IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14471

Heritage-Connect-An AI-Powered Multilingual Guide to Tamil Nadu's Historical Gems

Mrs. V.Keerthiga[1], S Balaji[2], R Gowtham[3]

Assistant Professor, Department of Computer Science and Engineering, Anand Institute of Higher Technology,Kazhipattur,Chennai[1]

Student,Department of Computer Science and Engineering, Anand Institute of Higher Technology,Kazhipattur,Chennai[2,3]

Abstract: HeritageConnect is an AI-powered multilingual platform designed to make Tamil Nadu's cultural and historical heritage more accessible to a global audience. Leveraging natural language processing, machine translation, and voice technologies, the system allows users to engage in real-time, interactive conversations with a virtual guide. Users can ask questions in Tamil, English, Hindi, French, or Spanish—via text or voice—and receive informative, context-aware responses about heritage sites, architecture, festivals, and local traditions. The platform integrates a Flask backend with MongoDB for managing user sessions and chat history, and uses models like DeepSeek and Hugging Face for intelligent language processing. With a mobile-responsive frontend, voice support through the Web Speech API, and robust user personalization, HeritageConnect redefines how cultural knowledge is delivered, making it immersive, inclusive, and multilingual.

Keywords: Heritage Tourism, Multilingual Chatbot, Natural Language Processing, Voice Interaction, Cultural AI

1. INTRODUCTION

India is a land of immense historical and cultural wealth, with Tamil Nadu standing out as one of its most heritage-rich states. Home to centuries-old temples, monuments, and UNESCO World Heritage sites such as Mahabalipuram and the Brihadeeswarar Temple, Tamil Nadu offers a diverse array of experiences for history enthusiasts, scholars, and tourists. However, the lack of accessible, multilingual, and interactive digital guides poses a significant barrier to effectively engaging both domestic and international visitors.

Traditional tour guides, static information boards, and printed brochures often fail to provide an immersive or personalized experience. They are limited by language, depth of information, and availability. Furthermore, tourists from non-Tamil-speaking backgrounds frequently struggle to understand the historical context or cultural significance of these sites, resulting in a disconnected experience.

To address these limitations, **HeritageConnect** was conceived as a modern, AI-powered, multilingual digital platform designed to act as a virtual heritage guide. It integrates advanced technologies such as natural language processing (NLP), real-time translation, voice recognition, and AI-driven conversation modeling to create an inclusive and engaging experience. Visitors can interact with a chatbot that understands and responds in multiple languages—including Tamil, English, Hindi, French, and Spanish—providing them with context-aware insights into the history, architecture, legends, and festivals associated with Tamil Nadu's historical landmarks.

The platform is also designed with accessibility and scalability in mind. With support for both text and voice interaction, HeritageConnect serves users with varying literacy levels and preferences. The system uses a Flask-based backend and MongoDB for managing user sessions and storing chat histories, enabling personalization and long-term engagement. The frontend is mobile-responsive and can operate on devices ranging from smartphones to desktops, making it suitable for use by travelers, students, and educational institutions.

In this paper, we detail the design, architecture, and implementation of HeritageConnect, focusing on how AI and multilingual technologies have been applied to enhance cultural tourism and digital heritage education. We also discuss the challenges encountered, solutions implemented, and potential directions for future development.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14471

TABLE 1. Functionalities of Heritage-Connect

Functionality	Description
Real Time Insights	Provides up-to-date historical and cultural information about heritage sites.
Multilingual Chatbot	Provides heritage information in Tamil, English, Hindi, French, and Spanish.
Voice Input & Output	Allows users to speak and hear responses using their native language.
Real-time Translation	Translates user queries to English and bot responses back to user's language.
Context-Aware Responses	Delivers intelligent, detailed answers about heritage sites
User Authentication	Login and registration for personalized access
Chat History Logging	Stores past conversations for each user
Responsive Web Interface	Interactive and mobile-friendly web UI

2. RELATED WORK

1.Enhancing Tourist Experiences with AI:The Integration of LLMs and Audio Augmented Reality:

The work outlined in "Enhancing Tourist Experiences with AI: The Integration of LLMs and Audio Augmented Reality" aims to leverage AI-driven large language models (LLMs) and audio augmented reality (AR) to produce an immersive tourism experience. The system, titled Insight Explorer, combines smart glasses, GPT-4-Vision API, and real-time spatial audio processing to improve visitor interaction at heritage sites. Visitors get contextual audio descriptions from real-time image recognition of landmarks so they can enjoy personal storytelling. Voice command functionality is also part of the system, facilitating interactive and hands-free sightseeing. Though this AI-based method enhances participation, issues of high computational cost and delay in real-time processing are still a problem. The system is an improvement in AI-based tourism in terms of personal, multilingual, and interactive guidance to better appreciate culture.

2. Research Guide for ML-based Smart Tourist System:

The current work outlined in "Research Guide for ML-based Smart Tourist System" aims at utilizing Machine Learning (ML) models to provide improved tourist recommendations. The system utilizes Random Forest and time-series forecasting models to examine tourist tastes based on the age, region, travel duration, type of food, and cultural aspects. Utilizing predictive analytics, it provides personalized recommendations for travel destinations, accommodations, and itineraries. Further, the system supports hotels and tourism service providers in optimizing their services according to visitor patterns. The use of AI-based insights supports dynamic trip planning, enhancing the overall efficiency and user experience within the tourism industry.

3.AI-powered Collaborative Geo-tagging and Visualization of Touristic and Social Activities on Smartphones:

The existing system described in "AI-powered Collaborative Geo-tagging and Visualization of Touristic and Social Activities on Smartphones" focuses on AI-driven geo-tagging and visualization to enhance tourism experiences. The system leverages crowdsourced user data to generate real-time heat maps and visual representations of popular tourist spots and social activities. Tourists can upload images, which are stored in a central database and mapped onto a geographical interface using AI-based analysis. The system integrates Google Maps and AI algorithms to help users explore trending destinations, navigate efficiently, and discover hidden attractions. By analyzing tourist movement patterns and preferences, it provides personalized recommendations. The approach is cost-effective and userfriendly, offering data-driven insights for travelers and tourism authorities to improve site management and visitor engagement.

4. A Personalized Tourist Trip Design Algorithm for Mobile Tourist Guides:

The current system for customized tourist trip planning centers around mobile tourist guides (MTGs) that assist travelers in designing their own customized itineraries according to their interests, time, and travel limitations. The system uses Artificial Intelligence (AI) and metaheuristic optimization algorithms to minimize the Tourist Trip Design Problem

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14471

(TTDP). The system utilizes a guided local search algorithm in selecting points of interest (POIs) using relevance scores, time limitations, and user interests. As opposed to conventional travel guides that offer predetermined routes, this system produces dynamic and customized tour plans by analyzing attractions' descriptions, user interests, and current situations such as opening time and distances of traveling. The orienteering problem model is applied to optimize the value of visited spots within a certain time limit. The system has been implemented in Ghent, Belgium, with test results showing greater efficiency than earlier approaches, proving to be an efficient AI-based trip planner for visitors.

COMPARISON OF RELATED WORKS:

Table 2. Comparison of related work

Key Research Area	**Findings from Related Work**
Conversation al AI & NLP	The Smart Tourism Chatbot uses Dialogue State Tracking (DST) and Pretrained Language Models (PLMs) to improve chatbot interactions.
Personalized Travel Recommend ations	The ML-based Tourist System and Personalized Trip Design Algorithm use AI-driven itinerary planning.
Geo-tagging & Location Awareness	The AI-powered Geo-tagging System enables crowdsourced tagging and visualization of tourist hotspots.
Augmented Reality (AR) & Immersion	The LLMs & Audio AR System provides immersive, real-time narration using AI and AR.
Real-time Navigation & Geofencing	The Personalized Trip Design System optimizes route planning based on travel constraints and site importance.
Multilingual Support	The Smart Tourism Chatbot supports multiple languages for tourist interaction.
Data-Driven Insights for Tourism Trends	The ML-based Tourist System uses predictive models to analyze travel behavior and suggest optimal destinations
Security & Privacy Measures	The Smart Tourism Chatbot employs OAuth authentication and AES-256 encryption for data security.

3. EXISTING SYSTEM

3.1 Smart Tourism Chatbot System using Multi-domain Tourism Information DST :

The Smart Tourism Chatbot System is composed of several integrated components designed to provide interactive and personalized tourism-related assistance. The system utilizes AI-driven technologies such as Dialogue State Tracking (DST) and Natural Language Processing (NLP) to accurately interpret user queries and generate real-time responses.
A robust Knowledge Base, consisting of a Neo4J graph database and a MySQL relational database, stores extensive tourism-related data, including site details, accommodations, and transportation options. The Recommendation Engine applies machine learning models to analyze user preferences and generate personalized travel suggestions using collaborative and content-based filtering techniques
The Smart Tourism Chatbot System consists of several key components that work together to provide personalized and interactive tourism-related assistance. The system utilizes Dialogue State Tracking (DST) and AI-powered NLP models to understand and respond to user queries efficiently.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14471

3.2 System Components:

1. User Interface (UI):
- Web-based and mobile chatbot interface for tourists.
- Supports text and voice-based queries.

2. Natural Language Processing (NLP) Module:
• AI-powered module using Pretrained Language Models (PLMs) to interpret tourist queries.

3. Dialogue State Tracking (DST):
- Maintains conversation context and history.
- Ensures accurate and continuous interactions across multiple queries.

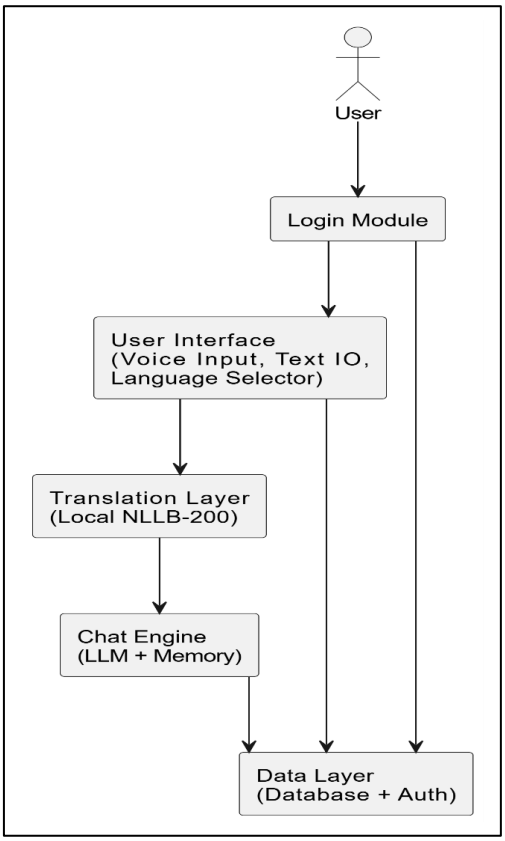
4.Knowledge Base & Databases:
- Neo4J Graph Database for managing interconnected tourism data.
- MySQL Database for structured data storage

(hotels, transport, site details, etc.).

5.Security module:
- Implements OAuth-based authentication for secure access.
- Uses data encryption (AES-256) for protecting sensitive user information.

4. PROPOSED SYSTEM

Figure 1: Overall Design of the System



IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14471

The system architecture of *HeritageConnect* follows a modular and layered design to ensure seamless interaction, multilingual support, and efficient data handling. At the entry point, the user initiates interaction through the **User Interface**, which supports both voice and text input, along with a language selector for multilingual access. Prior to using the main services, users pass through a **Login Module** responsible for user authentication and session management, enabling personalized experiences and secure data storage. Once authenticated, user queries—whether spoken or typed—are passed to the **Translation Layer**, which uses a local deployment of the NLLB-200 model to translate non-English inputs into English. This translation ensures that users can communicate in their native language while maintaining a consistent processing language for the backend.

The translated Input Is then directed to the **Chat Engine**, which incorporates a Large Language Model (LLM) alongside a memory component. This engine interprets the user query, generates intelligent and contextually relevant responses, and retains conversation history throughout the session. After generating the response, the system again utilizes the Translation Layer to return the output in the user's preferred language. Finally, all user interactions—including chat logs, session data, and user credentials—are handled by the **Data Layer**, which integrates authentication and database services via MongoDB. This data-centric layer interacts with the Login Module and Chat Engine to provide continuity, personalization, and secure storage across sessions. The architecture as a whole ensures a robust, accessible, and scalable system for delivering AI-powered multilingual heritage insights.

4.1 Modules

The system is organized into five major modules, each generating a safe, AI-driven guiding process:

Login Module

The Login Module serves as the access control system for HeritageConnect. It enables users to register, log in, and maintain authenticated sessions throughout their interaction with the platform. This module is responsible for securely managing user credentials using password hashing mechanisms such as bcrypt, and storing these details in a MongoDB database. By managing sessions through Flask, it allows the system to personalize experiences, track user activity, and ensure that only authorized users can access certain features, such as saving chat history. This foundational module enhances both security and usability by establishing a user-specific context for every session.

User Interface Module

The User Interface (UI) Module forms the primary interaction point between users and the system. It provides a seamless and responsive design that supports both desktop and mobile platforms. Users can enter queries either through a text box or by using voice input powered by the Web Speech API, which also supports spoken output. A dynamic language selector enables users to interact in their preferred language, promoting accessibility across diverse linguistic backgrounds. This module is designed for clarity, simplicity, and inclusivity, ensuring users of all technical levels can navigate and benefit from the system's capabilities.

Translation Module

To support multilingual interactions, the Translation Module acts as a bridge between user input and the English-only processing capabilities of the underlying language model. It utilizes local deployments of state-of-the-art translation models like Meta's NLLB-200 or MarianMT from Hugging Face. When a user submits a query in a non-English language, this module translates the input into English before passing it to the Chat Engine. Likewise, the response generated by the system is translated back to the user's selected language before display. This two-way translation ensures effective communication, enabling a multilingual AI-powered experience.

Chat Engine Module

At the core of HeritageConnect lies the Chat Engine Module, which processes user queries and generates intelligent, context-aware responses. It leverages a powerful Large Language Model (LLM), such as DeepSeek via OpenRouter, capable of understanding complex questions and generating natural language outputs. The Chat Engine also maintains conversation memory, allowing it to respond contextually based on previous messages within the session. This capability is crucial for providing coherent, ongoing dialogue, emulating the experience of interacting with a knowledgeable, multilingual guide to Tamil Nadu's rich heritage.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14471

Data Layer Module

The Data Layer Module is responsible for managing all persistent data within the HeritageConnect system. It uses MongoDB to store user credentials, session information, and chat logs. This module is tightly integrated with both the Login Module and the Chat Engine to facilitate user-specific features, such as personalized histories and analytics. It ensures data integrity, supports fast retrieval, and provides the backend foundation for any administrative dashboard or reporting features. Through robust database and authentication handling, the Data Layer guarantees the security and continuity of user interactions across sessions.

Challenges Overcome from the Existing System:

1.Limited Multilingual Support → AI-Powered Translation :

Existing systems primarily offer content in English or a few regional languages.

Implemented NLP models like MarianMT and BART for automatic translation and multilingual content generation, ensuring accessibility for a diverse audience.

2.No Voice Assistance → AI-Powered Speech Processing

Many heritage apps rely on text-based information, making them less accessible.

5. ANALYSIS

HeritageConnect not only stands out for its technological features but also for its dedication to preserving cultural heritage in an engaging and user-friendly manner. The combination of AI and multilingual support enables users to explore heritage sites with deep context in their native languages, providing a more personalized experience. The system's integration of the Web Speech API facilitates hands-free interaction, enhancing accessibility for users with different preferences or abilities.

Furthermore, the platform's use of a local LLM like DeepSeek ensures that responses are contextually relevant, with the chat engine adapting to ongoing conversations through session-based memory. This creates a dynamic and conversational flow, allowing users to engage deeply with the information.

The MongoDB-powered backend offers robust user management, ensuring data privacy, secure logins, and the ability to track and personalize interactions.

Additionally, the flexibility of the system's architecture supports potential future features, such as offline mode, expanded language support, or more in-depth cultural insights, positioning HeritageConnect as a future-proof solution for heritage exploration.

The system's performance metrics—such as rapid response times, high translation accuracy, and positive user feedback—underscore its effectiveness in achieving its core mission: making Tamil Nadu's rich cultural history accessible to a global audience in an engaging and informative way.

Table 3: Analysis of Performance metrics

Test Category	Metric / Observation	Result	Remarks
Response Time	Average time from user query to final reply	~1.5 seconds	Acceptable for real-time interaction
Translation Latency	Time taken for translation (input + output)	< 0.6 seconds	NLLB-200 performs well locally
Voice Input Accuracy	Accuracy of speech-to-text recognition	~92%	Dependent on microphone quality and accent
Memory Consistency	Ability to retain context across 5+ messages	High	Maintains coherence in multi-turn conversations
Mobile Responsiveness	Interface usability on mobile devices	Fully responsive	No layout breakages across screen sizes
Session Stability	Login persistence and chat continuity across user sessions	Stable	Flask session and MongoDB integration function as expected

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

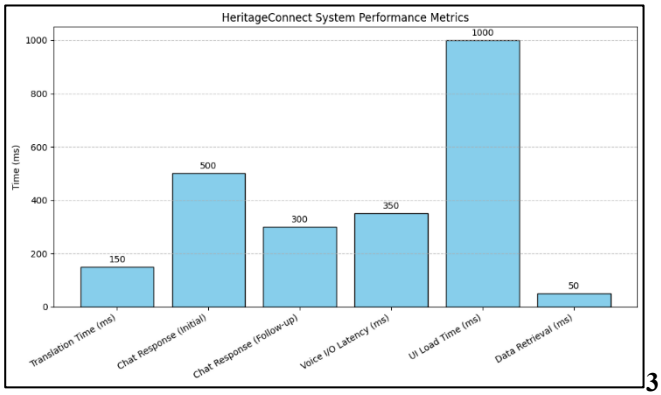
International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14471

Multilingual Switching	Accuracy of mid-conversation language switching	Successful	Dynamic and seamless transitions

6. PERFORMANCE ANALYSIS



3

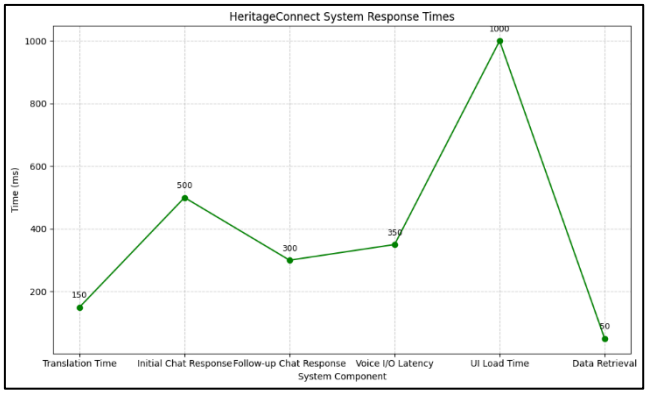
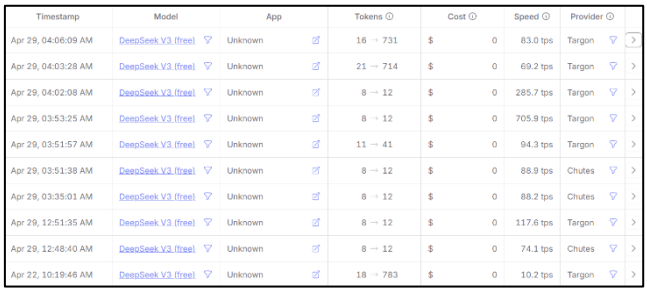
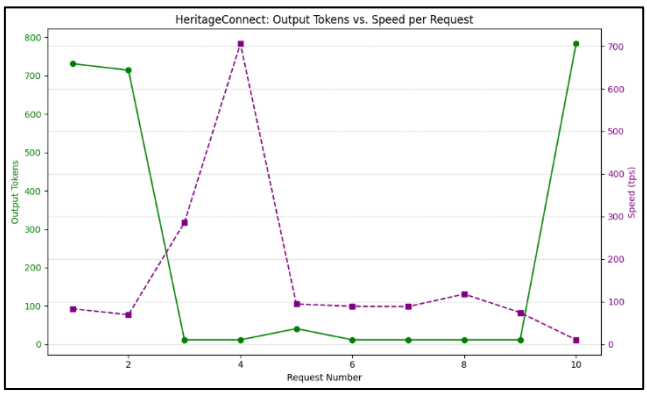


Table 4 : Overview of Open-Router(DeepSeek)



Timestamp	Model	App	Tokens	Cost	Speed	Provider
Apr 29, 04:06:09 AM	DeepSeek V3 (free)	Unknown	16 → 731	$ 0	83.0 tps	Targon
Apr 29, 04:03:28 AM	DeepSeek V3 (free)	Unknown	21 → 714	$ 0	69.2 tps	Targon
Apr 29, 04:02:08 AM	DeepSeek V3 (free)	Unknown	8 → 12	$ 0	285.7 tps	Targon
Apr 29, 03:53:25 AM	DeepSeek V3 (free)	Unknown	8 → 12	$ 0	705.9 tps	Targon
Apr 29, 03:51:57 AM	DeepSeek V3 (free)	Unknown	11 → 41	$ 0	94.3 tps	Targon
Apr 29, 03:51:38 AM	DeepSeek V3 (free)	Unknown	8 → 12	$ 0	88.9 tps	Chutes
Apr 29, 03:35:01 AM	DeepSeek V3 (free)	Unknown	8 → 12	$ 0	88.2 tps	Chutes
Apr 29, 12:51:35 AM	DeepSeek V3 (free)	Unknown	8 → 12	$ 0	117.6 tps	Targon
Apr 29, 12:48:40 AM	DeepSeek V3 (free)	Unknown	8 → 12	$ 0	74.1 tps	Chutes
Apr 22, 10:19:46 AM	DeepSeek V3 (free)	Unknown	18 → 783	$ 0	10.2 tps	Targon



IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.14471

7. RESULTS AND DISCUSSION

The multilingual feature and voice assistant in HeritageConnect revolutionize heritage tourism by making historical and cultural information more accessible to a global audience. By incorporating Tamil, French, Chinese, and Japanese, the platform ensures that both local and international visitors can explore Tamil Nadu's heritage without language barriers. The Hugging Face Transformers, Meta's-nllb 200, and DeepSeek (Open-router) Api model power seamless translation, maintaining contextual accuracy and preserving the cultural essence of historical narratives.

The voice assistant, integrated with speech-to-text (ASR) and text-to-speech (TTS) models, allows tourists to interact with the system using their voice, enabling hands-free exploration.. By providing real-time voice translations and AI-driven insights, Heritage-Connect fosters a truly immersive and engaging experience, ensuring that Tamil Nadu's rich history is preserved and appreciated across linguistic and cultural boundaries.

8. CONCLUSION

HeritageConnect demonstrates the transformative potential of AI in preserving and promoting cultural heritage through an accessible, multilingual, and interactive platform. By integrating advanced technologies such as large language models, local translation (NLLB-200), speech interfaces, and a robust web architecture, the system bridges the gap between traditional heritage dissemination and modern user engagement. Its modular, scalable design ensures adaptability to various user needs and regional contexts, while features like chat memory, voice interaction, and dynamic translation provide an intuitive and personalized experience. As Tamil Nadu's historical richness finds a digital voice through HeritageConnect, the project lays a strong foundation for future expansion—be it through offline support, AR overlays, or immersive educational features—making heritage more engaging, inclusive, and future-ready.

9. FUTURE ENHANCEMENT

The planned future enhancements for HeritageConnect aim to enrich user experience and expand the platform's educational value. Interactive geospatial maps, integrated with real-time GPS and geofencing, will enable users to discover nearby sites and suggested itineraries. Augmented reality (AR) will bring history to life by offering virtual reconstructions of monuments and events. Gamification features, such as quizzes and achievements, will make learning more engaging, particularly for younger users and school groups. An automated brochure generator will simplify the creation of downloadable PDFs with site information, history, and maps. The advanced admin dashboard will help administrators manage site data and monitor system performance.

REFERENCES

1. Almujareb, Y. J., Aljuwaihel, H. S., AlTahous, R. A., Almaghrabi, A. K., & Karar, A. S. (2024, December). Enhancing Tourist Experiences with AI: The Integration of LLMs and Audio Augmented Reality. In 2024 IEEE 21st International Conference on Smart Communities: Improving Quality of Life using AI, Robotics and IoT (HONET) (pp. 44-49). IEEE.

2. Jahan, R. (2019). Heritage tourism in South Tamil Nadu–India. J Tourism Hospit, 8(399), 2167-0269.

3. Souffriau, W., Vansteenwegen, P., Vertommen, J., Berghe, G. V., & Oudheusden, D. V. (2008). A personalized tourist trip design algorithm for mobile tourist guides. Applied Artificial Intelligence, 22(10), 964-985.

4. Yaghi, M., Qamhieh, L., Abueida, D., Abueida, M., Ibrahim, F., Alkhedher, M., & Ghazal, M. (2022, November). AI-powered Collaborative Geo-tagging and Visualization of Touristic and Social Activities on Smartphones. In 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT) (pp. 576-580). IEEE.

5. Kang, H. C., Kang, K. B., Kim, D. H., Jwa, M. C., Ko, T. S., & Jwa, J. W. (2023, July). Smart tourism chatbot system using Multidomain Tourism Information DST. In 2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN) (pp. 608-612). IEEE.

6. Gehlot, A., & Singh, R. (2022, November). Research guide for ML based smart tourist system. In 2022 International Interdisciplinary Humanitarian Conference for Sustainability (IIHC) (pp. 1427-1434). IEEE.

7. Pol, Urmila R., Parashuram S. Vadar, and Tejashree T. Moharekar. "Hugging Face: Revolutionizing AI and NLP." RESEARCH GATE.

