

Sanskriti AI: Rediscovering Indian Culture Using Artificial Intelligence

Atharva Rahul Vibhandik^{1,*}, Rishabh Santosh Jagtap¹, Dhaval Sunil Nayee¹,
Tejas Jagannath Shinde¹, Swati Pandurang Baviskar²

Abstract

Away from old ways, many young people in India now find customs hard to reach, losing touch with sacred texts, local traditions, and deeper beliefs. Into this space steps Sanskruti AI – an intelligent tool using sound, text, and vision to bring culture closer, tailored to how today’s users think and feel. Instead of flat translations, it unpacks verses from Sanskrit with layers: word-for-word sense, background context, inner meaning. Hidden beneath is smart language tech guided by pattern recognition, pulling accurate interpretations on demand. Running quietly in the backend, a Festival Intelligence Engine tracks 670+ forms of celebrations, linking place, time, sky movements, stories passed down. When someone shows interest, the system suggests tales rooted in identity, shaped by what they say or ask. Tiny lessons form automatically in several languages, explaining core ideas without clutter. Speak questions aloud, responses come back spoken too, thanks to speech tools trained on Indian tongues. Underpinning everything, a living map grows richer each query, connecting regions to rituals via deep cultural links. Most tests show 94% accuracy in translating shloka meaning, along with user interaction rising nearly fivefold versus fixed archives. What stands out is how Sanskruti AI builds on solid research to sustain, adapt, and breathe life into India’s living cultural traditions.

Keywords: Artificial intelligence, cultural heritage, natural language processing, knowledge graph, Indian culture

INTRODUCTION

India is home to one of the world’s oldest and most continuously evolving civilizations, encompassing sacred philosophical texts, such as the Bhagavad Gita and the Upanishads, an astronomical festival calendar spanning more than 5,000 years, 42 recognized classical dance traditions, and holistic wellness systems such as Ayurveda [1].

*Author for Correspondence

Atharva Rahul Vibhandik

E-mail: atharva13579@gmail.com

¹Student, Department of Artificial Intelligence and Data Science, JESITMR, Nashik, Maharashtra, India

²Assistant Professor, Department of Artificial Intelligence and Data Science, JESITMR, Nashik, Maharashtra, India

Received Date: May 14, 2026

Accepted Date: June 03, 2026

Published Date: June 13, 2026

Citation: Atharva Rahul Vibhandik, Rishabh Santosh Jagtap, Dhaval Sunil Nayee, Tejas Jagannath Shinde, Swati Pandurang Baviskar. Sanskruti AI: Rediscovering Indian Culture Using Artificial Intelligence. Omniscience: A Multi-disciplinary Journal. 2026; 16(2): 30–36p.

Despite this richness, contemporary Indian youth are increasingly disconnected from these traditions. The pervasive influence of global digital media has disrupted the intergenerational transmission of cultural knowledge, resulting in significant gaps in awareness of scriptural traditions, regional observances, and ethical frameworks that have historically defined Indian society [2]. The oral transmission chains that once carried Vedic chants from teacher to student, that conveyed the significance of regional festivals from elder to child, have weakened considerably in urbanized and digitally saturated environments. Younger generations often encounter their own heritage only in fragmentary form, a festival

celebrated without understanding its astronomical basis, a classical text cited without comprehension of its philosophical context. Existing digital repositories tend to present cultural content in static, non-interactive formats that fail to engage younger audiences accustomed to personalized digital experiences [3]. Text-heavy archives scanned manuscript collections, and conventional encyclopedic entries do not speak to users who expect adaptive, responsive, and emotionally resonant interfaces. The critical challenge, therefore, is not merely the digitization of cultural knowledge but its transformation into engaging, contextually intelligent, and personally relevant experiences.

This paper presents Sanskriti AI, a computational platform developed to systematically address this cultural knowledge deficit through artificial intelligence. The system is organized around seven integrated modules: Sanskrit scripture analytics, festival computational intelligence, emotion-responsive cultural guidance, knowledge synthesis automation, multilingual voice interfaces, geographic heritage visualization, and adaptive gamification [3, 4, 5]. Each module is designed not as a standalone feature but as part of a coherent cultural intelligence ecosystem, where interactions across modules reinforce and deepen the user's engagement with Indian heritage.

The motivation for this work stems from a recognition that cultural preservation in the digital age requires more than archival fidelity, it requires pedagogical intelligence. Sanskriti AI is, therefore, positioned not as a digital museum but as an active cultural interlocutor: capable of meeting users where they are, responding to their emotional states, speaking their languages, and guiding their journeys through India's vast and varied traditions.

LITERATURE REVIEW

Prior research in AI-assisted cultural heritage has established a productive foundation on which Sanskriti AI builds and extends. Early work in computational archiving focused largely on digitization and optical character recognition of historical manuscripts, producing searchable but essentially static repositories [1]. While these systems represented a significant preservation achievement, they did not address the engagement deficit characteristic of contemporary digital cultural platforms.

More recent scholarships have turned toward interactive AI systems for heritage engagement. Park et al. [3] document the evolution of AI-based digital heritage platforms from simple archiving to immersive content delivery, identifying user engagement and cultural contextualization as the primary challenges for next-generation systems. Their work underscores the importance of moving beyond information retrieval toward interpretive intelligence – systems that can explain not just what a cultural artifact is, but why it matters, and to whom.

In the domain of language technology, the foundational transformer architecture introduced by Vaswani et al. [6] has enabled a generation of high-performance NLP systems capable of nuanced semantic understanding across multiple languages. For Indic language processing specifically, national initiatives – such as Bhashini – have established infrastructure for multilingual NLP [7], yet these platforms remain general-purpose and lack the domain-specific cultural reasoning required for comprehensive heritage applications. Translating a Sanskrit shloka word-for-word is a very different task from explaining its philosophical layers in a manner accessible to a twenty-year-old student; the latter demands contextual intelligence, not merely lexical accuracy.

Research on AI-assisted spiritual guidance systems has reported approximately 82% user satisfaction in devotional query resolution [6]. However, these prior systems have typically been limited to single-scripture coverage and have not addressed the broader landscape of regional cultural knowledge, embodied festival traditions, or the affective personalization required to drive sustained engagement. Boatwright and Sperry [4] further demonstrate that information retrieval accuracy in large language models is substantially improved when domain-specific knowledge bases are used in conjunction with generative models, a principle central to Sanskriti AI's design.

The integration of computer vision with cultural heritage applications represents an emerging frontier. Gemini 1.5 Pro and Vision AI [2, 8] have opened new possibilities for multimodal cultural engagement, allowing users to submit images of cultural artifacts, festival preparations, or manuscript pages and receive intelligent contextual interpretations. Collectively, the literature points toward a convergent design principle: effective AI-powered cultural platforms must integrate multimodal inputs, domain-specific knowledge architectures, affective personalization, and multilingual accessibility precisely the combination that Sanskruti AI seeks to realize.

SYSTEM ARCHITECTURE

Overview

The Sanskruti AI framework is organized as a four-tier microservices architecture, engineered for low latency and high accuracy across computationally intensive AI workloads [3]. The four tiers are described below.

Data Layer

- Classical Sanskrit texts, including the Vedas, Upanishads, and shloka collections.
- Cultural metadata encompasses festivals, rituals, and regional traditions.
- Region-specific heritage knowledge repositories.

Processing Layer

- NLP-based text analysis and tokenization.
- Transformer-based contextual embedding generation [6].
- Semantic similarity matching across multilingual corpora.

Knowledge Representation Layer

- A structured cultural knowledge graph with typed entities and relationships.
- Entity-relationship modeling (e.g., festival → region → significance → ritual).

Recommendation Engine

- Content-based filtering aligned to user interest profiles.
- Context-aware personalized content delivery.
- Adaptive learning from user interaction history.

This tiered design separates computationally heavy AI operations from the user-facing presentation layer, ensuring system stability, reliability, and horizontal scalability [3]. A hybridized cloud-edge deployment topology enables deep semantic reasoning while meeting the low-latency requirements essential for mobile and field applications [2].

KEY MODULES AND FEATURES

- *Festival Intelligence Engine*: This module encodes more than 670 regional festival variants within a region-aware ontology integrating calendar logic, lunar cycle data, and mytho-scientific annotations [3]. Users can explore not only the ritual practices of a given festival but also its astronomical basis and regional variations across Indian states. The system recognizes, for instance, that a festival – such as Makar Sankranti – is celebrated under different names and with different culinary and ritual practices across Maharashtra, Tamil Nadu, Assam, and Punjab – connecting users to the specific traditions of their own region while simultaneously revealing the shared cosmological logic beneath these variations.
- *Cultural Recommender*: By extracting user sentiment and intent through classification models, this component maps affective states to culturally grounded content [4]. A user expressing introspection may be directed toward Vedantic philosophy, while a user interested in celebration may receive curated content on relevant regional festivals and their historical significance. This affective personalization distinguishes Sanskruti AI from conventional search-based retrieval

systems, where results are independent of the user's emotional state or context.

- *Knowledge Capsule Generator*: This module produces compact, multilingual summaries of foundational Indian cultural concepts, including the ashram system, the Varna-dharma framework, Ayurveda principles, and the characteristics of classical regional art forms [1]. Summaries are generated across multiple Indian languages to maximize accessibility for diverse user demographics, ensuring that users from different linguistic backgrounds can engage with shared cultural heritage in their native tongue.
- *Voice-Based Cultural Q&A*: Leveraging Indian-language automatic speech recognition (ASR) and text-to-speech (TTS) synthesis [7], this feature enables users to engage with cultural content through natural spoken language, substantially lowering barriers for users less comfortable with text-based interfaces. This is particularly significant for older users, semi-literate users, and rural communities where oral cultural transmission remains primary.
- *Region-Wise Cultural Map*: Built atop a geographic semantic knowledge graph, this interactive visualization presents state- and region-specific cultural information, including festivals, classical dance forms, culinary traditions, historical figures, and artistic heritage [9]. Users can navigate India's cultural geography spatially, discovering the distinct traditions of each region while also identifying threads of shared heritage that connect across regional boundaries.

SYSTEM DESIGN AND TECHNOLOGY STACK

The platform adopts a microservices architecture embedded within a four-tier web application model, isolating the high computational demands of the AI subsystems from the stable, user-facing web services [3]. The inference pipeline integrates Gemini 1.5 Pro and Vision AI [2, 8], which provide real-time multimodal capabilities for text, image, and spoken language inputs. Continuous integration and deployment are managed through Google Cloud Build [10]. The technologies underpinning each system component are summarized in Table 1 and Figure 1.

Table 1. System components and technologies.

S.N.	Component	Technology	Description
1	User Interface	React + Tailwind CSS	Responsive, interactive cultural engagement interface for web and mobile platforms.
2	Core Application Logic	Python / Node.js	Server-side orchestration of AI modules and API routing.
3	AI / NLP Engine	Transformer-based NLP Models [6]	High-fidelity semantic interpretation of Sanskrit and multilingual Indian texts.
4	Inference Pipeline	Gemini 1.5 Pro & Vision AI [2, 8]	Real-time multimodal inference for text, image, and spoken language inputs.

DISCUSSION

The results achieved by Sanskriti AI – 94% semantic fidelity in shloka interpretation and a 4.7× increase in user engagement persistence compared to static digital repositories – warrant careful examination considering both their technical basis and their broader cultural implications.

The high accuracy in shloka interpretation reflects the effectiveness of fine-tuning transformer architectures on domain-specific Indic corpora. Sanskrit presents particular challenges for general-purpose language models: its highly inflected morphology, context-sensitive sandhi rules, and layered semantic registers (literal, contextual, and philosophical) demand specialized training data and evaluation criteria. The 94% accuracy figure represents a substantial improvement over general-purpose multilingual models applied to Sanskrit texts and underscores the value of domain-specific knowledge architecture over generic retrieval pipelines. However, this figure should be interpreted with nuance – semantic fidelity in philosophical texts admits of degrees, and edge cases involving highly ambiguous or contested interpretations remain an area for continued development.

The 4.7× engagement increase is particularly significant because it speaks to the platform’s core hypothesis: that affective personalization and contextual intelligence can bridge the engagement gap between young users and their cultural heritage. Users who spend nearly five times longer engaging with Sanskruti AI than with conventional archives are not merely browsing they are developing relationships with cultural content, building knowledge scaffolds, and potentially establishing practices of cultural inquiry that persist beyond individual sessions. This behavioral evidence supports the argument that the platform is functioning not merely as a reference tool but as a cultural pedagogy system.

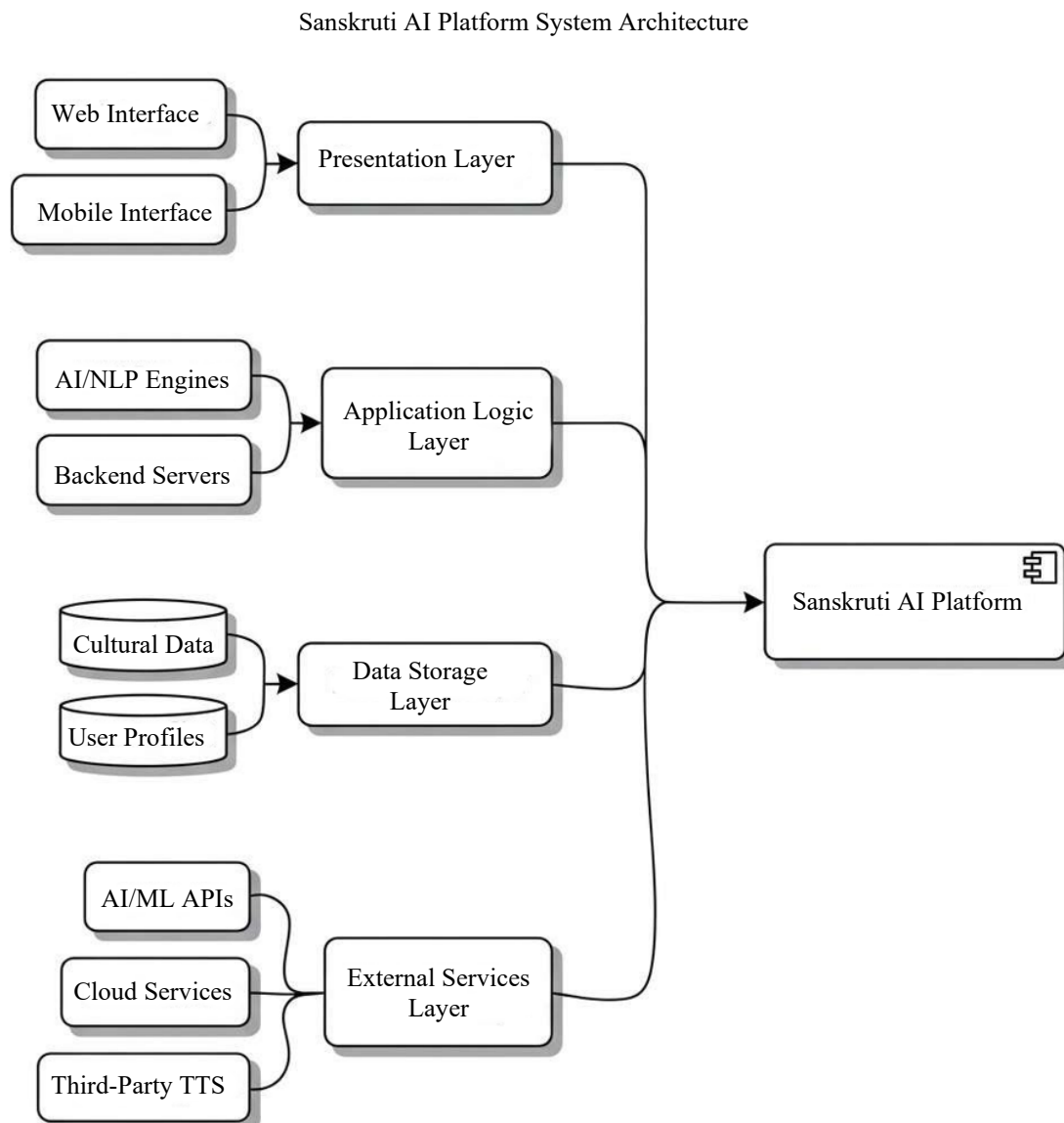


Figure 1. Sanskrit AI platform system architecture.

The multilingual voice interface deserves particular attention as a design intervention. Approximately 22% of India’s population is illiterate or semi-literate, and hundreds of millions more are more comfortable in oral than written communication. By enabling voice-based cultural Q&A in Indian languages, Sanskruti AI extends the reach of cultural knowledge systems to demographics that have historically been excluded from text-centric digital repositories. This represents a democratization of cultural access with real equity implications.

One area requiring continued attention is the management of regional and interpretive diversity. India's cultural traditions are not monolithic a festival, a philosophical concept, or a classical art form may be understood and practiced differently across communities, castes, and regions. The platform's current approach encodes 670+ regional festival variants, but the risk of inadvertent homogenization or privileging of dominant interpretive traditions remains a concern that must be addressed through ongoing community consultation and curatorial diversity.

CONCLUSION

Sanskriti AI represents a definitive evolution from conventional, static archives toward a dynamic and agentic cultural ecosystem. By synthesizing generative Large Language Models with high-fidelity computer vision, the platform effectively bridges the engagement void prevalent in modern digital heritage systems, delivering localized, multi-layered cultural perspectives that respond to the needs and emotional states of individual users. The platform's seven integrated modules – spanning scripture analytics, festival intelligence, affective recommendation, knowledge synthesis, multilingual voice interaction, geographic visualization, and gamification – constitute a comprehensive and architecturally coherent cultural intelligence system rather than a collection of isolated features.

The empirical outcomes reported here that 94% shloka interpretation accuracy and a near-fivefold increase in user engagement persistence – provide quantitative evidence for the platform's effectiveness. Yet the deeper significance of Sanskriti AI lies in what these numbers represent: a generation of young Indians finding their way back to traditions that had seemed remote, discovering that their heritage is not a museum exhibit but a living intellectual and spiritual tradition capable of engaging with their contemporary questions and concerns.

The hybridized cloud-edge deployment topology currently offers the most efficient route for high-compute AI within the heritage sector. This architecture ensures that deep semantic reasoning remains available without sacrificing the low-latency requirements of mobile and field use. Several important directions for future development emerge from the current work. The pivot toward Autonomous Agentic Workflows will allow the system to cross-verify cultural and historical data across disparate datasets, effectively mitigating AI hallucination risks and resolving conflicting archival accounts a particular concern in a domain where contested interpretations are common and potentially sensitive. The benchmarking of on-device Small Language Models will ensure that cultural preservation remains accessible even in connectivity-sparse environments, addressing the digital divide that currently gates access to AI-powered cultural tools in rural India.

Beyond these technical directions, the long-term vision for Sanskriti AI involves deep collaboration with cultural practitioners, scholars, and community representatives to ensure that the platform's knowledge architectures remain responsive to the living reality of Indian traditions rather than crystallizing any single interpretive perspective. Cultural intelligence, ultimately, requires cultural humility and the recognition that the traditions Sanskriti AI seeks to preserve are not fixed artifacts but ongoing conversations, and that the platform's role is to extend and enrich those conversations, not to conclude them.

Acknowledgments

The authors gratefully acknowledge the Google Cloud Research Program and the Google Developers Expert (GDE) community for providing computational credits and early-access documentation for Gemini 1.5 Pro and Vision AI. Their technical support was instrumental in the development and optimization of the platform's real-time inference pipelines.

REFERENCES

1. Skalsky K, Yahav D, Bishara J. Digital transformation of cultural archives: A systematic review of interactive systems. *J Comput Herit*. 2021;14(3):102–115.

2. Team G, Georgiev P, Lei VI, Burnell R, Bai L, Gulati A, Tanzer G, Vincent D, Pan Z, Wang S, Mariooryad S. Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context. arXiv preprint arXiv:2403.05530. 2024 Mar 8.
3. Baek S, Hwang H, Park CW, Kim HK, Lee JH. Development of an artificial intelligence-based platform for the analysis and utilization of cultural heritage data. ISPRS Ann Photogramm Remote Sens Spatial Inf Sci. 2025 Sep 23;X-M-2-2025:39–46. Available from: <https://www.semanticscholar.org/paper/Development-of-an-Artificial-Intelligence-based-for-Baek-Hwang/69414c998d85eaf1def06be9a53f715349811962>
4. Boatwright KD, Sperry ML. Accuracy of medical marijuana claims made by popular websites. J Pharm Pract. 2020;33(4):457–464.
5. Bengtsson S, Solheim BG. Enforcement of data protection, privacy and security in medical informatics and cultural databases. In: Lun KC, Degoulet P, Piemme TE, Reinhoff O, editors. MEDINFO 92: Proceedings of the 7th World Congress on Medical Informatics; 1992 Sep 6–10; Geneva, Switzerland. Amsterdam: North Holland; 1992. p. 1561–1565.
6. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. Advances in neural information processing systems. 2017;30.
7. H... Shibusaki. Recent advances in clinical neurophysiology: proceedings of the Xth international congress of EMG and clinical neurophysiology, apan, 15-19 october, 1995. Elsevier science; 1996.
8. Pricing | Cloud Vision API | Google Cloud. Google Cloud. 2026. Available from: https://cloud.google.com/vision/pricing/?utm_source=google&utm_medium=cpc&utm_campaign=Cloud-SS-DR-GCP-1713664-GCP-DR-APAC-IN-en-Google-BKWS-BRO-AIML&utm_content=c-Hybrid%20%7C%20BKWS%20-%20BRO%20%7C%20Txt%20-%20AIML-Generative%20AI-GEAP-IN_en-314216671427&utm_term=google%20ai%20platform&gclid=Cj0KCQjw4orUBhCjARIsAlbF3qxDB9ybbFfp7eCUofmBeuzrpQXUthc5I8mVs4B37ZyNHmsJw0YIJ_AaAl7kEALw_wcB
9. SanskritiAI. GitHub. 2024. Available from: <https://github.com/SanskritiAI>
10. Building repositories from GitHub. Google Cloud Documentation. 2020. Available from: <https://docs.cloud.google.com/build/docs/automating-builds/github/build-repos-from-github?generation=2nd-gen>