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: _____

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.

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.

iii. Supports bilingual communication in both English and Urdu.

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 emotion and sentiment 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

- The system requires a stable internet connection for real-time processing and communication.
- It must comply with data protection regulations, including GDPR, PDPL, and HIPAA standards.
- The system must support Urdu-English code-switching to accommodate bilingual communication.
- It should handle dialect variations and pronunciation differences commonly found across regions in Pakistan.

3. External Interface Requirements

3.1 Hardware Interfaces

- The system is compatible with VoIP/PSTN telephony hardware, including microphones, headsets, and telephony interface cards.
- The server infrastructure must be capable of supporting simultaneous audio streaming, speech processing, and real-time communication without performance degradation.

3.2 Software Interfaces

- The system integrates with internal university databases, APIs, and AI components such as ASR (Automatic Speech Recognition), NLU (Natural Language Understanding), and TTS (Text-to-Speech) through RESTful web services.
- It utilizes external libraries and APIs for language processing, speech recognition, and voice synthesis to ensure modularity and flexibility in development.

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

3.1. Automatic Call Handling
The system answers inbound calls automatically and interacts with callers in real time.

3.2. Speech Recognition & NLU
Uses speech recognition and natural language understanding to detect caller intent and provide accurate responses.

3.3. Multilingual Support
Supports communication in Urdu, English, or a hybrid (Urdu-English code-switching) for flexible conversations.

3.4. University Database Integration
Retrieves and presents information from the university database, including admissions, contacts, and office hours.

3.5. Escalation to Human Operator
Transfers complex or unresolved queries to a human operator for resolution.

3.6. Call Logging & Monitoring
Logs all call sessions and system interactions for monitoring, auditing, and improvement purposes.

3.7. Administrative Analytics Dashboard
Provides an admin dashboard to review call metrics, system performance, and user interaction insights.

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] World Bank, *Reimagining a Digital Pakistan* (2025). https://thedocs.worldbank.org

[2] Harvard Business Review, *Customer Experience in the Age of AI* (2022). https://hbr.org

[3] Neo4j Documentation (2025). https://neo4j.com/docs/

[4] LangChain Documentation (2025). https://python.langchain.com

[5] ElevenLabs (2025). https://elevenlabs.io