Software Requirements Specification (SRS)

AI Inbound Calling Agent

Project Manager: Dr. Muhammad Ilyas

Internal Advisor: Dr. Muhammad Saad Razzaq

Project Team: Muhammad Asmaad Saeed, Laiba Siraj, Fatima Zahid, Areeba Rehman

Submission Date: October 20, 2025

Project Manager's Signature: ______________________

Table of Contents

Document Information

Customer	University of Sargodha (UOS)
Project	AI Inbound Calling Agent
Document	Software Requirements Specifications
Document Version	1.0
Identifier	
Status	Draft
Author(s)	Muhammad Asmaad Saeed, Laiba Siraj, Fatima Zahid, Areeba Rehman
Approver(s)	Project Manager
Issue Date	October,2025
Document Location	
Distribution	1. Project Manager 2. Project Supervisor

Definition of Terms, Acronyms and Abbreviations

Term	Description
ASR	Automatic Speech Recognition
NLU	Natural Language Understanding
TTS	Text-to-Speech
KG	Knowledge Graph
GraphRAG	Graph-based Retrieval-Augmented Generation
OTP	One-Time Password
VoIP	Voice over Internet Protocol.
PTSN	Public Switched Telephone Network

TLS	Transport Layer Security
SRTP	Secure Real-Time Transport Protocol
GDPR	General Data Protection Regulation
PDPL	Pakistan Personal Data Protection Law
HIPAA	Health Insurance Portability and Accountability Act

1. Introduction

1.1 Purpose

This Software Requirements Specification (SRS) document defines the detailed requirements for the AI Inbound Calling Agent system. It outlines the system's purpose, features, operating environment, and design constraints to guide its development, implementation, and evaluation. The intended audience for this document includes project advisors, developers, system architects, testers, evaluators, and university stakeholders, as well as the project review and evaluation committee, who will use it as a reference for design validation, performance assessment, and future enhancements.

1.2 Project Overview

The AI Inbound Calling Agent is an intelligent, Urdu-speaking virtual assistant designed to automate inbound customer support for the University of Sargodha (UOS). It integrates Automatic Speech Recognition (ASR), Natural Language Understanding (NLU), and Text-to-Speech (TTS) technologies to manage real-time calls, understand natural language queries, and deliver accurate, context-aware responses in both Urdu, English and hybrid of both languages.[2][3]

1.3 Scope

The system's included functionalities will be:

i. Automatically answers inbound calls and interacts with callers in real time.

ii. Uses speech recognition and Natural Language Understanding (NLU) to identify intent and provide accurate responses.[2]

iii. Supports bilingual communication in both English and Urdu and hybrid of two.

iv. Retrieves essential information from the university's database, including admission schedules, departmental contacts, office timings, and event details and others.

v. Allows seamless call transfer to a human operator for complex queries.

vi. Provides an analytics dashboard to monitor call volume, response accuracy, and performance based on system logs and user feedback.

vii. Includes basic urgency detection to improve response quality and user experience.

The system's excluded functionalities will be:

I. Outbound or promotional calling.
II. Video or chat-based communication.
III. Integration with social media or non-telephony communication platforms.
IV. External Support for non-verbal communication modes (e.g., emails, or text chat).

2. Overall System Description

2.1 User Characteristics

Users of the AI Inbound Calling Agent include a diverse range of individuals and groups. Primary users are students, faculty members, and administrative staff who contact the university for information regarding admissions, schedules, departments, or general queries. Additionally, external callers, such as parents, prospective students, and the public, may use the system to obtain information or make inquiries. Administrators and technical staff have elevated privileges to monitor performance, manage system configurations, and ensure smooth operation. Stakeholders, including university management and project supervisors, oversee the system's performance, data insights, and alignment with institutional goals.

2.2 Operating Environment

The system operates in a **cloud or on-premises environment**, integrated with telephony systems using VoIP/PSTN gateways.

2.3 System Constraints

- It must comply with data protection regulations, including **GDPR, PDPL,** and **HIPAA** standards.
- The system is limited to **inbound voice-based communication** only and does not support outbound or promotional calling.
- It does not support **video or chat-based communication** interfaces.
- The system does not integrate with social media platforms or any non-telephony communication channels.
- It excludes external support for non-verbal communication modes, such as **emails** or **text-based chat systems**.
- The system requires a **stable internet connection** for real-time processing and data exchange.
- The system must support **Urdu-English code-switching** and handle **dialect variations** commonly used in Pakistan.

3. External Interface Requirements

3.1 Hardware Interfaces

- The server infrastructure must be capable of supporting simultaneous audio streaming, speech processing, and real-time communication without performance degradation.
- The system is compatible with VoIP/PSTN telephony hardware, allowing inbound calls to be received through telephone lines, mobile phones, and softphone applications[5].
- It supports use with standard microphones, headsets, and desktop or laptop computers for voice input and output.

- The **server infrastructure** (cloud-based or on-premise) must be capable of handling **simultaneous audio streaming**, **speech processing**, and **real-time communication** without performance degradation.
- No specialized or proprietary hardware components are required for system operation.
- For optimal performance, it would be preferred that the deployment server includes multi-core processors, 8 GB or more RAM, and reliable network connectivity to ensure smooth call handling and speech processing.

3.2 Software Interfaces

- It utilizes external libraries and APIs for language processing, speech recognition, and voice synthesis to ensure modularity and flexibility in development.• The system integrates with multiple software components to ensure seamless functionality and intelligent response generation.
- It connects with university databases to retrieve real-time information such as admission details, departmental contacts, and event schedules.
- The system integrates with **ASR (Automatic Speech Recognition)** modules to convert spoken input into text and with **NLU (Natural Language Understanding)** modules to interpret user intent and entities.
- It employs **TTS (Text-to-Speech)** and **voice synthesis** engines to generate natural Urdu and English voice responses.
- The system supports communication through **RESTful APIs** for data exchange between components and external services.
- It includes an urgency detection module to identify high-priority or time-sensitive queries and ensure appropriate routing or escalation.
- Integration with **Neo4j** enables graph-based knowledge retrieval through **GraphRAG reasoning,** improving contextual accuracy and response relevance.[1]

3.3 Communication Interfaces

- The system employs message queuing technologies such as RabbitMQ or Apache Kafka to manage real-time data exchange between system modules.
- All communication channels are secured using TLS (Transport Layer Security) and SRTP (Secure Real-Time Transport Protocol) to maintain data integrity, confidentiality, and protection against interception.

3. Functional Requirements

- Automatically answers inbound calls and interacts with callers in real time.
- Uses speech recognition and natural language understanding (NLU) to detect intent and respond accurately.[2]
- Supports communication in Urdu, English, or a hybrid of the two (Urdu-English code-switching).
- Retrieves and presents information from the university database, such as admissions, contacts, and office hours.[1]
- Transfers complex or unresolved queries to a human operator.
- Logs all call sessions and system interactions for monitoring and improvement.
- Provides an administrative analytics dashboard to review call metrics and system performance.

5. Non-Functional Requirements

5.1 Performance Requirements

- The system must respond to user input within 2 seconds to ensure real-time interaction.
- It should maintain an uptime of at least 99% for consistent service availability.

5.2 Safety Requirements

- The system must ensure no data loss during call handling, even in the event of a network interruption or system failure.
- All critical operations should include data backup and recovery mechanisms to prevent loss of information.

5.3 Security Requirements

- The system must implement end-to-end encryption and strict access control to protect all sensitive information.

- It must comply with GDPR, PDPL, and HIPAA standards to ensure user data privacy and legal compliance.
- The system will not use biometric or OTP-based authentication methods to maintain simplicity and avoid storing sensitive identifiers.

5.4 User Documentation

- A comprehensive User Manual, Administrator Guide, and Troubleshooting Documentation will be provided with the system deployment.
- These documents will assist end-users and technical staff in installation, operation, maintenance, and issue resolution.

6. Assumptions and Dependencies

- The system assumes continuous and stable internet connectivity for seamless operation.
- It assumes accurate performance of the ASR and NLU models for reliable speech processing.
- The system assumes the availability and proper functioning of telephony infrastructure (VoIP/PSTN gateways).
- It depends on the reliability of external cloud APIs and language model services for speech recognition, synthesis, and data retrieval.
- It also depends on regular maintenance and updates of software components and external integrations to ensure consistent performance.

7. References

[1] neo4j, "Neo4j documentation," neo4j, 2025. [Online]. Available: https://neo4j.com/docs/

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

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

[4]M. &. K. A. &. K. K. &. T. Z. &. A. H. Agro, "Code-Switching in End-to-End Automatic Speech Recognition: A Systematic Literature," 2025

[5]M. O. I. T. &. TELECOMMUNICATION, "Pakistan Personal Data Protection Bill," MINISTRY OF INFORMATION TECHNOLOGY & TELECOMMUNICATION, 2023.

[Online]. Available: https://share.google/oUXqmt3btzePKP7hK.