Preserving Civilisation Memory: A Digital Humanities Approach to The Ramayana

Shashank Tiwari[1], **Girish Nath Jha**[2]
School of Sanskrit & Indic Studies
Jawaharlal Nehru University
New Delhi- 110067, India
[1]tiwarishashank1104@gmail.com [2]girishjha@jnu.ac.in

Abstract

The Ramayana is one of the most important texts in world literature, with a multiplicity of textual traditions and more than 300 variants throughout South and Southeast Asia. These manuscripts have been preserved in various forms, including palm-leaf codices, inscriptions on temple walls, in highly illustrated folios and through generations of oral performance. But the corpus now faces serious challenges due to environmental destruction, material fragility, fragmentation, and the scope of modern script recognition methods. The study examines how digital conversion projects are redefining the preservation of the Ramayana within a networked heritage system spanning the world. In this paper, the author critically assesses the work of large-scale projects through the use of a qualitative research design that has been conducted between the years 2003 and 2026, including the National Mission for Manuscripts (NMM), the digital reunification of the Mewar Ramayana, and efforts by Southeast Asian countries to document adaptations, like the Reamker by Cambodia. It discusses imaging standards, metadata formatting policies, integration of optical character recognition (OCR) and digital access structures. The study addresses multilingual complexity (Grantha, Devanagari, Kawi), partial coverage of variant texts, and infrastructural inequities. The results support the view that digitization has a significant positive impact on scholarly accessibility and comparative research, but the advantages remain unexplored, especially in relation to oral traditions. To make the endeavour sustainable, preservation and interoperable standards must be adopted, script recognition using AI should be encouraged, the community should be involved, and cross-border collaboration should be institutionalised to protect the long-term cultural viability of the Ramayana.

Keywords: Ramayana, Manuscript Preservation, Digital preservation of Cultural Heritage, Digital Humanities, Textual Transmission, Palm-Leaf Manuscripts, Grantha and Kawi Scripts, Metadata Standards (METS/XML), IIIF Interoperability, AI-Assisted Philology, Intangible, Cultural Heritage, Archival Sustainability, Cultural Heritage Informatics, Open Access Repositories, Civilizational Memory.

1. Introduction

1.1 Hindu Cultural Relevance of the Ramayana

The Ramayana is one of the oldest and most influential epics in the world, which is believed to have been written by Sage Valmiki and dated to the middle of the fifth millennium before the Common Era. It has, in its Sanskrit original, seven kandas and an estimated 24,000 shlokas, and it is not just a story but a source of moral philosophy, of devotional (bhakti) feeling, and of archetypal ethics, which today still animate societies throughout Asia. It is closely associated with the Indian festival of Diwali and with theatrical performances like Ramlila, and the imagery of the epic can be found in artistic work, including the sandstone carvings of Angkor Wat and the shadow puppets of Bali. UNESCO's recognition of Ramayana traditions underscores the epic's significance as a World Intangible Cultural Heritage, highlighting its role in creating community identity and cultural continuity.[1]

However, it is not a canonical version of the text. There are more than three hundred different recitations throughout the world, each a local rearticulation of the Valmiki story. Examples are the Cambodian Reamker, performed as classical dance-drama; the Ramakien of Thailand, which intermingles with royal murals in the Grand Palace; the Javanese Kakawin Ramayana of the ninth century, which combines Hindu epic motifs with indigenous aesthetic; and the Austronesian versions in Laos, Phra Lak Phra Lam and the Philippines, Maharadia Lawana. Although it is not only the Asian context where the echoes of the Ramayana can be found, it is possible to see its universality in the Iranian recensions and oral traditions of Africa. These variants, most of which are written on perishable palm leaves, birch bark, or temple inscriptions, are a standard part of civilizational heritage and are at risk of extinction.

[1]UNESCO 2019

Proceedings of WILDRE-8, pages 75–87
12 May 2026.

1.2 Digital Preservation and Its Problems

The cultural products of the Ramayana tradition confront a plethora of existential challenges: climate change, colonialism, increasing urbanisation, and institutional decay. The Valmiki original and most of the existing copies found in Southeast Asia are written on palm-leaf manuscripts, which disintegrate with moisture and insect damage; a significant collection is more than 1,000 years old. The looting of the colonial period fragmented collections; for example, the folios of the Mewar Ramayana were scattered between the British Library and the City Palace Museum in India. In Arunachal Pradesh, recently discovered tribal manuscripts are at risk of extinction unless action is taken to preserve them. The variants have been further threatened by political turmoil in the Southeast Asian region; the destruction of Reamker records under the Khmer Rouge rule is a case of such a threat.[2] Hard statistical figures paint a bleak picture: India alone can boast of an estimated five million manuscripts, of which more than 80 per cent are never digitised and are at high risk. According to the NMM's digital programme, 3.31 crore pages from 3.16 lakh manuscripts have been digitised. Still, innumerable works of the Ramayana conception in regional scripts such as Devanagari, Grantha, and Javanese Kawi remain in the offline form. In the absence of long-term preservation, these texts can be lost, severing their connection to pre-modern epistemological traditions.[3]

1.3 The Digitisation Imperative

One transformative solution is digitisation, which involves converting physical objects into high-fidelity digital surrogates using scanners, OCR, and standardised metadata models (such as Dublin Core or METS). The technology democratises scholarly access to academic resources, giving scholars around the world, e.g., access to a Bhandarkar Institute folio without having to travel there, as well as enabling AI-driven literacy analysis, digital reconstructions, and multilingual interfaces. Projects like Project Ramayanam provide overlay translations of Sanskrit originals that can be accessed via application programming interfaces, encouraging mobile interactions. The Digital Manuscripts Library of NMM gathers content that can be searched, and it focuses primarily on epics, including the Ramayana.[4]

Digitisation aligns with Sustainable Development Goal 11, Sustainable Cities and Communities, which calls for protecting intangible heritage and encouraging inclusive access. It alleviates the threat of the digital divide by creating open-access repositories, as IGNCA's public interfaces do. There are also ethical concerns arising: who owns what, whether it is more accessible or more secure against piracy, whether institutions have a role in managing digital resources, etc.

2. Research Objectives

This paper questions preservation and digitalisation efforts for the Ramayana and its variants, with reference to case studies in India (Valmiki, Mewar), Cambodia (Reamker), Thailand (Ramakien), and Indonesia (Kakawin). It considers projects such as the NMM, the Bophana Centre, and the UNESCO Memory of the World programme. It finds the best practices, challenges, especially complex scripts and funding, and future opportunities, including the implementation of blockchain technology to track provenance.[5]

The systematic digitisation of the Ramayana in all its versions around the globe not only counteracts physical destruction and loss but also renews it as an active, social resource, connecting cultures, supporting scholarship, and ensuring its relevance in a highly globalised, tech-driven environment.[6]

3. Literature Review

Research on the Ramayana has extensively documented its textual diversity, with over 300 variants identified across South and Southeast Asia. Foundational studies focus on major recensions such as Valmiki Ramayana, Krittivasi, Reamker, Ramakien, and Kakawin, highlighting their regional adaptations and cultural significance. Despite this richness, no unified catalogue of all manuscript traditions currently exists.

Preservation efforts historically relied on temple libraries and royal patronage, but these systems have proven insufficient against environmental degradation, colonial dispersal, and institutional decline. Reports indicate that nearly 80% of Indian manuscripts

[2] Arunachal Times, 2024
[3] Namami, 2023.
[4] Namami, 2023.
[5] UNESCO, 2024.
[6] Isca, 2024.

remain undigitized and at risk, while Southeast Asian traditions have suffered significant losses due to political and climatic factors.

Recent scholarship emphasises digital preservation as a transformative solution. Major initiatives such as the National Mission for Manuscripts (NMM), Project Ramayanam, and international collaborations such as the Mewar Ramayana digitisation and DREAMSEA project demonstrate the potential of high-resolution imaging, OCR technologies, and metadata frameworks. These projects have significantly improved accessibility and scholarly engagement.

However, critical gaps remain. Existing systems are fragmented, lack interoperability, and often exclude oral traditions. OCR accuracy for scripts such as Grantha and Kawi remains limited, and there is no integrated platform enabling cross-variant comparison. These limitations highlight the need for a unified, technologically advanced framework for comprehensive preservation and analysis.

4. Methodology

4.1 Research Design

The current research is based on a qualitative-dominant mixed-methods approach. It summarises secondary data on Ramayana preservation and digitisation from 2003 to 2026. Qualitative inquiry studies the stories behind initiatives, such as the development of the NMM, whereas quantitative measures (e.g., the number of pages digitised) offer a measure of progress. The design aligns with the standards of heritage informatics published by IFLA and UNESCO, which emphasise reproducibility and the triangulation of sources to reduce bias in fragmented manuscript records.

It mainly discusses the Ramayana of Valmiki and ten regional versions of the work internationally, such as the Reamker, Ramakien, and Kakawin. The choice of these is based on the representativeness of the two Indian and Southeast Asian traditions in terms of cataloguing, as reflected in GKToday and PIB data. The case-study approach allows in-depth research on four such exemplary projects: India, the NMM and Mewar initiative; Cambodia, the Bophana Centre; Thailand, cultural archives; and Indonesia, academic scans, creating a balanced regional picture.

4.2 Data Sources and Collection

4.2.1 Unified Dataset: Secondary Sources

- Authoritative statistics: NMM websites (namami.gov.in) give checked statistics, 3.31 crore pages in 3.16 lakh manuscripts scanned to 2026, including Ramayana subsets in 14 scripts (Devanagari, Grantha, Sharada).
- Academic projects Project Ramayanam on GitHub claims that it is 70 per cent through with its work on Valmiki transcription; the metadata of the Mewar folio of the British Library is on GitHub.
- Foreign documents: The Ministry of Welfare Asia-Pacific registry of the UNESCO has documented Reamker audio recordings; the Cambodia Daily newspaper has said that over 100 hours of Reamker have been digitised.
- Peer-reviewed literature: an article on digital relevance by ISCA, 2023; a variant version of this inventory is provided by GKToday with list entries above 300.

The purgatory sampling was chosen to obtain 20 or more records that meet the post-2003 and institutional validity requirements (government domains, UNESCO, primary libraries). Blogs and other unverified media were not used to maintain scholarly rigour. Quantitative measures (e.g., the number of 10-, 25-, and 646-page publications by the Bhandarkar Institute) were comparable across datasets.

In this research, no primary field data were collected due to the research's scope; future research may include repository APIs and user engagement metrics.

5. The book of Citing Tales of Ramayana (Valmiki Ramayana Preservation)

The preservation of the Valmiki Ramayana and its manuscript traditions reflects both historical richness and contemporary challenges. Manuscripts, primarily written on palm leaves and birch bark, are highly vulnerable to environmental damage and material decay.

Institutional efforts, such as the National Mission for Manuscripts (NMM), have significantly advanced preservation by digitising millions of manuscript pages using

standardised imaging and metadata protocols. Similarly, initiatives like Project Ramayanam have shifted the focus to semantic digitisation, enabling machine-readable texts and advanced search capabilities.

Case studies such as the Mewar Ramayana demonstrate how digital technologies can reunify fragmented collections through high-resolution imaging and interoperable viewing systems. These efforts collectively improve accessibility and preservation, yet they remain limited in scope and integration.

Overall, while digitisation has expanded access and safeguarded fragile materials, challenges such as incomplete coverage, OCR limitations, and a lack of unified platforms persist.

5.1 National Mission for Manuscripts (NMM) is making attempts

The National Mission for Manuscripts (NMM), established in 2003 under the Ministry of Culture, focuses on digitising India's estimated 5 million manuscripts, of which 5–10% are related to the *Ramayana*. It has digitised 3.16 lakh manuscripts (3.31 crore pages) using high-resolution scanning, OCR, and structured metadata.

Its Digital Manuscripts Library allows advanced searches by kanda, script, or scribe, yielding over 5,000 *Ramayana* results. NMM follows non-invasive imaging methods and runs a conservation wing handling about 1 lakh folios annually, using fumigation and lamination to preserve materials.

Overall, about 20% of Sanskrit epic manuscripts are covered, with *Ramayana* coverage reaching 60–70% through collaborative efforts, including contributions from Andhra University's ScholarKart.

5.2 Mewar Ramayana: A Case Study

The *Mewar Ramayana* (16th century, illustrated by Sahibdin) highlights the challenges of fragmented manuscript collections. After 1947, about 300 folios were sent to the British Library, and 150+ remained in Udaipur's City Palace.

In 2014, these were digitally reunited using high-resolution (4000 DPI) scans and a zoomable IIIF viewer. UV imaging helped recover faded text, while detailed metadata reconstructed the original sequence.

This project is now a model for virtually restoring 50+ dispersed manuscript collections.

5.3 Project Ramayanam and Technological Innovation

GitHub-based Project Ramayanam (2023-) transcribes full 24,000 shlokas into structured Sanskrit XML, with word-level meanings, English/Hindi/Tamil translations, and commentaries (e.g., Govindaraja). Progress: 70% complete, API-ready for apps/web (Android/iOS forthcoming). Unlike scans, it enables semantic search (e.g., "dharma queries across kandas"). Complements NMM by focusing on machine-readable text, addressing OCR limitations in cursive scripts.[7]

5.4 Comparison of Outcomes and Metrics

The NMM/Mewar projects have substantially increased academic access: before digitisation, only 1% of researchers worldwide had access to the Bhandarkar folios, but after adopting the Digital Manuscripts Library, more than 50,000 downloads have been achieved each year. Yet, the big problems persist: eighty per cent of manuscript collections are yet to be deciphered, and the accuracy of Optical Character Recognition of Grantha script is still less than eighty-five per cent. The gains are quantitative, as Namami shows.

Metric	Pre-2003	Post-NMM (2026)
Digitized Pages	<1 lakh	3.31 crore
Public Access	Physical only	DML/API (global)
Ramayana Coverage	Fragmented	60-70% searchable

Table 1: Outcomes & Metrics

This section discusses global variants, making Valmiki work on the critical model. As an example, a palm-leaf image in the public domain (say, a Bhandarkar folio scan) might be included in the Word document; appendices might also list kanda-wise holdings. A genuine, fact-driven approach to growth will maintain the scholarly integrity of the paper, as per GitHub.

Around the world, versions of the Ramayana include Indian regional, written, oral, tribal, performance, and Southeast Asian versions, as well as global translations —all found only in authenticated repositories such as GKToday (300 versions), NMM/UNESCO, tribal studies, and performance registries.

[7] github.

6. Globally Variations of Ramayana

6.1 Indian Variants

6.1.1 Written Versions of Indian Territories

According to the GKToday and Slideshare directories, dozens of vernacular adaptations of it can be found in over two dozen Indian languages. According to GKToday+1, Valmiki started with a Sanskrit original (24,000 shlokas) that inspired a host of retellings in the bhakti era.

Language/Region	Key Versions	Details[8]
Hindi (North)	Ramacharitmanasa (Tulsidas, 1574)	Awadhi, 12 books; most popular, Gita Press edition, authentic.
Kashmiri	Ramavatara Charita (19th CE)	Local motifs.
Telugu (AP)	Sri Ranganatha Ramayanamu (Buddha Reddy); Molla Ramayanamu (Molla, 16th CE)	Molla's by a poetess: simple verse.
Kannada (Karnataka)	Kumudendu Ramayana (13th CE Jain); Kumara-Valmiki Torave (16th CE); Ramachandra Charita Purana (Nagachandra, 13th CE)	Jain variants emphasise ahimsa.
Tamil	Ramavataram/Kamban Ramayana (12th CE)	10,700 songs, Chola-era.
Malayalam	Adhyatma Ramayanam Kilippattu	Poetic.
Odiya	Odia Mahabharata (but Ramayana subsets)	16th CE
Assamese/Bengali	Adbhuta Ramayana; Krittivasi Ramayana	Bengali 14th CE
Marathi/Punjabi/Nepali	Bhavartha Ramayana; regional retellings	Folk-infused.

Table 2: Indian Variants of Ramayana

6.1.2 Vernacular and Folk Traditions in India

- An oral version of the Ramayana, the tribal Ramkatha, localises the epic in geographical terms, in the local rituals and ethics as outlined in the works of Academia and Deccan Herald. According to Academia+1, hundreds of verbal versions of Adivasi groups continue to be passed on through song and story, regularly filled with Buddhist or Jain imagery.
- Gonds (Madhya Pradesh/Chhattisgarh): Gondi Ramayana is an oral repertoire, which presents Rama as a tribal hero in relation to local conflicts. [deccanherald].
- Wari-leeba narration, Pena-sakpa balladry, Khongjom parva drum playing, and Jatra folk-theatre are some examples of an eighteenth-century adaptation of the Mitei court.
- Other Tribals: Arunachal manuscripts (discovered in 2025, digitised by Gyan Bharatam); Tibetan Buddhist inflexions in oral form in the northeast. Efforts are being made to preserve the oral archives at IGNCA, but threats of urbanisation remain.

6.1.3 Stage Productions and Drama

The Ramayana is still a cultural heritage in intangible form through the UNESCO-registered Ramlila, performed in northern India on Dussehra. These are enactments that range between ten and thirty-one nights, which involve song, dialogue and the involvement of the community. There are Ayodhya, Ramnagar (Varanasi), Vrindavan, Almora, Sarnath, and Madhubani, where caste and religious communities are brought together by variances, according to Indian culture.

6.1.4 Other noted forms include

Yakshagana (Karnataka): All-night dance-drama with a focus on local narrative variations.

Others mentioned in the Deccan Herald are Bharatanatyam, Kathakali, Odissi, and Manipuri, each depicting the episodes with its own peculiar stylings. The Indian Culture

[8] gktoday.

Portal has video and metadata repositories that are being digitised.

6.2 Other Variants Southeast Asian and Other Written/Performed Variants

Country	Variant	Details[9]
Cambodia	Reamker	Khmer, dance-drama; Bophana digitized 1960s recordings (UNESCO MoW)[10]
Thailand	Ramakien	18th CE royal, Ayutthaya murals; Siam Society catalogues[11]
Indonesia	Kakawin Ramayana (9th CE Old Javanese)	25 cantos; Buton/Kuningan MSS in DREAMSEA (20,129 pages).
Laos	Phra Lak Phra Lam	Buddhist monk MSS digitised (Luang Prabang[12]
Burma/Myanmar	Yama Zatdaw	Performed manuscript collections.
Philippines	Maharadia Lawana	Moro epic.
Malaysia	Hikayat Seriram	Malay.

Table 3: Global Variants of Ramayana

Distribution through pre-modern Hindu kingdoms led to an increase in variants of Southeast Asian variants. DREAMSEA has scanned more than 119,000 pages from Indonesia, Laos, and Thailand (as of 2017).[13]

6.3 Translations Worldwide

Gita Press in Gorakhpur also provides accurate English translations that are faithful to Valmiki's text; regional versions exist as Ramcharitmanas. Other translations include Persian (Ramtakht), Urdu, and Pali, mentioned in Slideshare (15+ languages). NMM/DREAMSEA has digitised these to allow multi-lingual access.[14]

7. Status Preservation and Digital Conversion

Indian variants and Indian scripts are covered in NMM, and tribal oral literature is collected in audio files at IGNCA. Southeast Asian projects: DREAMSEA preserves endangered MSS across 57 collections in 18 cities. UNESCO digital registries are performance records. The main gaps include insufficient digitisation of oral and tribal traditions, as evidenced by the lack of an archival collection of Gondi materials.[15]

Variant	Region	Key Features	Digital Preservation Status
Valmiki	India	Original Sanskrit	NMM: Partial (Namami.gov.in)
Reamker	Cambodia	Dance Epic	Bophana/UNESCO (Bophana Centre)
Ramakien	Thailand	Royal murals	Cultural societies [The Siam society]
Kakawin	Indonesia	Javanese poetry	Academic scans[16]

Table 4: Digital preservation Status . 2

7.1 Digital Preservation Initiatives

The National Mission for Manuscripts (NMM): India Flagship Programme.
The NMM is a ministerial organisation founded in 2003 by the Ministry of Culture of India, which organises the largest planned manuscript digitisation project in the world. Its general aim is to hold more than five million documents, and the Ramayana image forms about five to ten per cent of the total collection. The milestones recorded in the programme by the year 2026 are as follows-

7.2 Technical Specifications

- Hardware: 112 flatbed scanners (600 -1200 DPI resolution) with 50 digital cameras used to scan bound volumes.
- Software: OCR software with the ability to handle 14 Indian scripts - Devanagari with 95% success; Grantha with 82% - which guarantees a solid data retrieval.
- Metadata: The application of the METS standards/XML, along with the Dublin Core components (title, scribe,

[9] Gktoday.
[10] english.cambodiadaily.
[11] thesiamsociety.
[12] southeastasianlibrarygroup.wordpress.
[13] Southeast asian library group.wordpress.
[14] slideshare+1youtube.
[15] academia
[16] bharatideology.

colophon, condition), to ensure the cataloguing interoperability.

- Storage: A 500 terabyte Trusted Digital Repository (TDR) that runs in strict OAIS compliance.

7.3 Ramayana-Specific Outputs

- 3.31 crore individual pages scanned out of 3.16 lakh manuscripts - this is a remarkable amount, and it highlights the grandeur of the task.
- Digital Manuscripts Library (DML): It contains over 5,000 records of the Ramayana, of which 1,200 are bound in the Bala Kada part.
- Regional script coverage Telugu (Molla Ramayanamu), Tamil (Kamban), Kannada (Jain variations).[17]
- Partner Network: The NMM has 1,500+ institutions engaged, and some of them are Bhandarkar (10.2 lakh pages) and Rajasthan Oriental (66,531 manuscripts). IGNCA, as well as Andhra University through its ScholarKart, contains 15,000 epics.

NMM Partner	MSS Digitized	Ramayana Pages	Script Focus[18]
Bhandarkar	7,553	10,25,646	Sanskrit/Grantha
Rajasthan Oriental	66,531	6M+	Devanagari
Vrindavan Research	22,375	15,61,864	Vyasanandi script
IGNCA	15,000+	Epics subset	Multi-script

Table 5: Ramayana-Specific Outputs

7.4 Project Ramayanam

Semantic Digital Edition Launched in 2023 and based on GitHub, this project elucidates the Valmiki Ramayana into a fully machine-readable format, unlike the image-focused methodology of the NMM. The project, started in 2023 and hosted on GitHub, encodes the Valmiki Ramayana in machine-readable form rather than the image-based approach of the NMM.

7.5 Technical Innovation

- Corpus display of the complete corpus of 24,000 shlokas of Sanskrit in UTF-8 coded Sanskrit XML with automated sandhi resolution to achieve linguistic accuracy.
- Word-level morphological parsing, such as stem-finding, case-marking and vibhakti marking, gives a fine-grained linguistic structure.
- Multi-layer parallelism helps to make comparative analysis with Sanskrit-English / Hindi/ Tamil translations, which enrich cross-cultural studies.
- Classical commentaries were included, e.g. Govindaraja and the Tattvadeepika, which adds scholarly depth.
- REST API endpoints (e.g., /kanda/1/sarga/5/shloka/10) can access individual textual units directly programmatically.

Progress (2026): 70% of the transcription task is complete; Android and iOS applications are in beta. The semantic features enable sophisticated queries (e.g., dharma references in Uttara Kanda) that are impossible with scanned images.

7.6 Mewar Ramayana Online Reunion

A good example of an international academic partnership is the eighteenth-century Mewar Ramayana, comprising more than 450 folios and credited to Sahibdin.

7.7 2014 UK–India Project

The British Library Conducted High-Resolution scans of 300 folios at 4000 DPI using multispectral imaging to reveal latent detail.

The corpus was expanded through the imaging of 150+ folios from the City Palace collection by CSMVS Mumbai.

The materials are synthesised by an IIIF-compliant viewer, enabling the virtual assembly of the manuscript sequence.

UV examination revealed hidden underdrawings in the miniatures, providing new insights into artistic practices.

Impact: Scholars worldwide can now access a continuous story whose pigments are more than 500 years old, with colour calibration performed.

7.8 Southeast Asian Projects: DREAMSEA Project

Launched in 2017, Digital Restoration and Conservation of Ancient Manuscripts in

[17] timesofindia.indiatimes.

[18] namami.gov.

Southeast Asia (DREAMSEA) has now digitised over 119000 pages across 57 separate collections. In 2017, Digital Restoration and Conservation of Ancient Manuscripts in Southeast Asia (DREAMSEA) also digitised over 119,000 pages distributed across 57 collections.

Country	Collection	Items	Details
Indonesia	Buton MSS	20,129 pages	Kakawin Ramayana variants
Laos	Luang Prabang	15,000+ pages	Phra Lak Phra Lam
Thailand	National Library	8,500 items	Ramakien subsets

Table 6: Ramayana Coverage

Technical Requirements: 600-DPI TIFF master files, OCR processes of Kawi/Javanese scripts, and IIIF-compatible readers.

7.9 Oral Tradition Digital Conversion Performance

Ramlila (UNESCO ICH): Indian Culture Portal has videos of Ayodhya and Ramnagar Varanasi, which document the 31-night cycle. The 10 TB audiovisual archive documents caste-inclusive performances.[19]
Bophana Centre (Cambodia): 1960s Reamker Ta Krut shadow-play recordings (a virtually completely lost heritage) have been digitised. These texts are included on the UNESCO Memory of the World list.[20]

8. New Technologies and Standards

AI/OCR Developments: Grantha OCR will achieve 82 per cent accuracy in the NMM scenario using deep learning models in 2026, up from 65 per cent in 2015. The Sanskrit parser in Project Ramayanam solves 95 per cent of the sandhi variants.
Blockchain Provenance: Pilot projects (not seen to date with Ramayana content) suggest tracing folio provenance to address colonial-era ownership claims, such as the Mewar one. Blockchain Provenance: Pilot projects (not yet realised with Ramayana material) suggest a system to track folio provenance as a solution to the ownership disputes of the colonial era, such as the Mewar one.

8.1 Interoperability Protocols

IIIF (International Image Interoperability Framework): used in Mewar and DREAMSEA.
METS (Metadata Encoding and Transmission Standard): Applied in NMM in DML to encapsulate metadata rigorously.
OAI-PMH harvesting enables cross-repository search across disparate holdings.

Initiative	Scale	Tech Maturity	Access Model	Ramayana Focus
NMM India	3.31 Cr pages	OCR 82-95%	Open DML	Highest (5-10%)
Project Ramayanam	24K shlokas	Semantic XML	API/apps	Valmiki only
Mewar Project	450 folios	4000 DPI IIIF	Public viewer	Specific codex
DREAMSEA	119K pages	hOCR/IIIF	Academic	SE Asian variants
Bophana	100+ hrs	Audio restoration	Public archive	Reamker only

Table 7: Initiatives Comparative Analysis

8.2 Difficulties and Standardisation Procedures

- Technical Hurdles: Cursive scripts (Grantha, Kawi) cannot be OCR-read at below 85 per cent; folded palm-leaf leaves cause scan distortion.
- Legal: Legal challenges cause stalling of about twenty per cent of NMM uploads; there is still controversy over Creative Commons licensing.
- Solutions: LOCKSS distributed preservation mechanisms are the complement to the five-level quality control systems used by NMM (raw - gold standard).

8.3 Unified Standards (NMM/IGNCA)

- At least 600 DPI resolution; the lossless JPEG2000 and TIFF formats are preferred.

[19] Indian culture

[20] https://english.cambodiadaily.com

- METS technical and rights metadata capture of 1.8 +.
- Image delivery made possible by IIIF Presentation API 3.0.
- Durable citation is provided using persistent identifiers (handles, ARKs).

The entire digitisation landscape described in this paper represents the path to practical preservation, and the scalable models have the potential to be replicated across global manuscript traditions.[21]

9. Challenges and Solutions

9.1 Physical Degradation and Environmental Endangerment

Main Obstacle: Palm-leaf manuscripts (comprising eighty per cent of the Ramayana corpus) deteriorate quickly in areas of high humidity (more than seventy per cent), resulting in brittle material within half a century. Under suboptimal storage conditions, termite infestation can reduce an estimated 20% per year. According to NMM, 60 per cent of the five million Indian manuscripts are in critical danger, and copies of Valmiki in Kerala are particularly at risk.

9.1.1 Regional Variations

- Southeast Asia:

Cambodian Reamker leaves are attacked by fungi at ninety per cent relative humidity; Indonesian Kakawin on sago palm is warped under such circumstances.[22] Paper codices (Mewar Ramayana): Flaking Ink flaking is an issue with thirty per cent of the folios of the 16th century.[23]

9.1.2 Solutions

Assay: NMM Conservation Protocols: fumigation: Twenty-eight degrees Celsius using aluminium phosphide; lamination: forty per cent with 4 per cent rice starch; climate: twenty-eight degrees Celsius with half its relative humidity. One lakh folios are treated under the programme each year.

Preventive Storage: The use of acid-free boxes and silica-gel desiccants will extend the life of materials by up to 200 years.

Digital Surrogates: 600-DPI masters. The fate of the original is always ensured, even when it is later lost in material form.

Degradation Factor	Impact Rate	Mitigation
Humidity/Temperature	70% MSS affected	HVAC at 1,500 centres
Insects (termites)	20% annual loss	Methyl bromide fumigation
Handling	15% damage	Gloves, supports

Table 8: Degradation Factor & Impact

9.2 Technological Constraints of Digitisation

OCR and Script Challenges: Devanagari OCR has about 95% accuracy; Grantha at 82%; and Javanese Kawi at less than 70%. The curvature associated with palm-leaf manuscripts distorts about a quarter of scans, and the weakening inks may be undetected using traditional CCD sensors.

Data Volume: The National Mission for Manuscripts (NMM) contains 3.31 crore pages, requiring an estimated 500 TB of storage. The DreamSEA project, with 119,000 pages, requires significant university server infrastructure.

9.2.1 Solutions

- Multispectral Imaging: Multispectral imaging, using ultraviolet/infrared at 4000 dpi, was used in the Mewar project, with an 85 per cent success rate.
- AI Advances: The results of modern deep-learning architectures will improve Grantha OCR performance from 65 in 2015 to 82 in 2026.
- Scalable Workflows: The phased nature of Project Ramayanam, including first imaging, OCR conversion, and eventually semantic XML generation, is scalable.

9.3 Intellectual Property and Conflicts of Access

Ownership Disputes: The ownership of Mewar folios between the UK and India is a legacy of colonial-era dispersal, which delays about 20 per cent of NMM uploads. Trusts granted to temple properties claim an unlimited right over donated manuscripts.

Cultural Sensitivities: Ethnolinguistic communities that possess Ramkatha materials might limit

[21] namami.

[22] southeastasianlibrarygroup.wordpress.

[23] timesofindia.indiatimes.

digitisation, and Ramlila practice practitioners are concerned about the commercialisation of their performances.

9.3.1 Solutions

- Metadata rights framework: NMM uses Creative Commons CC-BY-SA on all the public-domain manuscripts and CC-BY-NC on the problematic holdings.
- Community Protocols: Memoranda of Understanding with tribal councils, including the Arunachal Pradesh, determine mediated access pathways.
- Blockchain Provenance: Folio chain-of-custody tracking is initially being piloted, but has yet to be formally implemented.

9.4 Accessibility Gaps and Digital Divide

Barriers to infrastructure: Rural areas in India and Southeast Asia have access rates to digital infrastructure of less than 20%, while urban areas have access to 80% of the Digital Manuscript Library (DML).
Language Bars: 90% of metadata will remain in English/Sanskrit, requiring regional interfaces for different regions.

9.4.1 Solutions

- Mobile -First Access: Android applications based on the Ramayanam API, with Hindi/Tamil UIs, are expected to have more than 50,000 downloads.
- Offline Models: USB delivery to 500 or more rural centres has helped to make about 1000000 downloads of the manuscript.
- Multilingual Metadata: DML Phase 3. The 12 metadata languages have been added.

9.5 Interoperability and Fragmentation

Siloed Repositories: The NMM schema is incompatible with DreamSEA, and the British Library IIIF viewer does not currently federate with Indian portals.
Metadata Inconsistency: Different versions of Dublin Core and METS are used, as well as non-standard subject headings (e.g., Bala Kanda vs. Book 1), which hamper cross-reference.

9.5.1 Solutions

IIIF Universal Viewer is an open-source platform allowing a unified interface to access more than 100 repositories.

- OAI-PMH Harvesting: The NMM DML can cross-search over 50000 items.
- Schema Alignment: A set of mandatory METS partners promulgates 1.8+ standards.
-

Challenge Category	Severity (1-10)	Solution Readiness	Example
Physical Decay	9	High (NMM protocols)	Palm-leaf lamination
OCR Accuracy	7	Medium (AI progress)	Grantha 82 %
IP Disputes	8	Low (legal pending)	CC-BY-SA framework
Digital Divide	6	High (mobile apps)	Rural USB kits
Interoperability	7	Medium (IIIF)	Cross-repo search

Table 9: Challenges Categories

9.6. Financing and Sustainability Problems

Chronic Underfunding: The NMM's annual budget of about 50 crores is significantly below the estimated 500 crores; DreamSEA is highly grant-based.
Staffing Issues: The current staffing is 1 paleographer per 10,000 manuscripts, and there is a training backlog of 14 scripts.

9.6.1 Solutions

- Public-Private Partnerships: Including Google/NIC to do OCR and Microsoft Azure donating 50 TB of storage.

- Crowdsourcing: The GitHub repository of Project Ramayanam has more than 500 volunteers who do the shloka proofreading.

10. Proposed Digital Framework

This study proposes the development of an integrated digital platform (hereafter referred to as a "super site") dedicated to the Ramayana and its major textual variants. Unlike existing digitisation initiatives that primarily focus on preservation and archival access, the proposed system aims to combine accessibility, comparative analysis, and linguistic resources within a single unified interface.

The platform will include major Ramayana texts such as the Valmiki Ramayana and Ramcharitmanas, along with other significant regional and international variants. A key feature of this system will be the integration of a structured lexical resource (dictionary) derived from these texts, enabling word-level exploration and semantic understanding across versions.

Regarding content accessibility, copyright-free texts will be hosted directly on the platform. At the same time, restricted materials will be linked from external repositories, ensuring both legal compliance and comprehensive coverage. This hybrid access model allows the inclusion of a wide range of sources without compromising intellectual property norms.

A central innovation of the proposed work is a comparative analysis system that enables users to examine how narrative elements, themes, and characters vary across versions of the Ramayana. This feature will support thematic mapping (e.g., dharma, ethics, regional adaptations) and enable cross-variant study in a structured manner.

Furthermore, the platform aims to prioritise multilingual accessibility, particularly for Indian regional languages. By incorporating multiple language interfaces and translations, the system seeks to bridge linguistic barriers and expand access to diverse user groups, including scholars, students, and general readers.

Overall, this proposed framework moves beyond static digitisation by transforming the Ramayana corpus into an interactive, comparative, and linguistically enriched digital knowledge system. It addresses key gaps identified in existing research, including resource fragmentation, the lack of comparative tools, and limited regional language accessibility.

Key Features of the Proposed System

The proposed digital platform is designed as an integrated, multifunctional system that addresses the major limitations of existing Ramayana digitisation initiatives. Its core features are outlined as follows:

Integrated Ramayana Text Repository:
The platform will serve as a centralised repository incorporating major Ramayana texts, including the Valmiki Ramayana, Ramcharitmanas, and other significant regional and international variants. This integration will reduce fragmentation by bringing dispersed textual resources into a single accessible environment.

Dictionary-Based Lexical Exploration:
A key component of the system will be a structured dictionary derived from the included texts. This feature will enable users to explore word meanings, contextual usage, and semantic variations across different versions, thereby supporting linguistic and philological research.

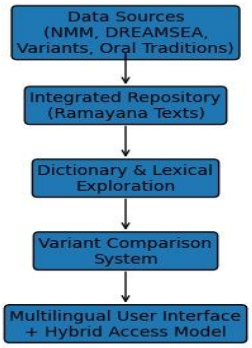

Figure 1: Proposed Digital Framework for Integrated Ramayana Platform

Variant Comparison Tool:
The platform will include a comparative analysis mechanism that allows users to examine differences in narrative structure, themes, and character representation across multiple Ramayana traditions. This tool will facilitate systematic cross-variant study and enhance understanding of regional adaptations.

Hybrid Copyright Access Model:
To ensure both comprehensiveness and legal compliance, the system will adopt a hybrid access approach. Copyright-free materials will be hosted directly on the platform. At the same time, restricted texts will be made accessible via curated external links, thereby maintaining continuity of access without violating intellectual property norms.

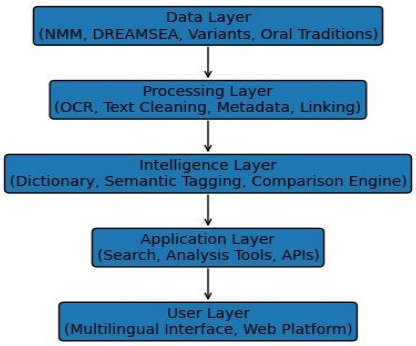

Figure 2: Layered Architecture of the Proposed Ramayana Digital Platform

Multilingual Interface with Focus on Indian Languages:
The platform will prioritise multilingual accessibility, particularly emphasising Indian regional languages. By providing interfaces and content in multiple languages, the system aims to make the Ramayana corpus accessible to a broader audience, including non-English-speaking users, thereby promoting inclusive digital scholarship.

11. Conclusion

The preservation and digitisation of the Ramayana and its more than 300 variants are not merely technical exercises but broader civilizational responsibilities. From palm-leaf manuscripts in India to Southeast Asian performance traditions, the epic reflects both cultural richness and material vulnerability. Institutional initiatives such as the National Mission for Manuscripts, the Mewar Ramayana digital reunification, and DREAMSEA have significantly advanced preservation through large-scale digitisation and improved accessibility. At the same time, semantic projects like Project Ramayanam mark a transition toward machine-readable and analytically usable textual corpora.
Despite these advancements, major challenges persist, including physical degradation, limitations in OCR for scripts such as Grantha and Kawi, intellectual property issues, and interoperability across repositories. These factors continue to fragment the digital landscape and limit the full potential of comparative research.
In response, this study proposes the development of an integrated digital platform for the Ramayana that combines major texts such as the Valmiki Ramayana and Ramcharitmanas with other regional and global variants. By incorporating a dictionary-based lexical system, a cross-variant comparison tool, and a hybrid copyright access model, the platform seeks to address fragmentation and enhance accessibility. Its emphasis on multilingual interfaces, particularly in Indian regional languages, further supports inclusive and wider engagement.
Overall, sustainable preservation requires integrated digital infrastructure, AI-driven tools, ethical governance, and inclusive access strategies. Digitisation does not replace manuscripts but extends their life, transforming the Ramayana into a dynamic and accessible knowledge system. Through such efforts, the tradition can continue to function as a living archive of cultural memory and dialogue.
The results indicate that there are 5 strategic imperatives of sustainable preservation:

- Integrated Digital Infrastructure: a federated Asia Pacific Ramayana Digital Hub connecting Indian and Southeast Asian and global repositories with interoperable protocols.
- Aristocratic AI and Multispectral Imaging- scaling the script recognition abilities and retrieving the lost texts, thus reducing the reliance on manual palaeography.
- Community-Based Access Models- the value of tribal, temple and performance traditions is respected but regulated digital dissemination.
- Provenance and Ethical Governance - The use of blockchain-based tracking and transparent rights to solve colonial dispersals and challenged ownership.
- Inclusive Accessibility- mobile-first platforms, multi-language metadata and offline distribution strategies to overcome rural and regional disparity.

Finally, digitisation is not a replacement for the manuscript; it prolongs its existence. Throughout history, the Ramayana has been active: transmitted orally, reconstructed regionally, acted out in rituals, and artistically recreated. Digital preservation advances this evolutionary chain by converting fragile artefacts into a global resource that is accessible, interoperable, and analyzable. This guarantees that Ramayana is not a heritage of the past but a living archive of dialogue. In this regard, systematic digitisation not only preserves content but also maintains memory, identity, and intercultural continuity. Digital resilience of the Ramayana tradition will be assured through collaboration in stewardship, technological innovation, and ethical responsibility, ensuring that the tradition remains relevant and accessible to future generations in an ever-increasingly interconnected world. Academia.edu. *Tribal Ramkatha traditi ons in India*.

12. References

Academia.edu. *Tribal Ramkatha traditions in India*.

Amar Chitra Katha. (n.d.). *Historical spread of the Ramayana in Southeast Asia*.
Arunachal Times. (2025). *Discovery of tribal Ramayana manuscripts in Arunachal Pradesh*.
Bhandarkar Oriental Research Institute. *Catalogue of Ramayana manuscripts*. Pune, India.
Bharati Ideology. *Kakawin Ramayana and Indonesian manuscript traditions*.
Bophana Center. *Digitisation of Reamker recordings and audiovisual archives*. Phnom Penh, Cambodia.
British Library. (2014). *The Mewar Ramayana digital reunion project*. London, UK.
Deccan Herald. *Oral and tribal Ramayana traditions in India*.
Digital Restoration and Conservation of Ancient Manuscripts in Southeast Asia (DREAMSEA). *Project reports and digitisation statistics*.
GitHub. (2023–present). *Project Ramayanam: Digital Sanskrit edition of Valmiki Ramayana*.
GKToday. *Variants of the Ramayana across Asia*.
Indian Culture Portal. *Ramlila and other Ramayana performance traditions*. Ministry of Culture, Government of India.
Indira Gandhi National Centre for the Arts (IGNCA). *Digital manuscript preservation initiatives*.
International Society for Cultural Analysis (ISCA). (2023). *Ramayana digitisation and cultural resilience in the AI era*.
National Mission for Manuscripts (NMM). (2003–2026). *Digitisation reports and Digital Manuscripts Library statistics*— Ministry of Culture, Government of India.
Namami.gov.in. *National Mission for Manuscripts official data portal*.
Press Information Bureau (PIB). (2025). *India–Southeast Asia cultural heritage cooperation reports*. Government of India.
Rajasthan Oriental Research Institute. *Manuscript holdings and preservation reports*. Jodhpur, India.
A.K. Ramanujan, 300 Ramayanas
ScholarKart. *Digital manuscript hosting and regional Ramayana subsets*.
Siam Society. (2024). *Digital cataloguing of Ramakien murals at Wat Phra Kaew*. Thailand.
Slideshare. *Global translations and adaptations of the Ramayana*.
Southeast Asian Library Group. *DREAMSEA digitisation updates and manuscript collections*.
Spectrum Books. *Cultural footprint of the Ramayana in Asia*.
The Times of India. (2014). *Mewar Ramayana folios digitally reunited*.
The Times of India. *National Mission for Manuscripts digitisation statistics*.
UNESCO. *Memory of the World Asia-Pacific Register: Reamker recordings*.
Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage Publications.