International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 04 | April - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

AI-POWERED VOICE RECEPTIONIST FOR ATHARVA COLLEGE OF ENGINEERING USING VAPI.AI, TWILIO AND N8N AUTOMATION

Sanica Zore[1], Avnish Maurya [2], Sahil Sandanshiv[3] ,Amitesh Sanaganabhatla[4]

[1]*Department of Electronics and Telecommunication Engineering, Atharva College of Engineering*
[2]*Department of Electronics and Telecommunication Engineering, Atharva College of Engineering*
[3]*Department of Electronics and Telecommunication Engineering, Atharva College of Engineering*
[4]*Department of Electronics and Telecommunication Engineering, Atharva College of Engineering*

Abstract -This paper presents the design and implementation of an AI-powered voice receptionist system, named "Kanika", developed for Atharva College of Engineering (ACE), Mumbai. The system automates the management of incoming phone calls using a no-code and low-code technology stack comprising Vapi.ai for voice AI orchestration, Twilio for telephony infrastructure, GPT-4o as the underlying large language model, Azure Neural TTS (hi-IN-SwaraNeural) for bilingual Hindi-English speech synthesis, and n8n for post-call workflow automation. Kanika handles frequently asked questions related to admissions, fee structures, examination schedules, and placement activities. It collects caller information, routes complaints to appropriate departments, logs all interactions to Google Sheets, books appointments via Google Calendar, and dispatches real-time alerts to department heads via Gmail and WhatsApp. Tested at Atharva College of Engineering, the system demonstrates 24/7 availability, natural bilingual Hindi-English conversation capability, and significant reduction in receptionist workload.

Key Words: Voice AI, AI Receptionist, Conversational AI, Vapi.ai, Twilio, n8n Automation, GPT-4o, Bilingual NLP, Hindi-English AI, Educational Technology, No-Code AI, Google Workspace Integration.

1.INTRODUCTION *(Size 11, Times New roman)*

In today's fast-paced academic environment, educational institutions face increasing pressure to improve administrative efficiency while maintaining a high quality of service for students, parents, and visitors. One of the most resource-intensive administrative functions is the management of incoming telephone calls. At Atharva College of Engineering (ACE), an AICTE-approved, NAAC-accredited institution affiliated to the University of Mumbai, the front-desk receptionist handles a large volume of daily calls covering admissions, fee structures, examination schedules, timetables, placement activities, and appointment requests.

A significant proportion of these calls involve repetitive queries that do not require human judgment and can be systematically handled by an intelligent automated system. The limitations of traditional receptionist models — restricted to office hours, susceptible to human error, and unable to handle concurrent calls — create gaps in service availability and caller satisfaction.

This paper presents the design, development, and deployment of "Kanika", an AI-powered voice receptionist built specifically for Atharva College of Engineering. Kanika is capable of handling inbound telephone calls in both Hindi and English, understanding natural unscripted conversations, answering college-specific queries, collecting caller information, routing complaints to appropriate departments, booking appointments, and logging all interactions for administrative review. The complete technology stack includes Vapi.ai, Twilio, GPT-4o, Azure Neural TTS, Deepgram, n8n, Google Sheets, Gmail, WhatsApp, and Google Calendar.

2. Body of Paper

2.1 System Overview

The AI Receptionist System is an automation-based solution developed using n8n to handle user queries in real time. It simulates a human receptionist by using workflows integrated with AI models for natural language processing. The system captures user inputs from chat platforms or web interfaces and processes them to provide accurate, instant, and context-aware responses, especially for institution-related queries.

The architecture includes input handling, workflow automation, AI processing, and response delivery layers. n8n acts as the central controller, routing queries, applying logic, and integrating with external APIs or databases. The system can also store conversation logs and supports integration with platforms like WhatsApp or web chat, ensuring efficient, scalable, and automated communication.

2.2 Software Components

.1 System Components

[1] Telephony Gateway: Twilio provides the real phone number and PSTN/SIP routing, forwarding incoming audio to Vapi.ai.

[2] Voice AI Orchestrator: Vapi.ai manages the full voice conversation lifecycle, coordinating speech recognition, LLM inference, and speech synthesis.

[3] Language Model: GPT-4o processes transcribed caller input and generates contextually appropriate responses guided by the Kanika system prompt.

[4] Speech Synthesis: Azure hi-IN-SwaraNeural converts GPT-4o text responses to natural Hindi-English audio streamed to the caller.

[5] Automation Engine: n8n processes post-call webhook data and orchestrates logging, alerting, and scheduling.

[6] Data Outputs: Google Sheets for call logs, Gmail and Twilio WhatsApp API for department alerts, Google Calendar for appointment events.

2.3 Software Implementation

The Kanika AI receptionist system was tested through simulated and live call scenarios at Atharva College of Engineering, covering all six call handling flows across Hindi, English, and Hinglish inputs. The system successfully handled all FAQ query types including admissions eligibility, fee structure enquiries, examination schedule queries, and placement information. Bilingual call handling was achieved with natural language switching without any explicit language selection required from the caller.

Appointment booking calls successfully triggered Google Calendar event creation with caller details populated in the event description. Complaint calls triggered both Gmail alerts and WhatsApp notifications within seconds of call completion. All calls produced accurate Google Sheets log entries including caller name, phone number, query summary, department classification, action type, and timestamp.

2.4 Working Principle

The Kanika AI Receptionist system follows a structured approach including requirement analysis, system design, prompt engineering, voice configuration, and workflow automation. Initially, common queries across categories like admissions, fees, exams, placements, and complaints were identified, along with emergency cases requiring immediate escalation.

The system is designed as a multi-layer pipeline integrating telephony, voice AI, intelligence, automation, and output layers. Using n8n, workflows automate post-call processing such as logging, alerts, and scheduling. AI behavior is controlled through optimized prompts, while bilingual voice support ensures smooth Hindi-English interaction, enabling efficient and scalable automated communication.

2.5 Advantages of the Proposed System

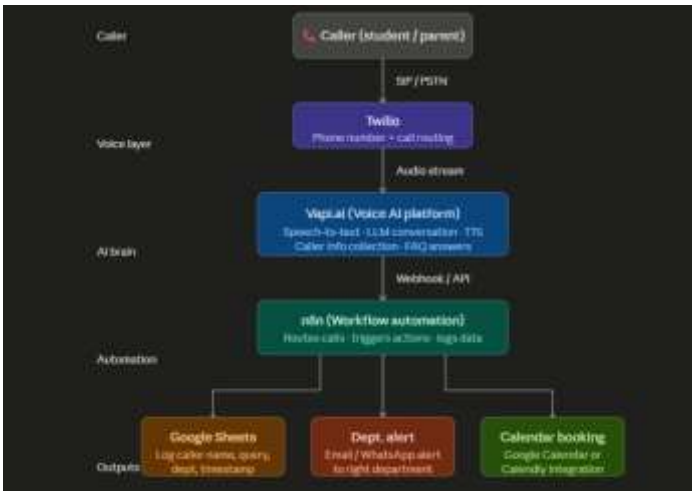
The proposed Kanika AI Receptionist system is a fully automated, end-to-end telephone call management solution for Atharva College of Engineering. The system intercepts all incoming calls to the college's Twilio-provisioned phone number and manages the complete caller interaction without human intervention for standard queries, escalating only genuine emergencies and complaints to human staff.

2.6 Limitations

The AI Receptionist system built using has certain constraints. It relies heavily on predefined prompts and knowledge bases, which may limit its ability to handle highly complex, ambiguous, or unexpected queries accurately. Voice recognition and language switching (Hindi-English) may also introduce errors in noisy environments or with varied accents, affecting response quality.

Additionally, the system depends on third-party services (AI APIs, telephony, and cloud platforms), making it vulnerable to latency, downtime, or API limitations. Real-time decision-making in critical scenarios (like emergencies) may not always be as reliable as human intervention, and continuous updates are required to maintain accuracy and relevance of information.

Fig -1: Block Diagram



3. CONCLUSIONS

This paper presents Kanika, an AI-powered voice receptionist system developed for Atharva College of Engineering, Mumbai. The system automates the complete lifecycle of incoming telephone call management, from initial greeting and natural language conversation to post-call logging, department alerting, and appointment scheduling.

The proposed system demonstrates that modern voice AI technologies can be practically deployed in educational institutions using accessible no-code and low-code platforms. By combining Vapi.ai, Twilio, GPT-4o, Azure Neural TTS, and n8n, the system achieves 24/7 availability, natural bilingual Hindi-English conversation, and seamless Google Workspace integration at a fraction of the cost of a traditional receptionist.

Testing confirmed that the system handles all six primary call categories with high accuracy across Hindi, English, and Hinglish inputs. Future work will integrate a live knowledge base using retrieval-augmented generation (RAG) for dynamic information, expand language support to Marathi, and incorporate administrative analytics dashboards derived from call log data.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to Atharva College of Engineering for providing the necessary facilities and support to carry out this work. The authors also extend their heartfelt thanks to Prof.Prajakta Pawar for her valuable guidance, encouragement, and continuous support throughout the development of this project.

REFERENCE

.[1] D. Jurafsky and J. H. Martin (2023), "Speech and Language Processing," 3rd ed., Prentice Hall, Ch. 26: Dialogue Systems and Chatbots.
[2] A. Vaswani, N. Shazeer, N. Parmar et al. (2017), "Attention Is All You Need," Advances in Neural Information Processing Systems (NeurIPS), Vol. 30.
[3] T. B. Brown, B. Mann, N. Ryder et al. (2020), "Language Models are Few-Shot Learners," Advances in NeurIPS, Vol. 33, pp. 1877–1901.
[4] M. Laranjo, A. G. Dunn, H. L. Tong et al. (2018), "Conversational Agents in Healthcare: A Systematic Review," Journal of the American Medical Informatics Association, Vol. 25, No. 9, pp. 1248–1258.
[5] R. Cui, S. Gallino, A. Moreno, and D. J. Zhang (2018), "The Operational Value of Social Media Information," Production and Operations Management, Vol. 27, No. 10, pp. 1749–1769.
[6] M. Winkler and K. Söllner (2018), "Unleashing the Potential of Chatbots in Education: A State-of-the-Art Analysis," Academy of Management Annual Meeting, Vol. 2018, No. 1.
[7] B. B. Chaudhary, S. Dandapat, A. Sharma, P. Bansal, and D. Varshney (2018), "A Hybrid Approach for Code-Mixed Language Identification," Proceedings of EMNLP, pp. 482–488.
[8] M. Fowler (2018), "Patterns of Enterprise Application Integration," Addison-Wesley Professional, Ch. 4: Messaging Channels.
[9] OpenAI (2024), "GPT-4o Technical Report," OpenAI. Available: https://openai.com/research/gpt-4o.
[10] Vapi.ai (2024), "Vapi Developer Documentation," Available: https://docs.vapi.ai.
[11] Twilio Inc. (2024), "Twilio Voice API Documentation," Available: https://www.twilio.com/docs/voice.
[12] Microsoft Azure (2024), "Azure Cognitive Services — Text to Speech," Available: https://azure.microsoft.com/en-in/products/cognitive-services/text-to-speech.
[13] Deepgram (2024), "Nova-2 Speech Recognition Model," Available: https://developers.deepgram.com/docs/nova-2.
[14] n8n GmbH (2024), "n8n Workflow Automation Documentation," Available: https://docs.n8n.io.