RaabtaAI

Project Advisor

Dr. Muhammad Saad Razzaq (Internal)

Project Manager

Dr. Muhammad Ilyas

Project Team

Muhammad Asmaad Saeed	BSCS51F22S037	Team Lead
Laiba Siraj	BSCS51F22S046	Team Member
Fatima Zahid	BSCS51F22S010	Team Member
Areeba Rehman	BSCS51F22S039	Team Member

Submission Date

6 October 2025

Table of Contents

1. Abstract

In Pakistan, organizations across banking, telecom, healthcare, and government sectors face overwhelming volumes of customer support calls daily. Traditional call centers often lead to long wait times, inconsistent support quality, and high operational costs.

The proposed system, **RaabtaAI** is a Native Urdu-First Inbound Calling Agent, aims to automate inbound customer interactions using AI-driven speech and language understanding, designed specifically for the Urdu-speaking population. The system responds naturally in a native Urdu voice, understands caller intent and emotions, and either resolves the query autonomously or escalates complex or urgent cases to a human agent.

It integrates Automatic Speech Recognition (ASR), Natural Language Understanding (NLU), Knowledge Graphs, and GraphRAG reasoning to deliver intelligent, real-time, and context-aware responses. The goal is to enhance accessibility, reduce costs, and provide consistent 24/7 service.

Expected benefits include advancement in Urdu NLP research, increased automation efficiency in customer care, and a reduction in manual call-handling burdens across industries.

2. Background and Justification

In today's digital landscape, customer support plays a central role in shaping user experience and organizational reputation. However, traditional inbound call centers are burdened with long wait times, high costs, and inconsistent service quality. These challenges are more pronounced in Pakistan, where many users communicate in Urdu or a mix of Urdu-English, yet most automated systems remain English-centric, creating a language and accessibility gap.

With customers expecting real-time, 24/7, and personalized support, organizations are under increasing pressure to reduce costs, optimize resources, and improve efficiency. Existing solutions such as AutoCalls.ai offer generic flow-building tools for call automation, but they lack specialization in areas like Urdu-first interaction, code-switch handling, emotion and urgency detection, or secure multi-step verification.

Our project addresses these gaps by developing a native Urdu-first, agentic AI system designed specifically for the Pakistani context. Unlike generic workflow builders, this system integrates Automatic Speech Recognition (ASR), Natural Language Understanding (NLU), and Text-to-Speech (TTS) with local telephony gateways for low latency, while also incorporating advanced security pipelines such as OTPs and voice biometrics. By automating routine queries and providing empathetic, culturally relevant interactions, the solution ensures operational efficiency, accessibility, and trust.

3. Project Methodology

Our proposed system follows a modular, scalable, and secure architecture composed of 11 interconnected modules, with real-time communication via APIs and message queues (RabbitMQ/Kafka).
At the center lies the **Dialogue Manager**, acting as the decision-making "brain" orchestrating modules based on user intent and system state.

Core Modules and Flow:

1. Telephony Gateway Module

- Handles inbound calls from PSTN/VoIP.
- Establishes encrypted audio sessions using TLS/SRTP and forwards the audio stream to ASR.

2. ASR (Automatic Speech Recognition) Module

- Converts Urdu or Urdu-English mixed speech into text.
- Employs noise reduction and acoustic modeling for better accuracy in regional dialects.

3. NLU (Natural Language Understanding) Module

- Extracts intent, entities, sentiment, and urgency from text using fine-tuned transformer models (mBERT/XLM-R).
- Detects emotion and priority level for escalation logic.

4. Dialogue Manager (Central Brain)

- Coordinates system responses based on NLU results.
- Routes calls, fetches data from databases, triggers TTS output, or activates human escalation.

5. Knowledge Graph (KG) & GraphRAG Reasoning Module

- Maintains an interconnected graph of domain knowledge (e.g.,UOS student's records, Uni-merit lists record, rules, FAQs).
- Uses GraphRAG (Graph-based Retrieval-Augmented Generation) to traverse relevant nodes for contextual answers.
- Combines semantic search (via embeddings) with graph reasoning to ensure factual and consistent responses.
- Example: For a query like *"2025 me CS ka merit kahan close hoa tha"*, the KG links multiple entities (CS + Merit + Year) for a unified, contextual reply.

6. Knowledge Base (KB)

- Stores static or FAQ-style information and integrates with the Knowledge Graph for hybrid retrieval (semantic + graph).

7. Database / API Integration Module

- Connects to live data sources (UOS systems, telecom CRMs, databases).

- Handles authenticated calls for account balance, complaint status, etc.

8. TTS (Text-to-Speech) Module

- Converts text responses to natural, fluent Urdu voice via ElevenLabs, UpliftAI, or Play.ht.
- Maintains sub-2s latency and uses emotional tone modeling.

9. Human Escalation Module

- Transfers calls to human agents when emotion, urgency, or complexity is detected.
- Passes session context and transcript securely.

10. Security & Compliance Module

- Ensures end-to-end encryption, OTP verification, and anonymization of logs.
- Complies with GDPR, PDPL, and HIPAA (for healthcare sector).

11. Logging & Analytics Module

- Records anonymized transcripts, response metrics, and sentiment analysis.
- Provides retraining data and analytical dashboards for model improvement.

4.Project Scope

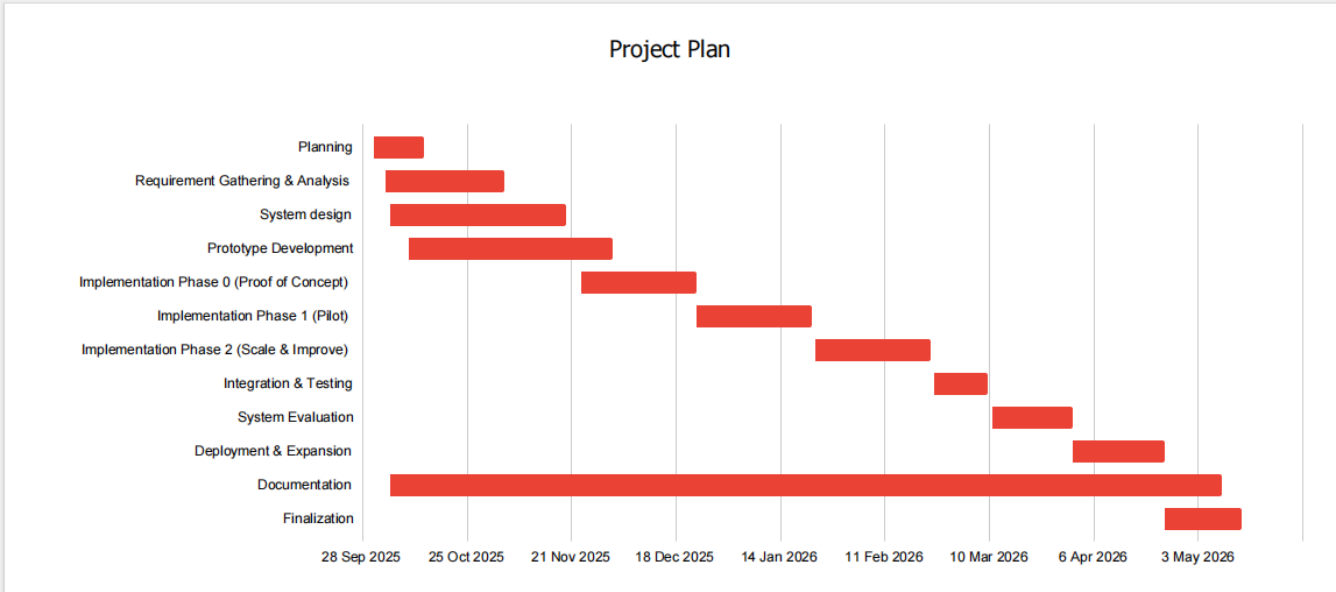
Included Functionalities:

- Automatically answers inbound calls and interacts with callers in real time.
- Uses speech recognition and natural language understanding (NLU) to identify intent and provide accurate responses.
- Supports bilingual communication in both English and Urdu.
- Retrieves essential information from the university's database, including admission schedules, departmental contacts, office timings, and event details.
- Allows seamless call transfer to a human operator for complex queries.
- Provides an analytics dashboard to monitor call volume, response accuracy, and performance.
- Includes basic emotion and sentiment detection to improve response quality and user experience.

Excluded Functionalities:

- Outbound or promotional calling.
- Video or chat-based communication.
- Integration with third-party CRM systems.
- External system integrations beyond the university's internal database.
- Can handle a limited number of calls simultaneously (up to five callers).

5.High level Project Plan



6. References

[1] World Bank, *Pakistan development update: Reimagining a digital Pakistan*, Apr. 2025. [Online]. Available: https://thedocs.worldbank.org/en/doc/e414b36ae736660edf8f0f3cb597b1e9-0310012025/pakistan-development-update-reimagining-a-digital-pakistan-april-2025

[2] Harvard Business Review, *Customer experience in the age of AI*, Mar. 2022. [Online]. Available: https://hbr.org/2022/03/customer-experience-in-the-age-of-ai

[3] PwC & Parloa, *Transforming customer service for a smarter, more human future*. Parloa, 2021. [Online]. Available: https://www.parloa.com/customers/pwc-and-parloa-transforming-customer-service-for-a-smarter-more-human-future/

[4] AutoCalls, *AutoCalls AI: Call flow builder*. AutoCalls, 2025. [Online]. Available: https://autocalls.ai/

[5] Gartner, *Market guide for conversational AI platforms*. Gartner Research, 2022. [Online]. Available: https://www.gartner.com/en/documents/5332563

[6] G. Cloud, "Speech-to-Text documentation," Google, 2025. [Online]. Available: https://cloud.google.com/speech-to-text/docs.

[7] Wikipedia, "Natural language understanding," WIKIPEDIA, 2025. [Online]. Available: https://en.wikipedia.org/wiki/Natural_language_understanding.

[8] "LangChain Docs," 5 October 2025. [Online]. Available: https://python.langchain.com/docs/integrations/retrievers/graph_rag/.

[9] "Knowledge graph," WIKIPEDIA, 2025. [Online]. Available: https://en.wikipedia.org/wiki/Knowledge_graph.

[10] 11ElevenLabs, "ElevenLabs," 11ElevenLabs, 2025. [Online]. Available: https://elevenlabs.io/.

[11] M. O. I. T. &. TELECOMMUNICATION, "Pakistan Personal Data Protection Bill," MINISTRY OF INFORMATION TECHNOLOGY & TELECOMMUNICATION, 2023. [Online]. Available: https://share.google/oUXqmt3btzePKP7hK.
[12] neo4j, "Neo4j documentation," neo4j, October 2025. [Online]. Available: https://neo4j.com/docs/.
[13] Agro, Maha & Kulkarni, Atharva & Kadaoui, Karima & Talat, Zeerak & Aldarmaki, Hanan. (2025). *Code-Switching in End-to-End Automatic Speech Recognition: A Systematic Literature Review.*
[14] Kaleem, Mohammed & O'Shea, James & Crockett, Keeley. (2014). *Development of UMAIR the Urdu Conversational Agent for Customer Service.*
[15] OpenAI. (2025). GPT-5 Technical Report. Retrieved from https://openai.com.