Structured Data

Unstructured Data

Data ETL

Intent Classification

Pre-processing

Semantics & Meaning Model

Language Model

Pronunciation Model

Speech Recognition

Security- Authentication

Entity Parsing

Knowledge Base

Voice Bot Architecture

**Quality Assurance & Enhancement**

Monitoring & Reporting

**User Voice Command**

Duration Model

Acoustic Model

Front -End

Voice Response Model

Dialog Management

Transfer to Call Centre

Logging Model

In order to address the increasing number of queries from customers more effectively and promptly for ABC Bank, we can develop an AI-based SaaS voice automation platform, to optimise voice AI solutions and automate their customer interactions via an intelligent human-like dialogue.

At a high level, a conversational bot can be divided into the bot functionality (the "brain") and a set of surrounding requirements (the "body"). The brain includes the domain-aware components, including the bot logic, user intent processing and ML capabilities. Other components are domain agnostic and address non-functional requirements such as CI/CD, quality assurance, and security.

The bot can leverages cutting edge automated speech recognition and natural language understanding technology boosted by AI-powered business algorithms to help accelerate the engagement strategy of the bank.

**Speech Recognition Engine** will be used to understand the human speech and transcribe the audio. We can use Azure LUIS, python speech recogniser, which will enable our bot to understand natural language by identifying user intents and entities.

For **Language modelling, semantics**, we can use some automation platform (say Vernacular.ai) which uses state-of-the-art, deep neural networks trained on thousands of hours of acoustic data for text to speech recognition.

In case, the voice bot is unable to service or cater to a particular customer query, it has been designed to direct the call immediately to one of the human experts “**Transfer to Call Centre model**”, minimising the navigation time on the conventional IVR.

**Acoustic model, Pronunciation model, Semantics** can be used to identify the users’ intents and important banking domain keywords. In order to implement this, we can use LM tuning, neural network and resolution metrics for word embeddings. To identify the pauses in between people communicating, we can use recurrent models which worked on real-time audios with minimal latency. Alongside, includes identifying speaker characteristics like gender, age, location, etc.

**Intent Classification** can be used to clearly understand user intent and direct it to correct response processing method. For choosing the best reply to be played back to the customer, we can use the **Knowledge Base**, **Entity parsing** that we trained to process the respective intents. We can utilize **SQL** and **NoSQL database** to capture user voice print, information and train intent classification model. Azure QnA Maker is a cloud-based API service that creates a conversational, question-and-answer layer over your data. Typically, it's loaded with semi-structured content such as FAQs. We can use it to create a knowledge base for answering natural-language questions. We can use Azure Logic Apps which is a serverless platform for building workflows that integrate user data, and calling respective services.

**Data ETL**, the bot will rely on information and knowledge extracted from the raw data by an ETL process in the backend. This data might be **structured (SQL database)**, semi-structured (CRM system, FAQs), or **unstructured** (Word documents, PDFs, web logs). An ETL subsystem will extracts the data on a fixed schedule. The content is transformed and enriched, then loaded into a knowledge data base.

**Dialog management** will take care of which language the response needs to be delivered. Azure Data Factory can be used to orchestrates and automates data movement and data transformation.

**For Security and Authentication**, we can utilize Azure AD (Azure Active Directory) using which users will authenticate through an identity provider such as the bot service handles the authentication flow and OAuth token management. Azure Key Vault will store credentials and other secrets using Key Vault.

**Logging and monitoring model** will be used to capture the metrics to improve the voice bot performance and help resolve technical challenges witnessed. Azure Application Insights can be utilized to log the bot's application metrics for monitoring, diagnostic, and analytical purposes. Power BI can be used to create monitoring dashboards for our voice bot.

**Quality assurance and enhancement model** will be addressing the performance improvement aspects of the voice bot. We can employee Azure DevOps which provides many services for app management, including source control, building, testing, deployment, and project tracking.

The reply is converted to audio which can be played to the customer for an end-to-end conversational feel.