

ATLAS-Audio Translation for Live Attraction System

Dr. Vinutha H P¹, Ramya G S², Sanjana S K³, Shruti S P⁴, Vidya D V⁵

¹ Professor, Department of Information Science and Engineering, Bapuji Institute of Engineering, Davangere, Karnataka, India.

^{2,3,4,5} Department of Information Science and Engineering, Bapuji Institute of Engineering, Davangere, Karnataka, India.

To Cite this Article: Dr. Vinutha H P¹, Ramya G S², Sanjana S K³, Shruti S P⁴, Vidya D V⁵, “ ATLAS-Audio Translation for Live Attraction System”, Indian Journal of Computer Science and Technology, Volume 05, Issue 03 (September-December 2026), PP: 503-509.



Copyright: ©2026 This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution License; Which Permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract: Tourists visiting unfamiliar regions often face language barriers that make it difficult to understand local attractions, historical landmarks, and cultural information. To address this challenge, this paper presents ATLAS (Audio Translation for Live Attraction Systems), a web-based application designed to facilitate multilingual communication through speech processing and language translation. The system integrates speech-to-text conversion, automatic language detection, text translation, and text-to-speech synthesis to transform spoken input into translated audio output. Users can provide audio through the application, obtain a textual transcription, translate the content into a selected target language, and listen to the translated speech. The system is developed using Python and Flask, with dedicated modules for audio processing, speech recognition, language detection, translation, and speech synthesis. In addition, ATLAS incorporates tourism-related information to help users explore destinations and access relevant attraction details in text and audio formats. By combining these capabilities in a centralized platform, the proposed system aims to improve the accessibility of tourism information for people who speak different languages. The application demonstrates the potential of integrating speech technologies and artificial intelligence into tourism support systems. Further evaluation is required to measure translation quality, speech-recognition accuracy, and response time across different languages and real-world acoustic conditions.

Key Word: ATLAS, Speech Recognition, Automatic Language Detection, Machine Translation, Text-to-Speech, Artificial Intelligence, Multilingual Communication, Smart Tourism.

I. INTRODUCTION

The tourism industry plays an important role in promoting cultural heritage, historical landmarks, and regional attractions. However, tourists visiting unfamiliar destinations often face language barriers that make it difficult to understand local information and communicate effectively. Recent advancements in Artificial Intelligence (AI), Natural Language Processing (NLP), and speech-processing technologies provide opportunities to overcome these challenges through automated translation systems. ATLAS (Audio Translation for Live Attraction Systems) is a web-based application designed to facilitate multilingual communication and improve access to tourism information. The system integrates Speech-to-Text (STT), automatic language detection, machine translation, and Text-to-Speech (TTS) technologies to process spoken information and generate translated audio output. Users can provide audio input, obtain a textual transcription, translate the content into a selected target language, and listen to the translated speech. The application also provides information about tourist destinations and attractions, enabling users to explore places and access relevant details in text and audio formats. ATLAS is developed using Python and the Flask web framework, with dedicated modules for audio processing, speech recognition, language detection, translation, and speech synthesis. By combining these functionalities within a centralized platform, the system aims to reduce language-related difficulties and make tourism information more accessible to visitors from different linguistic backgrounds. The proposed application demonstrates the potential of integrating speech technologies and artificial intelligence into tourism services. Further evaluation can help assess its translation quality, speech-recognition accuracy, processing time, and effectiveness under different real-world conditions.

II. SCOPE OF THE PROJECT

The scope of ATLAS is to develop a web-based application that facilitates multilingual communication for tourists. The system converts spoken audio into text, detects the language, translates the content into a selected target language, and generates audio output. It aims to help tourists understand information about unfamiliar places and overcome language barriers. The application also provides information about tourist destinations and attractions in India through text and audio. The system integrates speech processing, language translation, and speech synthesis within a single platform. Future enhancements may include support for additional languages, improved translation accuracy, and real-time translation in noisy environments.

III. EXISTING SYSTEM AND PROPOSED SYSTEM

Existing System: Existing tourism systems mainly depend on human guides, printed materials, audio guides, and mobile translation applications to provide information to visitors. These systems may offer information in limited languages or require tourists to switch between different applications for translation and audio playback. Communication becomes challenging when tourists cannot understand the local language. Some systems also depend on internet connectivity and manual interaction, which may affect convenience during travel. Traditional audio guides generally provide predefined information rather than translating user-provided speech. Therefore, an integrated multilingual system is needed to simplify communication and improve access to tourism information.

Proposed System: The proposed ATLAS system provides a web-based platform to overcome language barriers in tourism through speech processing and translation technologies. It converts audio input into text using speech recognition and identifies the spoken language automatically. The recognized text is translated into the user's preferred language using an AI-based translation module. The translated text is then converted into speech, enabling users to listen to the information in their selected language. The system also provides information about tourist destinations and attractions in India in text and audio formats. By integrating these features into a single application, ATLAS aims to improve the accessibility of tourism information and enhance the travel experience for multilingual visitors.

IV. SYSTEM DESIGN

System design: ATLAS follows a modular architecture consisting of input, processing, speech generation, tourism information, and presentation components. These components work together to convert spoken input into translated text and audio output.

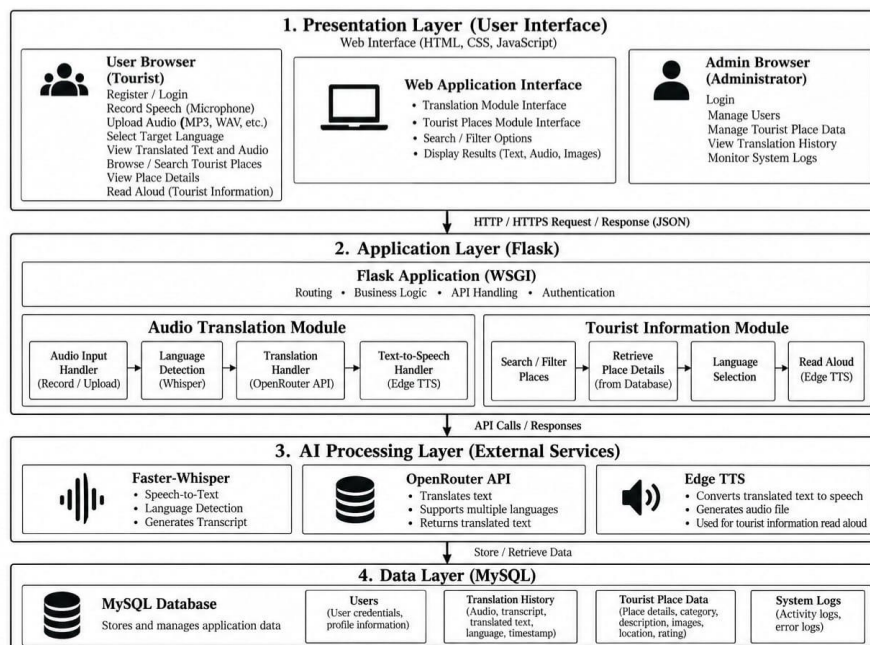


Figure no 1: System Design

- Presentation Layer (User Interface):** The presentation layer provides the web interface through which users interact with the ATLAS application. It supports user registration and login, audio recording and uploading, target-language selection, and displaying translated text and audio. It also provides access to tourism-related information, allowing users to explore destinations and view relevant attraction details.
- Application Layer (Flask):** Developed using Python and Flask, this layer handles user requests, authentication, routing, and communication between different modules. It coordinates audio processing, translation, speech generation, and tourism-related functionalities.
- Audio Translation Module:** This module accepts audio input and prepares it for speech recognition. Faster-Whisper converts speech into text and detects the spoken language, while the OpenRouter API translates the recognized text into the selected target language.
- Text-to-Speech Module:** This module uses Edge TTS to convert translated text into spoken audio. The generated audio is presented to users through the web interface for convenient listening.
- Tourism Information Module:** This module provides information about tourist destinations and attractions. It helps users explore places and access relevant details through the application's tourism features.
- AI Processing Layer:** This layer integrates Faster-Whisper for speech recognition, OpenRouter for translation, and Edge TTS for speech generation. These services work with the Flask application to complete the translation process.
- Data Layer:** The data layer uses the configured database and SQLAlchemy models to manage user information and other application records. It supports data storage and retrieval where implemented.

8. Complete System Workflow: Users provide audio through recording or file upload, and Faster-Whisper generates the transcript. OpenRouter translates the text, Edge TTS generates the translated speech, and ATLAS displays the available results to the user.

The ATLAS system was developed using Python, Flask, SQLAlchemy, Faster-Whisper, OpenRouter API, and Edge TTS. Functional testing was performed on the implemented modules to evaluate audio transcription, language translation, speech generation, and tourism information features.

V. RESULT

The implemented ATLAS application integrates speech recognition, language detection, text translation, and speech synthesis into a unified web-based platform. The system accepts audio input through microphone recording or file upload and uses Faster-Whisper to generate text transcriptions and identify the spoken language. The recognized text is translated into the selected target language using the OpenRouter API, and Edge TTS converts the translated text into an audio file for playback. The application also provides tourism-related information to help users explore destinations and understand tourist attractions. The modular architecture enables different components to work together, providing a structured workflow from audio input to translated speech output. The system demonstrates the potential of integrating artificial intelligence and speech-processing technologies to improve multilingual communication and accessibility of tourism information. Its performance can be further evaluated using measurable parameters such as transcription success rate, language detection accuracy, translation completion rate, audio generation success rate, and average processing time. Further testing with different languages, audio samples, and background noise conditions can help identify limitations and improve the reliability and effectiveness of the proposed system.

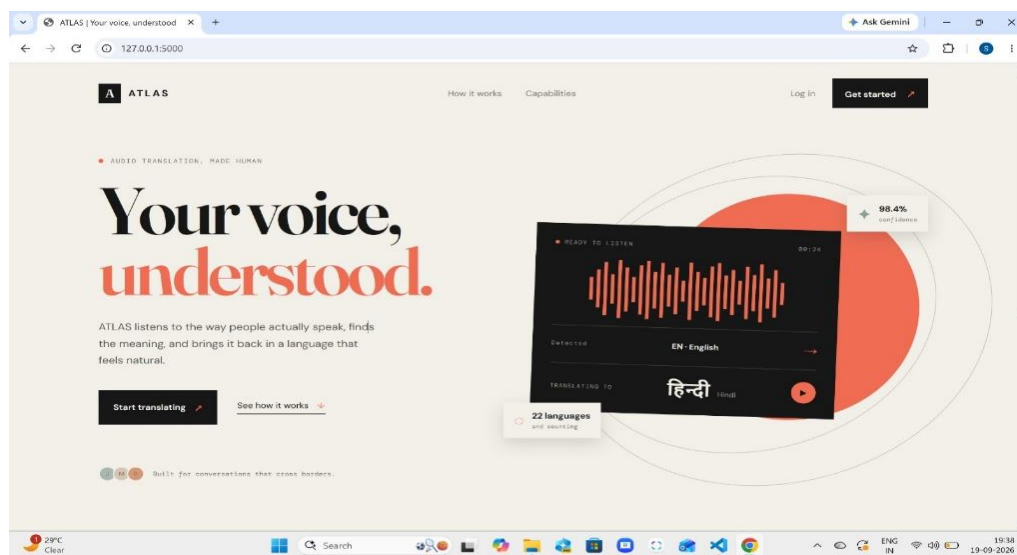


Figure no 1: ATLAS home page

The home page introduces the ATLAS audio translation system and its main purpose. It provides navigation options for users to learn about the system and begin using its features.

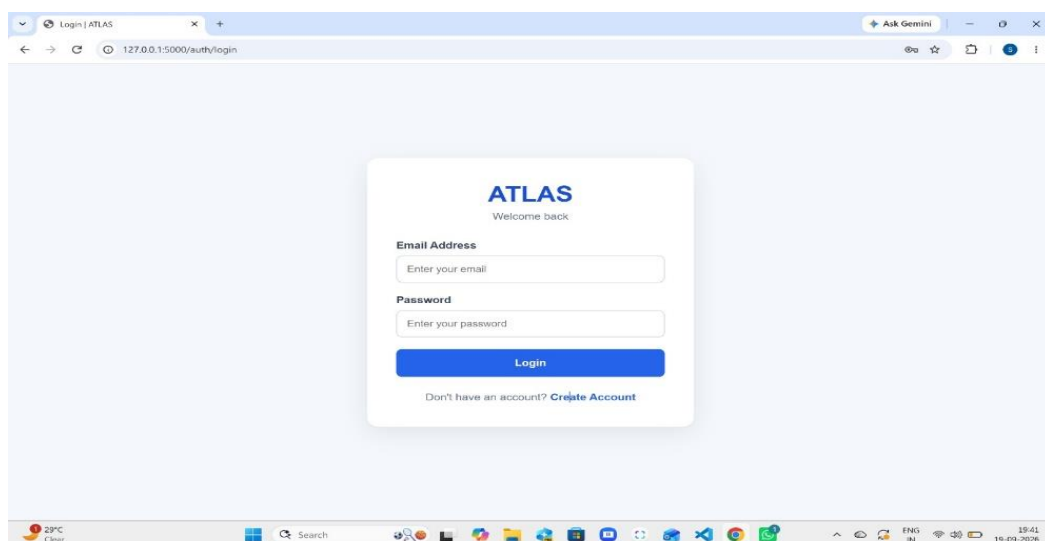


Figure no 2: User Login Page

The login page allows registered users to access the ATLAS application using their email address and password. It also provides an option to create a new account.

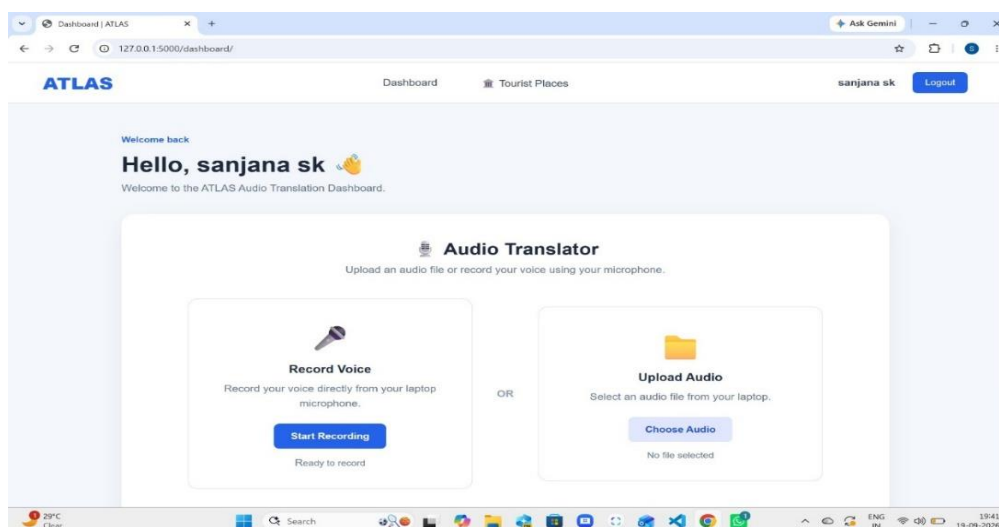


Figure no 3: Audio Translation Dashboard

The dashboard allows users to record audio through a microphone or upload an existing audio file. It provides the interface for initiating the audio processing and translation workflow.

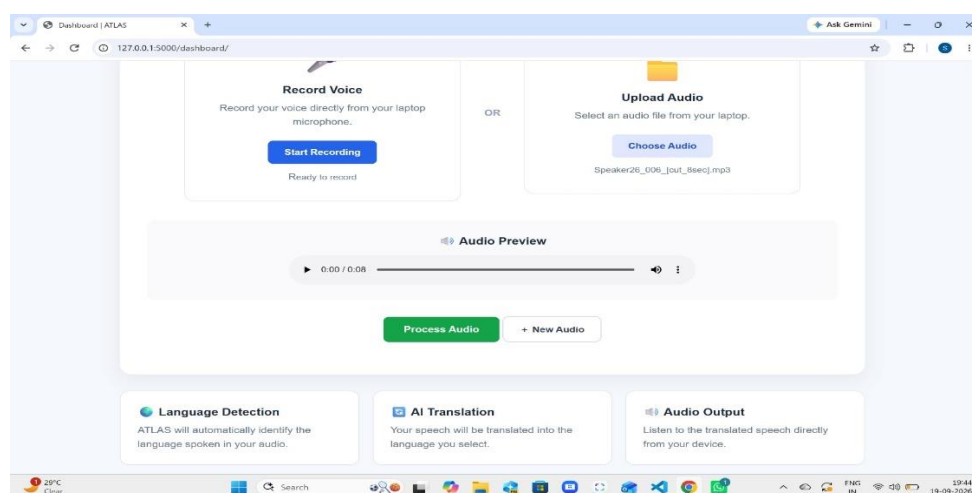


Figure no 4: Audio Input and Processing

This screen displays the selected audio file with an audio preview before processing. Users can listen to the recording and click Process Audio to begin speech analysis.

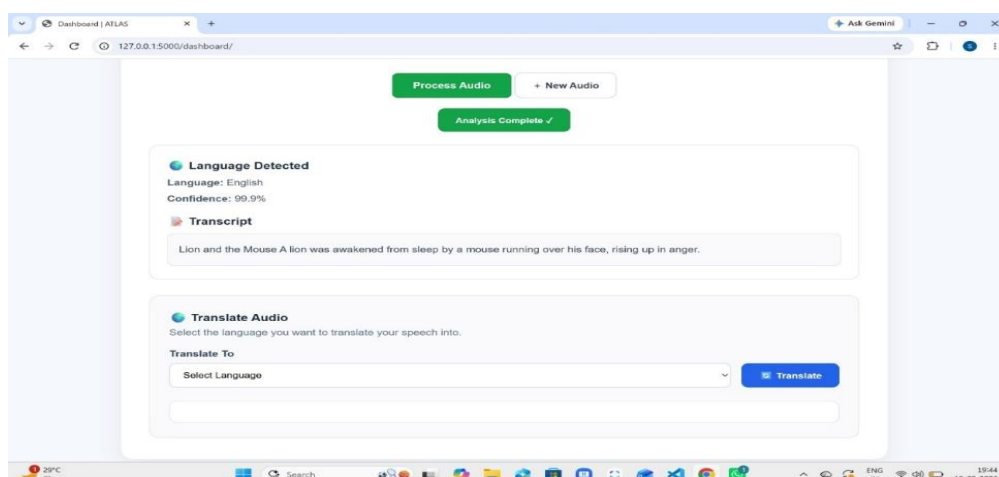


Figure no 5: Language Detection and Transcription

The system identifies the language spoken in the uploaded audio and displays the recognized speech as text. This allows users to review the transcription before translating it.

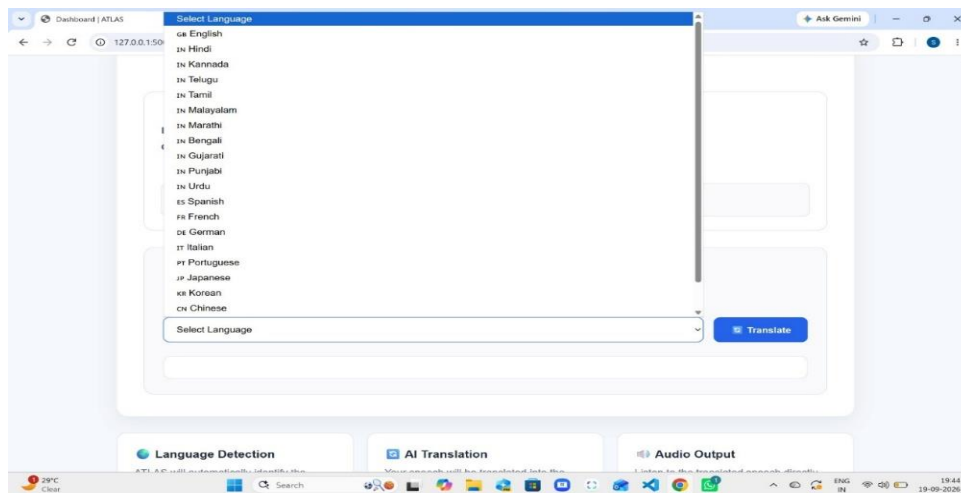


Figure no 6: Target Language Selection

The interface provides a list of available target languages, including Kannada, Hindi, Tamil, and other languages. Users can select their preferred language and initiate the translation process.

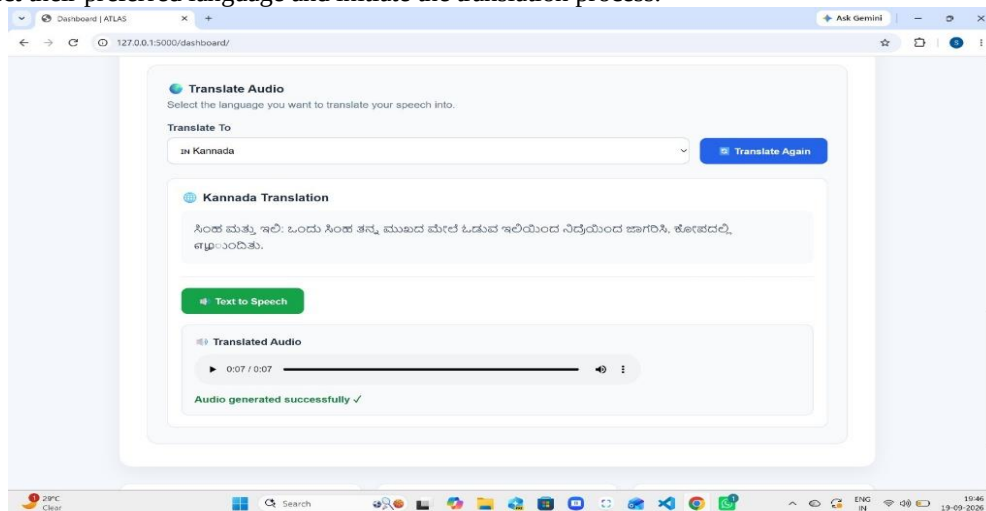


Figure no 7: Translated Text and Audio Output

The translated text is displayed in the selected language, and users can generate its corresponding speech output. The generated audio can be played directly through the application.

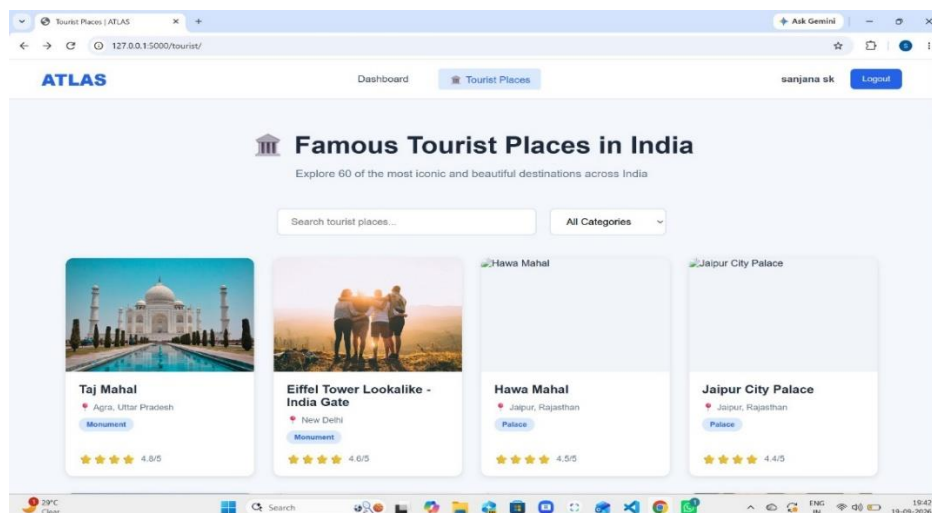


Figure no 8: Tourist Places Listing

The tourist places page presents popular destinations across India with their names, locations, categories, and ratings. Search and category filters help users find destinations of interest

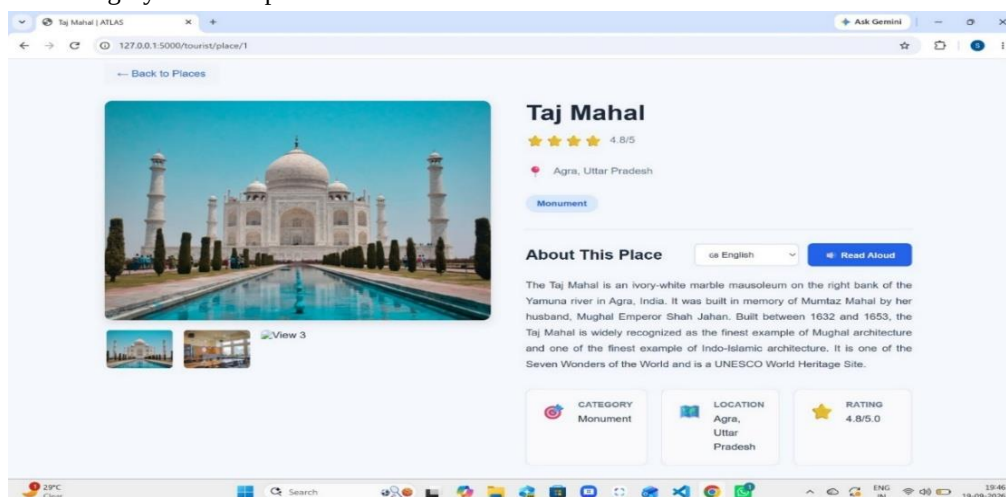


Figure no 9: Tourist Place Details

This page displays detailed information about a selected tourist attraction, including its description, location, category, rating, and images. It helps users learn more about the destination.

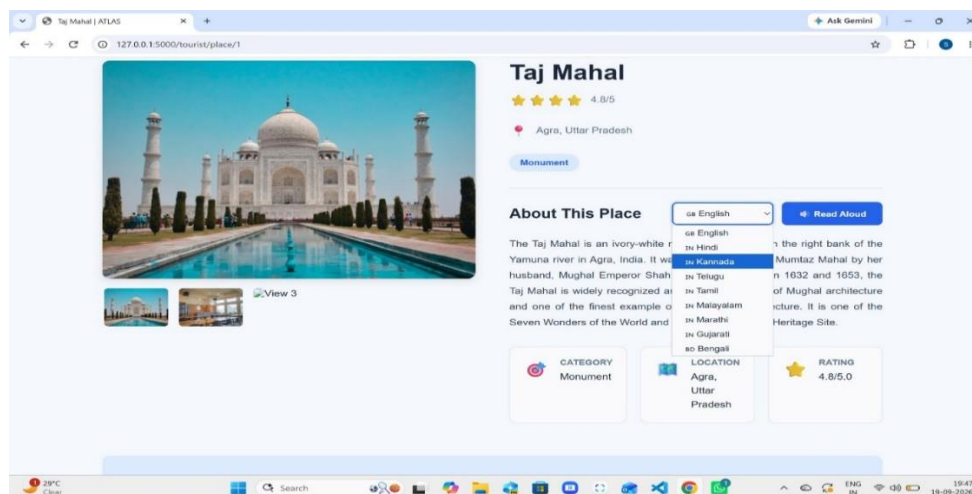


Figure no 10: Multilingual Tourist Information

The page includes a language selection option and a Read Aloud feature for tourist descriptions. These features help users access destination information in their preferred language through audio output.

VI. DISCUSSION

The primary contribution of ATLAS lies in integrating automatic speech recognition, language translation, and text-to-speech synthesis into a unified web-based platform for multilingual communication. The system uses Faster-Whisper for speech transcription and language detection, the OpenRouter API for translation, and Edge TTS for speech generation. By combining these components into a single workflow, ATLAS reduces the need to switch between separate applications and provides a streamlined process from audio input to translated speech output.

In addition to its core translation functionality, ATLAS provides a tourism information interface with attraction descriptions and a read-aloud feature for supported languages. The effectiveness of the system can be evaluated using measurable criteria, including Word Error Rate (WER) for transcription quality, translation adequacy and correctness assessed against human-verified reference translations, speech generation success rate for successful audio output, and average processing time measured from audio submission to the availability of each output. The overall success rate can also be calculated as the percentage of test cases that complete the required processing stages successfully. These metrics would provide a quantitative basis for evaluating reliability and identifying performance bottlenecks.

However, system performance may vary depending on audio quality, background noise, language selection, network connectivity, and the availability of suitable text-to-speech voices. A controlled evaluation using representative audio samples and multiple supported languages is therefore necessary to determine its practical effectiveness. Overall, ATLAS demonstrates the practical integration of existing AI services for multilingual and tourism-related communication, while establishing a foundation for future improvements in accuracy, processing efficiency, and language support.

VII. CONCLUSION

ATLAS (Audio Translation for Live Attraction System) presents a web-based solution for multilingual communication by integrating speech recognition, language translation, and text-to-speech synthesis into a unified workflow. The system uses Faster-Whisper for speech transcription and language detection, the OpenRouter API for translating recognized text, and Edge TTS for generating translated speech. It also provides tourism-related information with a read-aloud feature to help users access information about tourist attractions in supported languages. The modular architecture demonstrates the practical integration of existing AI services into a single application. However, further quantitative evaluation is required to assess transcription quality, translation correctness, speech generation reliability, and processing time. Future enhancements may include expanding language and voice support, improving performance under noisy audio conditions, and conducting comprehensive usability and accuracy testing. Overall, ATLAS provides a foundation for developing accessible, AI-assisted multilingual communication and tourism applications.

REFERENCES

1. A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust Speech Recognition via Large-Scale Weak Supervision," Proceedings of the 40th International Conference on Machine Learning (ICML), vol. 202, pp. 28492–28518, 2023. Available: <https://proceedings.mlr.press/v202/radford23a.html>
2. SYSTRAN, "Faster-Whisper: Faster Whisper Transcription with CTranslate2." GitHub repository. Available: <https://github.com/SYSTRAN/faster-whisper>
3. OpenRouter, "OpenRouter API Documentation." Available: <https://openrouter.ai/docs>
4. Edge-TTS Contributors, "edge-tts: Python Module for Microsoft Edge's Online Text-to-Speech Service." Python Package Index. Available: <https://pypi.org/project/edge-tts/>
5. Python Software Foundation, "Python Documentation." Available: <https://docs.python.org/3/>
6. Pallets Projects, "Flask Documentation." Available: <https://flask.palletsprojects.com/>
7. Oracle, "MySQL Reference Manual." Available: <https://dev.mysql.com/doc/>
8. Mozilla, "Web Speech API." MDN Web Docs. Available: https://developer.mozilla.org/en-US/docs/Web/API/Web_Speech_API