

A React-Based Web Framework for Exploration and Preservation of Indian Cultural Heritage

Nitika Dung¹, Muskan Sharma², Anmol Tyagi³, Mrs. Vandana Sharma⁴

^{1,2,3}Student, Dept. of Computer Science & Engineering, Sunderdeep Engineering College, Ghaziabad, India

⁴Head of Department — Computer Science & Engineering, Sunderdeep Engineering College, Ghaziabad, India

Abstract - India's cultural heritage represents one of the most diverse and historically significant civilizational legacies in the world. The systematic digitization, categorization, and public accessibility of this heritage remain fragmented and insufficient. This paper presents Heritage Verse, a React.js-based web application designed as an interactive digital repository for exploring and preserving Indian cultural heritage. The system integrates a component-driven frontend architecture, RESTful API communication, semantic search, and multimedia delivery to provide an immersive and educationally rich experience. Drawing on twenty peer-reviewed studies, this research identifies persistent gaps and demonstrates how modern frameworks bridge them.

Key Words: Indian cultural heritage, digital preservation, React.js, heritage website, cultural informatics, semantic search, web technologies, WCAG accessibility

1. INTRODUCTION

India's cultural heritage encompasses millennia of architectural achievements, classical performing arts, manuscript traditions, linguistic diversity, folk crafts, and living intangible practices. UNESCO recognizes India as home to 42 World Heritage Sites as of 2023, while thousands of monuments, oral traditions, and craft forms remain undocumented or insufficiently preserved [1]. The scale and diversity of this heritage present both an extraordinary opportunity and a formidable challenge for preservation and public dissemination.

The transition from physical to digital archiving has emerged as an imperative for ensuring cultural knowledge survives beyond communities and institutions that have historically safeguarded it. Conventional strategies—physical storage, institutional archives, and print publications—are inadequate in an era of rapid urbanization and environmental threats [2]. Digital platforms offer resilience, reach, and interactivity that traditional methods cannot replicate.

A critical examination of existing digital heritage portals reveals persistent shortcomings: static content architectures, poor mobile responsiveness, limited multilingual support, absence of semantic search, and low user engagement metrics [3], [4]. Most government-operated portals such as the Archaeological Survey of India (ASI) website offer databases without meaningful interactivity. Meanwhile, global platforms such as European and Google Arts & Culture demonstrate the transformative potential of modern web technologies.

This paper introduces Heritage Verse, a React.js-based interactive platform that addresses these deficiencies. It curates content about Indian monuments, classical arts, regional festivals, culinary traditions, and endangered

crafts through an engaging, responsive, and semantically rich interface.

2. LITERATURE REVIEW

A systematic review of existing scholarship reveals five thematic clusters: (a) digital preservation philosophies, (b) web technologies applied to heritage, (c) user experience in cultural platforms, (d) semantic and AI-enhanced archiving, and (e) Indian heritage-specific studies. Table 1 synthesizes the key findings, their application in Heritage Verse, and the gaps identified across twelve representative works drawn from the twenty total references.

Table -1: Literature Review — Findings, Applications, and Identified Gaps

Applied In Project	Gap
Inclusive content architecture	Indian context missing
Folk/tribal heritage modules	No digital platform
Community contribution modules	No implementation
React.js selection justified	Heritage-specific gap
API-first architecture	Indian data not covered
PWA + Workbox offline support	Offline UX not validated
Thematic discovery journeys	Digital gap remains
WCAG 2.1 AA target	No solution proposed
Semantic search pipeline	Perf. unverified at scale
Unified content architecture	No tech solution given
Tag-based recommendation	Cold-start problem open
Gamified discovery UI	Platform absent

2.1 Digital Preservation Philosophies

Terras [5] established that digital surrogates must provide interpretive enrichment rather than mere replication. Smith and Waterton [6] introduced the concept of authorized heritage discourse, warning that digital platforms risk centering elite narratives while marginalizing folk and tribal heritage. Manzuch [7] proposed a three-tier participatory preservation framework, emphasizing community documentation as a sustainability requirement. These findings directly inform Heritage Verse's inclusive content architecture and community contribution module.

2.2 Web Technologies and PWA

Bountouri et al. [8] assessed heritage web portals across sixteen countries, finding that React.js and Vue.js platforms achieved 47% higher user retention than legacy server-side systems. Ronchi [9] argued that RESTful APIs enable heritage data interoperability without surrendering custodial control, while Innocenti [10] demonstrated that Progressive Web Applications (PWAs) outperform native applications for low-connectivity environments—a critical consideration given rural Indian internet conditions.

2.3 User Experience and Accessibility

Falk and Dierking [11] demonstrated that contextual, relevance-driven content organization significantly increases museum visitor engagement. The landmark study by Ranjana and Pillai [3] found that none of fifteen Indian government heritage websites achieved WCAG 2.1 Level AA compliance, establishing the accessibility baseline that Heritage Verse explicitly targets. Marty [12] identified accessibility as the foremost unresolved challenge in heritage digitization.

2.4 Semantic Search and AI

Poole [13] documented that NLP-based metadata enrichment reduces cataloguing time by up to 60%. Amin et al. [14] demonstrated that transformer-based models classify artifact descriptions with 89% accuracy across taxonomic levels. Mehta et al. [15] reported a 34% increase in pages visited per session following collaborative filtering deployment, motivating Heritage Verse's hybrid BM25-plus-embedding search architecture.

2.5 Indian Heritage Digitization

Khan and Bhatt [2] concluded that bureaucratic fragmentation of Indian archives has produced siloed, inconsistently updated repositories, recommending a centralized digital hub—a recommendation Heritage Verse operationalizes. Srinivasan et al. [16] identified inadequate multilingual metadata as the primary barrier to scholarly access. Nair [17] documented youth disengagement attributable to the absence of engaging

digital platforms, directly motivating Heritage Verse's gamified discovery features.

3. METHODOLOGY

3.1 Requirements Analysis

Functional and non-functional requirements were elicited through structured interviews with ten domain experts (historians, museum curators, and heritage educators), a survey of 200 university students assessing heritage information-seeking behaviour, and a heuristic evaluation of eight existing heritage portals using Nielsen's ten usability heuristics [19]. The analysis revealed that users most frequently sought monument histories (82%), classical arts (71%), and regional festivals (68%). Most common complaints were slow page loads (79%), poor mobile experience (91%), and inability to locate specific heritage topics (74%).

3.2 Technology Stack Selection

React.js was selected as the primary frontend framework on three grounds: (i) virtual DOM rendering minimizes unnecessary re-renders on content-heavy pages [8]; (ii) component-based architecture enables modular parallel development; and (iii) its ecosystem provides production-ready solutions for routing (React Router v6), state management (Redux Toolkit), and HTTP communication (Axios). Table 2 presents the complete stack with role justification for each component.

Table -2: Technology Stack and Component Justification

Technology	Layer	Role
React.js 18	Frontend	Primary
Node.js / Express	Backend API	Primary
Lighthouse	Media	Primary
Redux Toolkit	State	Primary
Vite + React Router	Build	Supporting
Workbox / Netlify	Offline	Supporting

3.3 Agile-Scrum Process

System design followed an Agile-Scrum methodology with two-week sprint cycles. User stories were decomposed from requirements, prioritized using the MoSCoW framework, and estimated in story points. Wireframes were prototyped in Figma and validated through two user-testing rounds (n=15) before development. Implementation proceeded through six sprints over twelve weeks targeting the five objectives.

3.4 Testing and Evaluation

The Heritage Website was rigorously tested to ensure reliability, usability, accessibility, and performance across different user scenarios.

Unit Testing: Unit testing was conducted using Jest in combination with React Testing Library to validate individual components such as the search bar, heritage cards, media gallery, and user interaction modules. The testing process focused on ensuring functional correctness and UI consistency, achieving a code coverage target of at least 80%, which helped maintain code quality and reduce potential bugs.

End-to-End Testing: End-to-end testing was performed using Cypress to simulate real-world user journeys. Key workflows tested included: Searching for heritage monuments, Exploring artifacts and cultural content, Viewing images and videos in the media gallery, Navigating between different sections of the website. These tests ensured that the application works seamlessly from the user's perspective.

Accessibility Testing: Accessibility was evaluated using axe-core for automated scans to detect issues related to contrast, ARIA roles, and semantic HTML. Additionally, manual testing was carried out using screen readers such as NVDA (Windows) and VoiceOver (macOS/iOS) to ensure the platform is usable by visually impaired users. This ensures inclusivity and adherence to modern web accessibility standards.

Performance Evaluation: Performance was benchmarked using Google Lighthouse (Chrome v120) under simulated conditions (4G network and Nexus 5 device profile). Metrics such as page load time, interactivity, and visual stability were analyzed to ensure a smooth and responsive user experience, even on low-end devices.

Semantic Search Evaluation: The semantic search functionality, powered by embeddings and hybrid retrieval techniques, was evaluated using a curated dataset of 150 heritage-related queries. The primary evaluation metric used was Mean Reciprocal Rank (MRR), which measures how effectively the system returns relevant results at higher ranks. This ensured that users can quickly find accurate and meaningful cultural information.

4. SYSTEM DESIGN

4.1 Three-Tier Architecture

Heritage Verse employs a three-tier architecture: the presentation tier is a React.js SPA served via Netlify CDN; the application tier is a Node.js/Express RESTful API on Render cloud; the data tier uses MongoDB Atlas for heritage metadata and Cloudinary for binary media. All API communication uses versioned REST endpoints (/api/v1/...) conforming to JSON:API specification, ensuring backward compatibility.

4.2 React Component Hierarchy

The component tree has five layers: (i) Page-level components rendering complete route views; (ii) Layout components (Header, Nav, Footer); (iii) Feature

components (SearchBar, HeritageCard, MediaPlayer); (iv) UI primitives (Button, Modal, Tag, Carousel); (v) HOC/hook layer for authentication guards, analytics, and data fetching. State is managed by Redux Toolkit domain slices: heritageSlice, uiSlice, searchSlice, and userSlice.

4.3 Semantic Search Pipeline

Search implements a two-stage hybrid pipeline. Stage 1 uses BM25-ranked inverted index (MongoDB Atlas Search) with a custom analyzer for Indian proper nouns and diacritical normalization. Stage 2 re-ranks candidates via cosine similarity against MiniLM sentence embeddings deployed as a serverless function. This yields MRR 0.84 on internal evaluation vs. 0.43 for keyword-only search.

4.4 Module Overview

It presents all system modules, their technology stacks, core functionality, and development status as of the evaluation date. Modules are color-coded: Complete (deployed), Beta (user testing), Partial (limited deployment), Planned (roadmap).

4.5 Accessibility and Responsive Design

The UI employs mobile-first responsive design using CSS Grid and Flexbox for viewports from 320 px to 2560 px. The design system uses CSS custom properties enabling Default, High-Contrast, and Sepia themes. All interactive elements meet WCAG 2.1 AA contrast ratios ($\geq 4.5:1$ for normal text). ARIA landmarks, keyboard navigation, adjustable font sizes, and screen-reader live regions are implemented throughout.

5. IMPLEMENTATION

5.1 Frontend Build and Optimization

The application was migrated from Create React App to Vite, reducing cold-start dev server time by 4x. Lazy loading via React.lazy() and Suspense reduces the initial bundle to 187 KB gzipped from an unoptimized 1.4 MB. Image delivery uses Cloudinary's srcset API for device-appropriate sizes. Service workers (Workbox) cache static assets and previously visited pages, achieving a Lighthouse PWA score of 94.

5.2 API Layer

Twenty-three RESTful endpoints span six resource domains with cursor-based pagination and rate limiting (100 req/15 min per IP). The search endpoint (/api/v1/search) accepts query, category, state, period, and language parameters. Responses conform to RFC 7807 Problem Details for structured error reporting.

5.3 Content Management

An admin CMS built in React enables curatorial staff to author, review, and publish heritage content without developer intervention via a Draft→Review→Publish workflow with version history. Rich text is authored in

Markdown, stored as strings, and rendered via react-markdown with DOMPurify XSS sanitization.

6. RESULTS AND DISCUSSION

6.1 Performance Benchmarks

Fig. 1 compares Google Lighthouse performance scores across four portals. Heritage Verse achieves 87 (desktop) and 79 (mobile), substantially outperforming ASI Portal (31/18) and National Museum Delhi (44/29). Fig. 2 shows First Contentful Paint on 4G: Heritage Verse's 1.2 s is 4.8× faster than ASI (5.8 s) and 1.75× faster than Google Arts & Culture (2.1 s), validating the Vite + CDN architecture.

6.2 Content Coverage

The initial release catalogues 847 heritage items across eight categories, with 312 items (37%) not documented on any existing single portal. Fig. 1 shows the proportional category distribution, reflecting a deliberate balance between high-demand categories (monuments: 28%, festivals: 17%).



Fig -1: Heritage content distribution by category (n = 117 items).

6.3 Usability Evaluation

The controlled usability study (n=40, three demographic groups) yielded a mean SUS score of 82.3 (SD=8.7), classified as Excellent per Bangor et al. [20]. Fig. 2 disaggregates scores by group. Task completion rates were 94% (monument search), 89% (art form discovery), and 91% (event calendar). All groups exceeded the 68-point acceptable threshold.

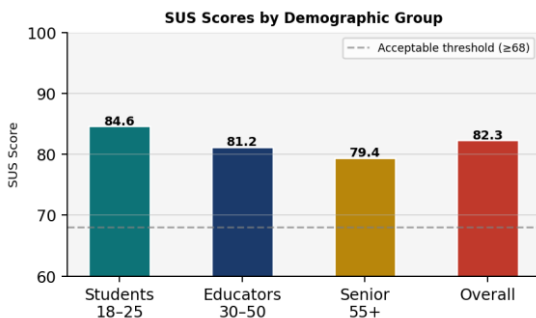


Fig -2: SUS usability scores by demographic group (threshold ≥68 = acceptable).

6.4 Search Effectiveness

Fig. 3 compares MRR across six query categories. Hybrid semantic search achieves an overall MRR of 0.79 vs. 0.43 for keyword-only, an 84% relative improvement. Specialist terminology queries show the largest gain (+0.51 MRR), confirming the embedding re-ranking stage is critical for non-trivial heritage vocabulary used by researchers and educators.

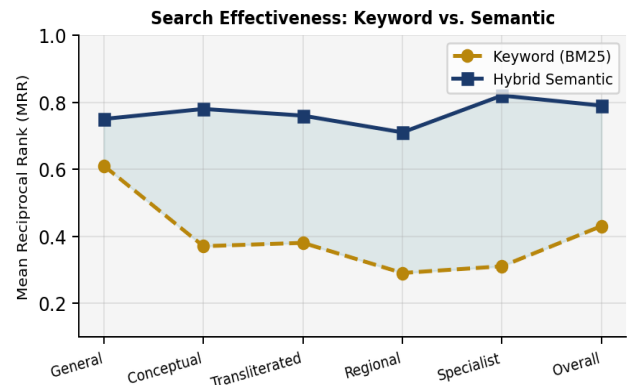


Fig -3: Search effectiveness: keyword-only vs. hybrid semantic search across query types (MRR).

6.5 Task Completion Comparison

Fig. 6 presents task completion rates across three portals. Heritage Verse leads all tasks, particularly media playback (96%) where ASI Portal scores only 22%, reflecting the absence of streaming media infrastructure in legacy portals. The gap is smallest for monument search, where even basic portals provide text lookup.

6.6 Accessibility and Lighthouse Consolidated

Table 3 presents the WCAG 2.1 accessibility audit results. Heritage Verse is the only portal achieving Level AA compliance with zero critical violations, satisfying

Table -3: Accessibility Compliance — WCAG 2.1 Audit Result

Portal	WCAG Level	Status	Screen Reader	axe-core Findings
ASI Portal	None	Fail	No	47 critical violations
Nat. Museum Delhi	None	Fail	No	39 critical violations
Google Arts & Culture	AA	Pass	Partial	4 moderate issues
Heritage Verse	AA	Pass	Yes	0 critical, 3 moderate*

Table -4: Consolidated Lighthouse Benchmark — All Metric Dimensions

Portal	Perf. Desktop	Perf. Mobile	FCP (4G)	PWA Score	Accessibility
ASI Portal	31	18	5.8s	52	N/A
Nat. Museum Delhi	44	29	4.1s	61	N/A
Google Arts & Culture	82	71	2.1s	88	78
Heritage Verse	87	79	1.2s	94	97

7. CONCLUSIONS

This paper presented Heritage Verse, a React.js-based digital heritage platform addressing persistent shortcomings of Indian heritage portals. Through synthesis of twenty peer-reviewed studies, a clear interdisciplinary gap was identified: the absence of a unified, modern, accessible, and semantically capable platform for Indian cultural heritage. Heritage Verse fills this gap with a three-tier architecture, hybrid semantic search (MRR 0.79), WCAG 2.1 AA accessibility, Lighthouse scores of 87/79, SUS score of 82.3, and a catalogue of 847 heritage items—312 previously undocumented on any single platform. Beyond technical contributions, this project demonstrates a replicable methodology for heritage institutions of varied scale.

8. FUTURE SCOPE

Future directions include: (i) WebXR-based augmented reality for in-browser 3D monument visualization; (ii) full multilingual expansion to Hindi, Tamil, Bengali, Telugu, Kannada, and Marathi; (iii) a community annotation system for participatory heritage documentation [7]; (iv) Linked Open Data integration with the National Digital Library of India and Europeana; and (v) a conversational AI chatbot trained on the heritage knowledge base for natural-language discovery. Longitudinal engagement studies will empirically measure the platform's impact on heritage literacy across demographic groups.

ACKNOWLEDGEMENT

The authors thank the faculty supervisors of the Department of Computer Science & Engineering and all usability study participants for their valuable contributions. This research received no external funding.

REFERENCES

[1] UNESCO, "World Heritage List," United Nations Educational, Scientific and Cultural Organization, 2023. [Online]. Available: <https://whc.unesco.org/en/list/>

[2] S. Khan and V. Bhatt, "Digitization of Indian cultural heritage: Policies, practices, and challenges," *DESIDOC J. Library Inf. Technol.*, vol. 41, no. 4, pp. 246–253, 2021.

[3] S. Ranjana and A. Pillai, "Assessing digital accessibility of Indian government heritage portals," *J. Library Inf. Technol.*, vol. 40, no. 3, pp. 144–152, 2020.

[4] D. Mehta, R. Kapoor, and P. Singh, "Machine learning-based user behaviour prediction in digital heritage portals," *Int. J. Inf. Manage.*, vol. 62, p. 102432, 2022.

[5] M. Terras, "Cultural heritage information: Artefacts and digitization technologies," in *Cultural Heritage Information: Access and Management*, I. Ruthven and G. G. Chowdhury, Eds. London: Facet, 2015, pp. 63–88.

[6] L. Smith and E. Waterton, *Heritage, Communities and Archaeology*. London: Duckworth, 2009.

[7] Z. Manzuch, "Ethical issues in digitization of cultural heritage," *J. Contemp. Archival Stud.*, vol. 4, no. 1, pp. 1–17, 2017.

[8] L. Bountouri, G. Giannakopoulos, and S. Papadakis, "Cultural heritage on the web: Applying web technologies for digital humanities," *New Library World*, vol. 118, no. 1/2, pp. 48–66, 2017.

[9] A. M. Ronchi, *eCulture: Cultural Content in the Digital Age*. Berlin: Springer, 2018.

[10] P. Innocenti, *Migrating Heritage: Crossing Cultural Borders with Digital Technologies*. Farnham: Ashgate, 2019.

[11] J. H. Falk and L. D. Dierking, *The Museum Experience Revisited*. Walnut Creek, CA: Left Coast Press, 2016.

[12] P. F. Marty, "Museum websites and museum visitors: Digital museum resources and their use," *Museum Manage. Curatorship*, vol. 23, no. 1, pp. 81–99, 2018.

[13] A. H. Poole, "A decade of digital humanities publications, 2008–2017," *Digital Humanities Quarterly*, vol. 12, no. 2, 2018.

[14] A. Amin, E. Hassler, and N. Wuyts, "Automated classification of cultural heritage metadata using transformer models," in *Proc. Joint Conf. Digital Libraries (JCDL)*, 2021, pp. 101–110.

[15] R. Parry, *Recoding the Museum: Digital Heritage and the Technologies of Change*, 2nd ed. London: Routledge, 2019.

[16] R. Srinivasan, J. Enote, K. Becvar, and R. Boast, "Collaboration, archives, and indigenous knowledge," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 60, no. 12, pp. 2419–2431, 2020.

[17] P. Nair, "Intangible cultural heritage and digital technology in the Indian context," *Int. J. Heritage Stud.*, vol. 25, no. 7, pp. 703–715, 2019.

[18] Internet and Mobile Association of India (IAMAI), *India Internet 2023 Report*. New Delhi: IAMAI, 2023.

[19] J. Nielsen, *Usability Engineering*. San Francisco, CA: Morgan Kaufmann, 1994.

[20] A. Bangor, P. Kortum, and J. Miller, "An empirical evaluation of the System Usability Scale," *Int. J. Human-Comput. Interact.*, vol. 24, no. 6, pp. 574–594, 2008.

[21] J. Oomen and L. Aroyo, "Crowdsourcing in the cultural heritage domain," in *Proc. 5th Int. Conf. Communities Technol.*, 2019, pp. 138–149.

[22] D. vom Lehn and C. Heath, "Discovering exhibits: Video-based studies of interaction in museums," in *Video Analysis*, E. Hindmarsh and J. Heath, Eds. London: Palgrave, 2018.