facing provisions (Akhtar et al., 2024), as they are not included as a case study element.

This draft should therefore not be interpreted as being representative of all planned 30-records at either institution. The Results section should be modified after the collection and coding of the data have been completed, or the paper should be restated as a doctrinal and conceptual research paper and the proposed empirical protocol should be explicitly marked as future research. The appropriate legal framework in Pakistan is the Copyright Ordinance, 1962, antiquities and heritage Acts at the provincial level, the rules of archives and public records, access to information laws, privacy laws, and digital-governance laws. There is a lack of authoritative and up-to-date statutory texts or an accurate history of how they have been used to access Punjab Archives digital collections in the sources retrieved for this analysis. A secondary study on Pakistan's Copyright Ordinance, which covers licensing provisions, is not a replacement for a reference to the copyright ordinance and amendments in an authoritative legal record (Murtiza, 2020). A review of Pakistani heritage legislation makes no mention of the focus on architectural heritage and archaeological sites/monuments, and nowhere does it say that antiquities provisions are applicable to digital reproductions or AI generated material (Khan et al., 2022). In addition, studies of the right-to-information regime in Punjab highlight the 2013 legislation and how it was put into practice, but fail to clarify the governance of specific archives requests and archives records (Afzal & Yongmei, 2023; Bashir & Nisar, 2020). These texts focus on questions to ensure the legal verification of an item and not on definitive conclusions regarding the rights in or reuse of any sample of a digital item.

The international materials must also be distinguished legally and according to the subject matter. The 2003 UNESCO Charter and the 2015 Recommendation are digital- and documentary-heritage standards that provide emphasis on preservation and access; these are not obligations under the treaty (Lixinski, 2020). UNESCO's 2021 Recommendation on the Ethics of Artificial Intelligence provides a policy framework for states, unlike a case-specific licensing principle for archival collections (Rouhani, 2023). The 1954 Hague Convention on the Protection of Cultural Property in the Event of Armed Conflict and the 1970 UNESCO Convention on the Means of Prohibiting and Preventing Acts of Cultural Insecurity and of Interfering with the Free Movement of Cultural Property are not to be taken for granted to provide any information about copyright in a digital surrogate and do not provide any indication of the regulation of the use of AI by an archive in relation to cultural property (Chamberlain & Pineschi, 2024). Confirmed treaty status and domestic applicability is required prior to reporting as findings.

4.2 Comparative case: JATAN

JATAN is not to be considered as a legal or administrative equivalent, but an institutional comparator relative to India. It is a government initiative in the form of a digital repository popularly identified in published research as one that came to be in 2014 (Sharma, 2021), and has been documented as a problem area in sectoral collections management, documentation and digitization in general. The account is contextual evidence of the Indian museum sector; it does not indicate that JATAN has poor documentation, or that the rights statements, metadata, access controls or AI practices of the repository are poor. None of the evidence in this collection contains an inventory or a stratified sample of JATAN records. The same coding categories are thus not reported for JATAN as in the case of Punjab Archives.

The two cases can therefore only be compared at the level they can be supported by the available sources. The rationale for the selected digitisation program and its preservation was documented in the published research material for Punjab Archives (Akhtar et al., 2024). According to published research, JATAN was established as a repository of the Government and is situated in a context where documentation and management of museum are a wider issue for the sector (Sharma, 2021). Both sources show that there is not enough evidence to compare the copyright notices, licensing, provenance practices, language coverage, correction mechanisms, or AI policies of the institutions. Neither case is established by the sources to be the use of AI. Any such comparison would be beyond the evidence.

4.3 Governance implications and evidentiary limits

The proposed governance elements should thus be viewed as a framework for subsequent coding and not as actual governance findings or as a set of principles. Some concerns are already present in the related standards and scholarship, such as preservation, access and authenticity, which are addressed by the digital-heritage instruments (Lixinski, 2020); and participation, accessibility, reliability, dignity and responsibility, which are mentioned as some of the main ethical concerns arising from cultural-heritage AI scholarship (Pansoni et al., 2023; Tiribelli et al., 2024); and the broader human-rights and inclusion-oriented policy framework set out in the UNESCO's AI Recommendation (Rouhani, 2023). There are no indications in the available evidence that either of the cases implements these elements, nor is there any indication of how they might be implemented in the planned institutional comparison. AI-based search and translation tools may improve the discoverability of archival materials, but their outputs should remain traceable to the original documents and their institutional sources. Bilateral cooperation, compatible metadata standards, transparent access policies, and responsible data-sharing arrangements could facilitate scholarly research while respecting the legal authority of custodial institutions and the cultural interests associated with the collections.

Comparison dimension	Punjab Archives, Lahore	JATAN, India	Evidence status / next step
Digitization activity	Published study reports a digitization programme.	Published research identifies a government-created digital repository.	Verify with current official institutional sources.
Item-level provenance / metadata	Not established by evidence currently assembled.	Not established by evidence currently assembled.	Inspect a documented sample and record fields present/absent.
Copyright notices and reuse licences	Not yet coded.	Not yet coded.	Collect current terms, item-level rights statements, and licence text.
Language and community attribution	Not yet coded.	Not yet coded.	Record language coverage, attribution, and relevant protocols.
AI use or AI policy	No direct evidence established in the current draft.	No direct evidence established in the current draft.	Do not infer AI use from digitization; seek policy/project-level evidence.
Correction, human review, and accountability	Not yet coded.	Not yet coded.	Inspect policies, correction channels, and provenance/transformation disclosures.

Table 2. Pakistan–India institutional comparison matrix

This comparison matrix is a research matrix and is not a finalised finding. Fill it out only after viewing and dating institutional records and policies.

Based on the available evidence, there is no specific governance gap when it comes to AI labelling, rights allocation, human oversight, community consultation, correction, data protection and cross-border reuse, in relation to Archives or JATAN. These conclusions must be based on coded records and institutional primary materials. We see that the gap at this stage is evidential in that the published sources provide context for the digitization process and evidence of the existence and wider sectoral context of JATAN, but they do not provide the level of detail required for reporting the results of coding the sheet at the record level. The legal conclusions are also preliminary, pending review of the statutes, amendments and treaty status with authoritative sources. This separation makes it harder to overlook a document and report it as not being found.

5. Discussion

It is not evident that AI automatically transfers ownership of the cultural memory under digitization, but rather that it increases the number of rights and governance decisions surrounding digitized materials related to their access, copying, computational reuse, interpretation and control of output. However, in the cases analyzed, this evidence does not show that either of the institutions (Punjab Archives or JATAN) are using AI, nor document either institution's item-level permissions. The defensible finding is then with regard to the conditions under which AI might change rights, not an observed change in ownership. While digitization can aid in preservation and access, a digitized surrogate does not constitute a copyright finding that an institution owns the copyright in the underlying work, nor grant permission to reuse it for training or to create AI-based outputs. Issues with authorship, authenticity, and interpretive bias are also highlighted in scholarship on cultural-heritage AI (Tiribelli et al., 2024). This further emphasizes the importance of separating the concept of legal title, copyright, from the custodial responsibility, public access and community interests

of sampled records for Pakistan. This is also relevant for community rights. Publicly available access can expand access but can also take away the ability of communities to impact the sensitive materials, descriptions, or downstream uses. Community-centred digital heritage systems demonstrate that access does not lead to community control over access and use (Rayburn et al., 2024). Therefore, when designing the proposed governance process, consultation and differential access, where applicable, should be taken into account, and not assumed to be a failure. Internationally, digital-heritage standards encourage preservation, access and focus on authenticity, and the UNESCO AI Recommendation provides a policy framework as one of its recommendations does not include a licence requirement for collections (Lixinski, 2020; Rouhani, 2023). While these tools offer valuable reference points for copyright norms, they are not definitive on copyright allocations or who can authorize the use of AI processing in Pakistan and India. Whether collections or AI services are digitised, it is feasible that decisions about storage, access, reuse, and technological infrastructure will be made outside of the organization or nation where the original material is stored. This is a practical as well as legal matter.

Digital sovereignty in this instance should have to do with accountability of the capacity to set and enforce terms of stewardship, rather than with data localization. Cross-border partnerships may help increase preservation and access, but should ensure that responsibilities, the ability to reuse, attribution, security and paths for correction are made explicit. For Pakistan, a proportionate institutional response would involve a provenance inventory of the items, with machine-readable and clear reuse terms attached, and ensuring that there is a distinction between the source and the AI-generated or transformed source. All AI-driven pilots should include documentation of why the AI is being used and what it uses to build, demand that consequential outputs from the AI be reviewed by humans, include a correction path, and evaluate the privacy and community effects before deployment. The measures are not statements of existing practice but practical proposals. The evidence also allows for other reasons why there is no evidence of AI governance: institutions might not be using AI; the information might not be in a policy document; or information about AI control may not be included in policies as a statement, but rather as part of routine procedures. Limited inferences about institutional practice and community experience because of a small, purposeful evidence base, inaccessible primary materials, no interviews, and no direct evidence of the use of AI. JATAN can only be a partial representation of Pakistani archives due to institutional and national context disparities. Lastly, copyright, archival and data protection requirements could shift, and the legal analysis should be corroborated with authoritative and up-to-date sources prior to operational conclusions. The cases as a whole do not provide grounds to conclude that the institution is non-compliant nor do they demonstrate that the ownership of AI has been re-distributed. They reveal why Pakistan should be the next to come with a trustworthy documentation and rights-informed governance before the reuse with Artificial Intelligence becomes commonplace.

Instrument / framework	Primary relevance	Legal character to distinguish	Verification needed
UNESCO 2003 Charter on the Preservation of the Digital Heritage	Digital preservation and access	Charter / normative guidance; do not describe as a binding treaty without legal basis	Official text, scope, and relevance to institutional practice.
UNESCO 2015 Recommendation on Documentary Heritage	Preservation and access to documentary heritage, including digital form	Recommendation; generally non-binding	Official text and how it informs policy analysis.
UNESCO 2021 Recommendation on the Ethics of AI	Human rights, inclusion, responsibility, and AI governance	Recommendation; policy framework, not item-specific licence	Official text and Pakistan/India policy relevance.
1954 Hague Convention and protocols	Protection of cultural property in armed conflict	Treaty obligations depend on participation and applicable provisions	Treaty status, protocols, domestic implementation, and scope.
1970 UNESCO Convention	Illicit import, export, and transfer of cultural property	Treaty; scope is not automatically digital copyright or AI training	Treaty status and relevance to the specific legal question.
WIPO / copyright instruments and national copyright law	Copyright, related rights, and possible AI-related questions	Binding status and domestic effect vary by instrument and jurisdiction	Applicable provisions, amendments, exceptions, and authoritative legal texts.

Table 3. International instruments and legal relevance

Legal status and policy relevance are separated in the above table. Treaty participation, domestic implementation and amendments, as well as the scope of each instrument, should be checked before submission against authoritative sources, this table is a research aid and not a substitute for verification.

6. Conclusion

The significance of this study is the understanding that the potential of AI for digital cultural memory does not depend on the ability to demonstrate that the ownership of this memory is being transferred, but that it raises further questions and decisions about access, reuse, interpretation, and stewardship. This governance-based finding is based on evidence available to them for the cases examined, and is not a finding of institutional use of AI or permissions at the item level. Before permitting computational reuse, it is crucial that heritage institutions document provenance, rights, access requirements, sensitivity, and uncertainty at the record-level in ways that are verifiable and documented. Stewardship would be more accountable if reuse terms were clear, if records of inputs and uses to the AI were transparent, if the AI was subject to human review, if mechanisms were built in for humans to correct the AI, and if privacy and community impacts were considered beforehand. These recommendations provide a proportional starting point for Pakistan and can serve as a benchmark against India, and also acknowledge that these cases are based on unique institutional and legal contexts. The results are limited by the small sample size, the fact that primary sources were not always available, the lack of interviews and the absence of direct evidence of AI usage, with legal changes over time ensuring the conclusions are valid for the period under study and can only be confirmed if there are authoritative documents. However, further studies should focus on the institutional records and permissions themselves, research the community experience by interviewing, and examine the implementation of the policies in practice. This type of work can be used to see if the proposed safeguards are being met in reality and whether the cross-border arrangements are clarifying responsibility for the safeguards. For now, the key takeaway is not that ownership has shifted, but that documenting and managing rights should come before the regular reusing and reuse of AI. Therefore, documentation and governance are prerequisites for computational reuse rather than remedies to be implemented after AI-powered services have already impacted access and interpretation.
Offering a six-category framework for evaluating whether institutional documentation distinguishes between physical possession, copyright, access permissions, community interests, and AI accountability, the study makes a convincing analytical and methodological contribution. It is not until the planned record level analysis is completed and the applicable legal texts are verified, that its empirical contribution can be limited.

References

[1] Afzal, J. and Yongmei, C. (2023) 'Federal and provincial legislation regarding "Right to Information" for good governance in Pakistan', *Discover Global Society*, 1(1). Available at: https://doi.org/10.1007/s44282-023-00015-7.

[2] Akhtar, M.N., Tariq, M.A., Batool, S. and Ali, S. (2024) 'Digitization process of archives: a case study of the Punjab Archives, Lahore', *Qlantic Journal of Social Sciences and Humanities*, 5(1), pp. 240–247. Available at: https://doi.org/10.55737/qjssh.391567320.

[3] Bashir, M. and Nisar, M.A. (2020) 'Expectation versus reality: political expediency and implementation of Right to Information laws', *Public Administration Quarterly*, 44(1), pp. 3–30. Available at: https://doi.org/10.1177/073491492004400101.

[4] Bushey, J. (2023) 'AI-generated images as an emergent record format'. In: *2023 IEEE International Conference on Big Data (BigData)*, pp. 2020–2031. Available at: https://doi.org/10.1109/bigdata59044.2023.10386946.

[5] Carroll, S.R., Garba, I., Rodríguez, Ó.L.F., Holbrook, J., Lovett, R., Materechera, S.A., Parsons, M.A., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J., Anderson, J. and Hudson, M. (2020) 'The CARE Principles for Indigenous Data Governance', *Data Science Journal*, 19. Available at: https://doi.org/10.5334/dsj-2020-043.

[6] Chakraborty, D. (2023) 'Digitizing the other half of the Indian subcontinent's Partition history', *Scholarly Editing*, 40. Available at: https://doi.org/10.55520/qnvf6mmf.

[7] Chamberlain, K. and Pineschi, L. (2024) 'The 1970 UNESCO and 1995 UNIDROIT Conventions and the 1954 Hague Convention and its Protocols'. In: *Oxford University Press eBooks*, pp. 727–748. Oxford: Oxford University Press. Available at: https://doi.org/10.1093/law/9780192846884.003.0047.

[8] Ghosh, A.K. (2024) *Recovering knowledge commons for the Global South*, 5(1). Available at: https://doi.org/10.55492/dhasa.v5i1.5011.

[9] Gîrbacia, F. (2024) 'An analysis of research trends for using artificial intelligence in cultural heritage', *Electronics*, 13(18), article 3738. Available at: https://doi.org/10.3390/electronics13183738.

[10] Helm, P., Bella, G., Koch, G. and Giunchiglia, F. (2024) 'Diversity and language technology: how language modeling bias causes epistemic injustice', *Ethics and Information Technology*, 26(1). Available at: https://doi.org/10.1007/s10676-023-09742-6.

[11] Hussain, F., Komal, S., Haq, M.I. ul, Rehman, N.U., Anwar, N., Hashmi, T.M. and Khan, M.S. (2025) 'Language policy in the age of artificial intelligence: implications for Urdu and Punjabi language in Pakistan', *Pakistan Journal of Life and Social Sciences (PJLSS)*, 23(2). Available at: https://doi.org/10.57239/pjlss-2025-23.2.00254.

[12] Israel, M.J. and Amer, A. (2022) 'Rethinking data infrastructure and its ethical implications in the face of automated digital content generation', *AI and Ethics*, 3(2), pp. 427–439. Available at: https://doi.org/10.1007/s43681-022-00169-1.

[13] Jaswal, B.A. (2016) 'Digitization of archival collections by libraries in Pakistan: issues, strategies, challenges and opportunities', *Pakistan Journal of Information Management and Libraries*, 17, pp. 97–108. Available at: https://doi.org/10.47657/201617897.

[14] Khan, N.A., Nuti, C., Monti, G. and Micheli, M. (2022) 'Critical review of Pakistani current legislation on sustainable protection of cultural heritage', *Sustainability*, 14(6), article 3633. Available at: https://doi.org/10.3390/su14063633.

[15] Kuan, L. and Majid, S. (2025) 'Digitization of cultural heritage in China and Pakistan: a comparative analysis under the Belt and Road Initiative', *Journal of Political Stability Archive*, 3(3), pp. 799–807. Available at: https://doi.org/10.63468/jpsa.3.3.53.

[16] Lixinski, L. (2020) 'Digital heritage surrogates, decolonization, and international law: restitution, control, and the creation of value as reparations and emancipation', *Santander Art and Culture Law Review*, pp. 65–86. Available at: https://doi.org/10.4467/2450050xsnr.20.011.13014.

[17] Murtiza, G. (2020) 'Does Pakistan's intellectual property law conform to international intellectual property law? An overview', *Journal of Arts & Social Sciences (JASS)*, 7(2), pp. 245–255. Available at: https://doi.org/10.46662/jass-vol7-iss2-2020(245-255).

[18] Oguamanam, C. (2025) 'Information systems and digitization of traditional knowledge: trends in cultural heritage and memory institutions and the WIPO Genetic Resources Treaty', *The Journal of World Intellectual Property*, 29(1), pp. 130–161. Available at: https://doi.org/10.1111/jwip.70005.

[19] Oruç, P. (2025) *Digitising cultural heritage*. Hart Publishing. Available at: https://doi.org/10.5040/9781509959310.

[20] Ovádek, M., Schroeder, P.A. and Zglinski, J. (2024) 'Where law meets data: a practical guide to expert coding in legal research', *European Law Open*, 3(4), pp. 820–848. Available at: https://doi.org/10.1017/elo.2024.23.

[21] Palladino, C. and Bodard, G. (2023) *Can't touch this: digital approaches to materiality in cultural heritage*. Ubiquity Press. Available at: https://doi.org/10.5334/bcv.

[22] Pansoni, S., Tiribelli, S., Paolanti, M., Frontoni, E. and Giovanola, B. (2023) *Design of an ethical framework for artificial intelligence in cultural heritage*, pp. 1–5. In: *Proceedings of the 2023 IEEE Conference on Ethics in Engineering, Science and Technology (ETHICS)*. Available at: https://doi.org/10.1109/ethics57328.2023.10155020.

[23]Rayburn, A., Punzalan, R.L. and Thomer, A.K. (2024) 'Persisting through friction: growing a community-driven knowledge infrastructure', *Archival Science*, 24(1), pp. 61–82. Available at: https://doi.org/10.1007/s10502-023-09427-5.

[24] Rouhani, B. (2023) 'Ethically digital: contested cultural heritage in digital context', *Indiana Magazine of History (Indiana University)*, 7(1), pp. 1–16. Available at: https://doi.org/10.14434/sdh.v7i1.35741.

[25] Sharma, K. (2021) 'Museum collection storage in India: a decade in review', *Museum International*, 73, pp. 144–155. Available at: https://doi.org/10.1080/13500775.2021.1956759.

[26] Taj, A. and Gala, B. (2024) 'Digitization projects for cultural heritage materials'. In: *Advances in Library and Information Science (ALIS) book series*, pp. 238–255. SAGE Publishing. Available at: https://doi.org/10.4018/979-8-3693-2782-1.ch013.

[27] Tiribelli, S., Pansoni, S., Frontoni, E. and Giovanola, B. (2024) 'Ethics of artificial intelligence for cultural heritage: opportunities and challenges', *U-PAD Unimc – Open Digital Publications (University of Macerata)*, 5(3), pp. 293–305. Available at: https://doi.org/10.1109/tts.2024.3432407.

[28] Wadhwa, S. and Chowdhary, R. (2024) 'Digital technology for literature revitalization: a framework for Sindhi libraries in India', *Preservation, Digital Technology & Culture*, 53(4), pp. 231–240. Available at: https://doi.org/10.1515/pdtc-2024-0030.

[29] Wagner, A. and de Clippele, M.-S. (2023) 'Safeguarding cultural heritage in the digital era – a critical challenge', *International Journal for the Semiotics of Law – Revue Internationale de Sémiotique Juridique*, 36(5), pp. 1915–1923. Available at: https://doi.org/10.1007/s11196-023-10040-z.