International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 04 April 2025, pp: 1003-1009 www.ijaem.net ISSN: 2395-5252

# Prompt Bridge: A Multilingual Prompt Engineering Framework to Bridge Language Barriers in AI Communication

Dr. Santosh K C[1], M Mokshitha[2], Mahima C[3], Mohammed Anas[4], Abhishek K[5]

[1]*Associate Professor, CSE dept. Bapuji Institute of Engineering and Technology, Davangere-04*
[2-5] *U GStudents, CSE dept. Bapuji Institute of Engineering and Technology, Davangere-04*

Date of Submission: 15-04-2025 Date of Acceptance: 25-04-2025

**ABSTRACT:** PromptBridge is an AI-powered communication system designed to eliminate language barriers between users and intelligent systems. The platform enables users—especially those from non-English-speaking backgrounds such as farmers, local vendors, and rural communities—to interact with AI in their native language. It performs a four-step workflow: accepting user input in the preferred language, translating it into English or a machine-readable format, optimizing the prompt for accurate AI interpretation, and translating the AI's response back into the original language. By combining prompt engineering and natural language processing, PromptBridge ensures more inclusive, accessible, and effective communication with AI models. This system not only enhances user experience but also broadens the reach of AI solutions to underserved populations, empowering them with technology regardless of linguistic limitations.

## I. INTRODUCTION

In the era of Artificial Intelligence, language remains a major barrier limiting access to AI-powered solutions.

Many users, especially in rural or regional areas, struggle to communicate with AI systems due to language constraints.PromptBridge is designed to bridge this gap by enabling users to interact with AI in their native languages.

The system translates user inputs into English or a machine-readable format for better AI understanding.

It then applies prompt optimization techniques to ensure accurate and relevant AI responses.

The responses are translated back into the user's original language for clarity and comprehension.

PromptBridge supports effective multilingual communication, making AI tools more inclusive. This project combines natural language processing and Prompt engineering in a streamlined workflow. It is especially beneficial for farmers, small business owners, and individuals with limited English proficiency.

PromptBridge aims to democratize access to AI technology by removing language as a barrier.

## EVOLUTION PROCESS

The development of PromptBridge began with identifying a critical challenge: the language barrier preventing non-English-speaking users from effectively interacting with AI systems. This led to the conceptualization of a solution that could translate user input into a machine-readable format, optimize it for AI processing, and then return the AI's response in the user's native language. The team gathered requirements focusing on natural language processing, prompt engineering, and user-friendly interface design. After selecting suitable technologies, including Python and NLP libraries, the system was structured into a four-step workflow: input translation, prompt optimization, AI response generation, and output translation. Individual components were developed and integrated into a functional prototype, followed by rigorous testing using diverse inputs. Based on user feedback, the system underwent refinement to improve translation accuracy and contextual understanding. The final product was deployed as a web-based application. Future improvements include voice-based input/output, support for more

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 04 April 2025, pp: 1003-1009 www.ijaem.net ISSN: 2395-5252

regional languages, and enhanced AI response handling

## II. LITERATURE REVIEW

| S.No. | Title | Author(s) | Year | Key Highlights | Relevance to Project |
|---|---|---|---|---|---|
| 1 | Automated Content Generation Using Generative AI Models | OpenAI | 2021 | Explores GPT-3 for generating human-like content using Few-Shot Learning; highlights challenges in coherence and personalization. | Demonstrates potential of prompt-based content generation in multilingual settings. |
| 2 | The State and Fate of Linguistic Diversity and Inclusion in the NLP World | Joshi et al. | 2020 | Critiques NLP's focus on high-resource languages; calls for inclusive data and models. | Supports the need for PromptBridge to include low-resource languages. |
| 3 | Data Statements for NLP: Toward Mitigating System Bias and Enabling Better Science | Bender & Friedman | 2018 | Proposes "data statements" to reduce bias; emphasizes demographic transparency. | Highlights ethical considerations in multilingual AI systems. |
| 4 | No Language Left Behind: Scaling Human-Centered Machine Translation | Fan et al. | 2022 | Meta AI's initiative to translate 200+ languages using multilingual models; stresses inclusive and open-source development. | Reinforces the vision of PromptBridge in reaching linguistic inclusivity. |
| 5 | OPUS-MT: Building Open Translation Systems for the World's Languages | Tiedemann &Thottingal | 2020 | Builds open translation systems for Indian and regional languages; advocates open-source tools. | Aligns with PromptBridge's open and accessible architecture for all languages. |
| 6 | Fine-tuning Multilingual Transformers for Low-Resource Languages | Hasan et al. | 2021 | Shows improved results via fine-tuning mBERT and XLM-R on low-resource datasets. | Informs PromptBridge's strategy for domain-specific prompt tuning. |
| 7 | Multilingual Tools in Education: A Comparative Study | Sowa et al. | 2021 | Demonstrates educational benefits of AI tools in native languages; improves engagement and learning. | Validates the societal impact of language-inclusive AI tools. |
| 8 | Unsupervised Cross-lingual Representation Learning at Scale | Conneau et al. | 2020 | Introduces XLM-R with strong zero-shot results; minimizes labeled data dependency. | Encourages PromptBridge's reliance on universal representations for prompt generation. |
| 9 | Advances in Natural Language Processing | Hirschberg & Manning | 2015 | Surveys major NLP developments; calls for improved non- | Provides foundational support for multilingual evolution in prompt |

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 04 April 2025, pp: 1003-1009 www.ijaem.net ISSN: 2395-5252

| S.No. | Title | Author(s) | Year | Key Highlights | Relevance to Project |
|---|---|---|---|---|---|
| | | | | English language handling. | engineering. |
| 10 | METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments | Banerjee & Lavie | 2005 | Proposes METEOR metric; better at evaluating semantic accuracy in translations. | Offers reliable evaluation method for assessing PromptBridge outputs. |

**EXISTING SYSTEM**

Several AI-driven chatbot systems have been developed to support multilingual and domain-specific communication in India. A multilingual chatbot for Indian languages enables fixed-response interactions across regional languages. Sevak, an open-source engine, supports both written and spoken inputs, allowing code-mixed communication in sectors like railways, healthcare, and judiciary. AI-based farmer chatbots, such as AgriBot and a fertilizer-predicting assistant, utilize NLP and computer vision to guide farmers in crop management and agricultural decisions. SchemeBot is designed to inform users about government schemes in healthcare, agriculture, and women's welfare using AI and machine learning. Collectively, these systems demonstrate the importance of multilingual support and intelligent, accessible interfaces tailored to specific domains, providing valuable insights for designing robust solutions like PromptBridge.

**PROBLEM STATEMENT**

In today's multilingual and multitasking world, users often struggle to access intelligent assistance that supports both regional languages and seamless voice interaction. Most AI assistants are limited by language barriers, lack of contextual understanding, or require constant switching between input methods. There is a growing need for an intelligent, multilingual conversational system that can understand, process, and respond to user queries—whether spoken or typed—in real time. This project aims to develop a smart assistant that combines speech recognition, translation, prompt optimization, and generative AI to offer reliable, natural, and multilingual support for everyday information retrieval and conversation**.**

**PROPOSED SYSTEM**

The proposed system is a multilingual conversational AI assistant that supports both text and voice-based interaction. It allows users to communicate in their native language and intelligently processes queries by integrating speech recognition, language translation, prompt optimization, and AI-powered response generation. For voice mode, it uses Whisper for transcribing speech to text, and for text mode, it directly accepts user input. The system translates all inputs to English using the NLLB model, optimizes long queries, and sends them to the Gemini AI model for context-aware responses. Responses are then translated back to the user's language and optionally spoken using gTTS. The assistant also integrates real-time APIs for weather, news, quotes, and definitions, ensuring dynamic, reliable, and human-like interaction.

**OBJECTIVES**

1. To develop a multilingual interface that allows users to interact with AI systems in their native or preferred languages, starting from today's need for inclusive and accessible technology.
2. To implement real-time translation and prompt optimization techniques that convert user queries into contextually accurate prompts for AI models and translate responses back into the original language without loss of meaning.
3. To support low-resource and regional languages by integrating scalable and adaptable language processing modules that address the communication needs of rural communities and non-English-speaking users.
4. To ensure user-friendly interaction with AI by creating a seamless end-to-end workflow that minimizes dependency on third-party translation tools and enhances the clarity, accuracy, and cultural relevance of AI-generated responses.

**SYSTEM REQUIREMENTS**

➢ **Hardware Requirements**

1. Processor: Intel Core i5 or higher (recommended: i7/i9 or equivalent AMD Ryzen for faster model execution)
2. RAM: Minimum 8 GB (recommended: 16 GB or more for smoother multitasking and model loading)

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 04 April 2025, pp: 1003-1009 www.ijaem.net ISSN: 2395-5252

3. Storage: At least 10 GB of free disk space (for dependencies, models, and logs)
4. GPU (Optional but recommended): NVIDIA GPU with at least 4 GB VRAM (for faster Whisper and Gemini processing if local LLMs are used)
5. Microphone: Built-in or external mic (for capturing user speech input)
6. Speakers/Headphones: For playing the text-to-speech

➢ **Software Requirements**

1. Operating System: Windows 10/11, macOS, or Linux (Ubuntu 20.04 or higher)
2. Python Version: Python 3.8 or above
3. Libraries and Frameworks:

- FastAPI (for backend API)
- Transformers (for BART, NLLB, and Gemini integration)
- sounddevice (for recording voice input)
- pydub and gTTS (for text-to-speech)
- faster-whisper (for speech-to-text)
- langdetect (for detecting language)
- scikit-learn, nltk (for TF-IDF keyword extraction)
- requests and urllib (for API calls)
- uvicorn (for running FastAPI server)

4. Browser: Google Chrome, Firefox, or any modern browser (for testing API and UI)

## III. METHODOLOGY

PromptBridge system adopts a modular, layered approach designed to facilitate seamless and accurate multilingual interactions between users and AI. The updated methodology leverages advanced models for voice recognition, language translation, and prompt optimization. The development pipeline is organized into five core phases:

**1. Voice & Text Input Processing**

The system begins by accepting input in two forms: text and voice. For voice-based inputs, Whisper-Medium, an open-source automatic speech recognition (ASR) model developed by OpenAI, is employed to transcribe speech into text with high accuracy, especially in low-resource Indian languages. The transcribed or directly typed text is then normalized and passed to the language detection module.

**2. Language Detection**

Language identification is performed using robust pre-trained models capable of handling a wide range of Indian regional languages and dialects. Accurate language detection ensures reliable downstream translation and semantic understanding, especially for linguistically complex or code-switched sentences.

**3. Translation & Prompt Optimization**

Once the input language is determined:

- The message is translated into English using NLLB (No Language Left Behind) by Meta AI. This neural machine translation model is optimized for high-fidelity translation in low-resource and multilingual settings, ensuring minimal semantic drift.
- The translated prompt is then refined using facebook/bert-base. This model is used for semantic enhancement and prompt optimization, focusing on improving clarity, preserving intent, and aligning the prompt structure for maximum understanding by generative models.

**4. AI Interaction Module**

The AI Interaction Module is the core component of the PromptBridge system, responsible for interpreting the user's processed input and generating meaningful, context-aware responses. After the input has been cleaned, translated to English (if necessary), and optimized for clarity, it is forwarded to this module for AI-driven understanding and response formulation.

This module leverages Google's Gemini API, a powerful generative language model, to analyze the user's prompt. Gemini is capable of handling a wide range of natural language tasks, such as question answering, content summarization, creative writing, and reasoning. The system sends the prompt to Gemini via an API request and awaits a response that is both semantically relevant and linguistically coherent.

The interaction within this module is driven by natural language processing and deep learning principles. Gemini uses a transformer-based neural network architecture to model the context and intent of the user's prompt

The response generated is then passed on to the Response Translation and Output Delivery module, where it is localized back into the user's original language and rendered as either text or synthesized voice, depending on the user's preference.

**5. Response Translation & Output Delivery**

The generated English response is translated back to the user's original language using NLLB, ensuring that nuances and contextual meanings are preserved. The final output is then

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 04 April 2025, pp: 1003-1009 www.ijaem.net ISSN: 2395-5252

delivered via a conversational interface that supports:

- Text-based and audio-based responses,
- Smooth dialog flow, and
- Interactive feedback loops for correction or rephrasing.

System Integration & Evaluation

The entire system is built using Python and modern web technologies. Open-source libraries, NLP toolkits, and APIs are seamlessly integrated to ensure:

- Real-time interaction,
- Scalability for broader deployment,
- Support for multiple Indian languages.

Extensive testing was conducted using real-world inputs in various regional languages (e.g., Hindi, Tamil, Bengali, Telugu) to validate the system's accuracy, latency, and semantic fidelity across the translation and interaction pipeline.

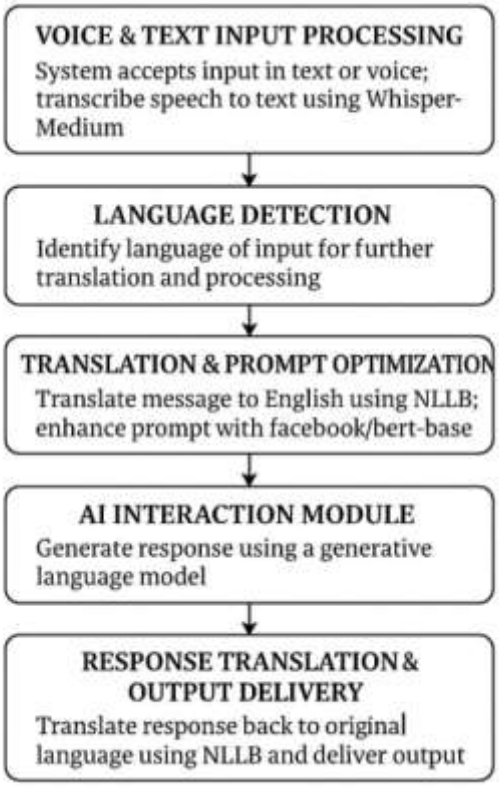


Fig. Methodology diagram of PromptBridge

## IV. IMPLEMENTATION

**Step 1:** System Initialization

All essential modules are initialized, including the Google Gemini API for response generation, the Whisper model for speech-to-text conversion, the NLLB model for translation, the gTTS module for speech output, and external APIs (GNews, OpenWeather, etc.) for enrichment.

**Step 2:** Mode SelectionThe assistant prompts the user to select either text or voice mode. This determines whether input will be processed as raw text or recorded audio.

**Step 3:** Input Acquisition

If the user selects text mode, the input is captured directly as a string. In voice mode, the system records audio in 5-second blocks using the sounddevice library. These are transcribed into text using the Whisper model with parameters: beam_size = 1 and vad_filter = True, optimizing transcription accuracy and silence removal.

**Step 4:** Language DetectionThe input text is passed through the langdetect library to determine its source language L. This sets the path for translation: Detected language L

Apply translation function T: $T(L) \rightarrow E$

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 04 April 2025, pp: 1003-1009 www.ijaem.net ISSN: 2395-5252

**Step 5:** Translation to English
The input is translated to English using the NLLB translation model. This ensures a consistent language base for summarization and AI response generation.

**Step 6:** Prompt Optimization (for long prompts)
If the translated prompt length > 300 characters, it is cleaned and summarized. Cleaning involves removing extra whitespaces, redundant punctuation, and non-informative tokens. Summarization is performed using BART or an extractive TF-IDF method.
The TF-IDF score for word w in document d is calculated using:
$TF\text{-}IDF(w, d) = tf(w, d) \times \log(N / df(w))$
where tf(w, d) is term frequency, N is the total number of documents, and df(w) is document frequency.

**Step 7:** API Enrichment (if applicable)
If the input contains keywords like "weather", "news", "quote", "definition", etc., corresponding APIs are triggered to fetch real-world data. The output of these APIs is passed to Gemini for summarization into a conversational response.

**Step 8:** AI Response Generation
The optimized or original prompt is sent to the Gemini AI model: $R = Gemini(P)$
where P is the cleaned/summarized English prompt and R is the generated English response.

**Step 9:** Translate Back to User's Language
The response R is translated from English back to the user's original language using the inverse function:
$T^{-1}(E) \rightarrow L$

**Step 10:** Speech Synthesis (Optional)
If voice response is enabled, the translated response is converted into speech using gTTS. The resulting audio is played using pydub.

**Step 11:** Output Delivery
The assistant returns the translated response (and keywords, if any). In voice mode, the transcribed input is also returned. The system then loops for the next input or ends based on user instruction.

## V. ADVANTAGES

**1. Multilingual Communication:**
Supports users in multiple Indian and international languages, making it inclusive and accessible to a wider audience.

**2. Dual Interaction Modes:**
Offers both text and voice-based interaction, providing flexibility and ease of use for different types of users.

**3. Real-Time Translation:**
Automatically detects and translates input/output languages, enabling smooth communication without language barriers.

**4. Intelligent Response Generation:**
Leverages the Gemini AI model to generate context-aware, meaningful, and human-like replies.

**5. Prompt Optimization:**
Handles long or complex queries efficiently by summarizing and extracting key information, improving response quality.

**6. Integration with Live APIs:**
Can fetch real-time information like weather, news, quotes, and definitions, making the assistant highly practical and **informative.**

**7. Voice Output Capability:**
Converts text responses to speech using gTTS, making it useful for visually impaired users or hands-free scenarios.

**8. Scalable and Modular:**
Built with modular components, making it easy to extend, customize, or integrate into other systems or applications.

## VI. APPLICATIONS

**1. Multilingual Virtual Assistants:**
The system can serve as a voice-enabled virtual assistant that supports multiple languages, making it useful for users in diverse linguistic regions.

**2. Customer Support Chatbots:**
Businesses can deploy it for handling customer queries via voice or text in native languages, enhancing accessibility and customer satisfaction.

**3. Educational Tools:**
It can assist students in learning new concepts, definitions, or summaries by interacting in their preferred language and using AI-generated explanations.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 04 April 2025, pp: 1003-1009 www.ijaem.net ISSN: 2395-5252

**4. Accessibility for Visually Impaired Users:**

With voice input and output capabilities, the assistant can help visually impaired users access information through natural conversation.

**5. Information Kiosks:**

The assistant can be integrated into public kiosks or help desks to provide real-time weather, news, or general information in various languages.

**6. Language Translation Companion:**

It can act as a real-time translation and communication tool between people speaking different languages.

**7. Smart Home or IoT Integration:**

The assistant can be extended to control smart home devices using natural language voice commands across languages.

## VII. CONCLUSION

The Multilingual Conversational AI Assistant successfully demonstrates the integration of speech recognition, language translation, prompt optimization, and AI-driven response generation within a unified system. By supporting both voice and text interactions in multiple languages, the assistant ensures accessibility, convenience, and user-friendliness for a wide range of audiences. Its ability to fetch real-time information and deliver context-aware, human-like responses makes it a powerful tool for daily assistance, education, and customer support. The modular design and use of advanced models like Whisper, NLLB, and Gemini further enhance the system's accuracy and flexibility. Overall, the project provides a scalable and intelligent solution that bridges communication gaps and elevates the user experience through natural, multilingual conversation.

## VIII. FUTURE SCOPE

The Multilingual Conversational AI Assistant has great potential for future enhancements and broader applications. One key direction is the integration of more regional and global languages using advanced translation models to support an even wider user base. The system can also be improved by incorporating emotion detection and sentiment analysis to make responses more empathetic and human-like. On-device deployment using lightweight models could enable offline access and better privacy, especially in rural or low-connectivity areas. Additionally, integrating visual input and output (e.g., image captioning, OCR, or GUI support) could expand usability across domains such as education, healthcare, and smart homes. Future versions can also include personalized responses, user profiles, and advanced memory features for maintaining conversational context over time.

## REFERENCES

[1] Dou, H., Tian, X., Lyu, X., Zhu, J., Li, J., & Guo, L. (2025). Speech Translation Refinement using Large Language Models.

[2] Liu, H., Chen, A., Chen, K., Bai, X., Zhong, M., Qiu, Y., & Zhang, M. (2025). Adaptive Inner Speech-Text Alignment for LLM-based Speech Translation.

[3] Mundnich, K., Niu, X., Mathur, P., Ronanki, S., Houston, B., Elluru, V. R., ... & Kirchhoff, K. (2024). Zero-resource Speech Translation and Recognition with LLMs.

[4] Goel, R., Ammar, W., Gupta, A., Vashishtha, S., Sano, M., Surani, F., ... & Yu, Z. (2023). PRESTO: A Multilingual Dataset for Parsing Realistic Task-Oriented Dialogs.

[5] Apple Machine Learning Research (2024). SELMA: A Speech-Enabled Language Model for Virtual Assistant Interactions.

[6] ResearchGate Publication (2025). Its Advancements in Natural Language Processing (NLP) and Its Applications in Voice Assistants and Chatbots.

[7] ScienceDirect Article (2024). Language Translation Effects in Chatbots: Evidence from a Field Experiment.

[8] Nature Article (2024). Joint Speech and Text Machine Translation for Up to 100 Languages.

[9] Kruk, F., Herath, S., & Choudhury, P. (2025). BanglAssist: A Bengali-English Generative AI Chatbot for Code-Switching and Dialect-Handling in Customer Service.

[10] Qin, L., Chen, Q., Zhou, Y., Chen, Z., Li, Y., Liao, L., Li, M., Che, W., & Yu, P. S. (2024). Multilingual Large Language Model: A Survey of Resources, Taxonomy and Frontiers.

[11] Biswas, S., Erlei, A., & Gadiraju, U. (2025). Mind the Gap! Choice Independence in Using Multilingual LLMs for Persuasive Co-Writing Tasks in Different Languages.

[12] ElevenLabs (2025).Exploring Conversational AI Breakthroughs in 2025.