

AI Based Conversational IVR Modernization Framework

Abstract:

Traditional Interactive Voice Response systems, as implemented in large-scale public service systems, use traditional static menus based on DTMF input from keypresses. Traditional IVR systems result in longer call times, a worse caller experience, and inflexibility, especially for non-technical callers. The current project promotes the concept of Conversational IVR Modernization Framework, where conversational Artificial Intelligence is used to enhance the traditional IVR systems in such large-scale public service systems by providing additional voice interaction capabilities.

By taking IRCTC as a reference system, the proposed system indicates that the existing VoiceXML-based IVR system can be upgraded without replacing the traditional system completely. A middleware integration layer is proposed between the traditional IVR system and the AI-driven conversational interface. The proposed system receives the voice inputs from the user, interprets the voice as text using AI-based functionality, and uses the simulated data to provide the output using voice synthesis.

Moreover, the proposed approach is more user-friendly and reduces navigation complexities while responding with enhanced efficiency. Also, it is expandable to different domains such as banking, telecom, or government helpline services, which makes it an effective solution for voice IVR transformation.

Introduction:

Many organizations use IVR systems to handle customer calls automatically. These are systems that allow the user to get desired information by just calling and selecting options without talking to a human operator. IVR systems find their applications in areas such as transportation, banking, telecommunication, and public services.

Most traditional IVR systems work using fixed menus. Users have to listen to voice instructions and press keypad numbers such as "Press 1 for PNR status" or "Press 2 for train timings." Although this method works, it often becomes confusing and time-consuming. Users may have to go through many options before getting to the required information.

On large public platforms like IRCTC, a huge number of users depend on IVR systems for train-related enquiries. Because of long menus, unclear instructions, language difficulties, and repeated inputs, users often feel frustrated. This results in longer call times and many users disconnecting the call before getting the required information.

With recent developments in Artificial Intelligence and speech technologies, IVR systems can now be made more user-friendly. Modern conversational IVR systems allow users to speak naturally and directly ask questions such as "Check my PNR status." The system understands the user's voice and provides the required information without using keypad-based navigation.

The main aim of this project is not to replace the existing IRCTC IVR system, but to improve it by adding an AI-based conversational layer. This is done using a middleware approach that connects the old IVR system with modern AI services. This method allows intelligent voice interaction while keeping the existing system unchanged.

By taking IRCTC as a case study, this project shows how traditional IVR systems can be upgraded into modern conversational IVR systems. The proposed approach improves user experience, reduces call complexity, and provides a scalable solution for future IVR modernization.

Problem Statement:

This project aims to modernize existing IVR (Interactive Voice Response) systems built on VoiceXML (VXML) by integrating them with modern Conversational AI platforms such as ACS and BAP Service. By reusing and extending legacy assets, the solution will enable these systems to support conversational interfaces while minimizing rework. The approach focuses on enhancing usability, improving user experience, and reducing the technical burden of transitioning legacy IVRs to AI-driven workflows.

Objective of the Project:

- To examine the current IVR system in use for various sites, including IRCTC.
- To develop a conversational IVR framework that enables users to converse through natural voice commands rather than navigating through the keypad.
- To integrate conversational AI with legacy IVR systems with a middleware-based approach to the existing infrastructure.
- To convert user voice input to text and to identify user intent using AI technologies.
- To fetch the necessary information based upon the query given by the end user from a data source.
- In order to provide voice-based responses to users in a simple and understandable manner.
- For improved user experience by reducing navigation complexities and call response time.
- To show how the scalable and reusable modernization model of IVR can be applied to other domains, i.e., banking, telecom, and government services

Traditional System – Legacy IVR:

The available IVR system used on many platforms is the traditional menu-based operation. It has fixed flow of instruction that users need to follow, and using their phone keypads, they have to respond.

How the Old IVR Works:

User makes a call

- The user calls the customer care/enquiry number provided.
- Pre-recorded Voice Message plays

The system plays recorded messages such as:

- “Press 1 for PNR status”
- “Press 2 for train timings”

User presses keypad numbers (DTMF input)

- In respond, the user simply presses the phone keypad buttons.
- Signalling from a keypad consists of signals known as DTMF signals (Dual Tone Multi Frequency).

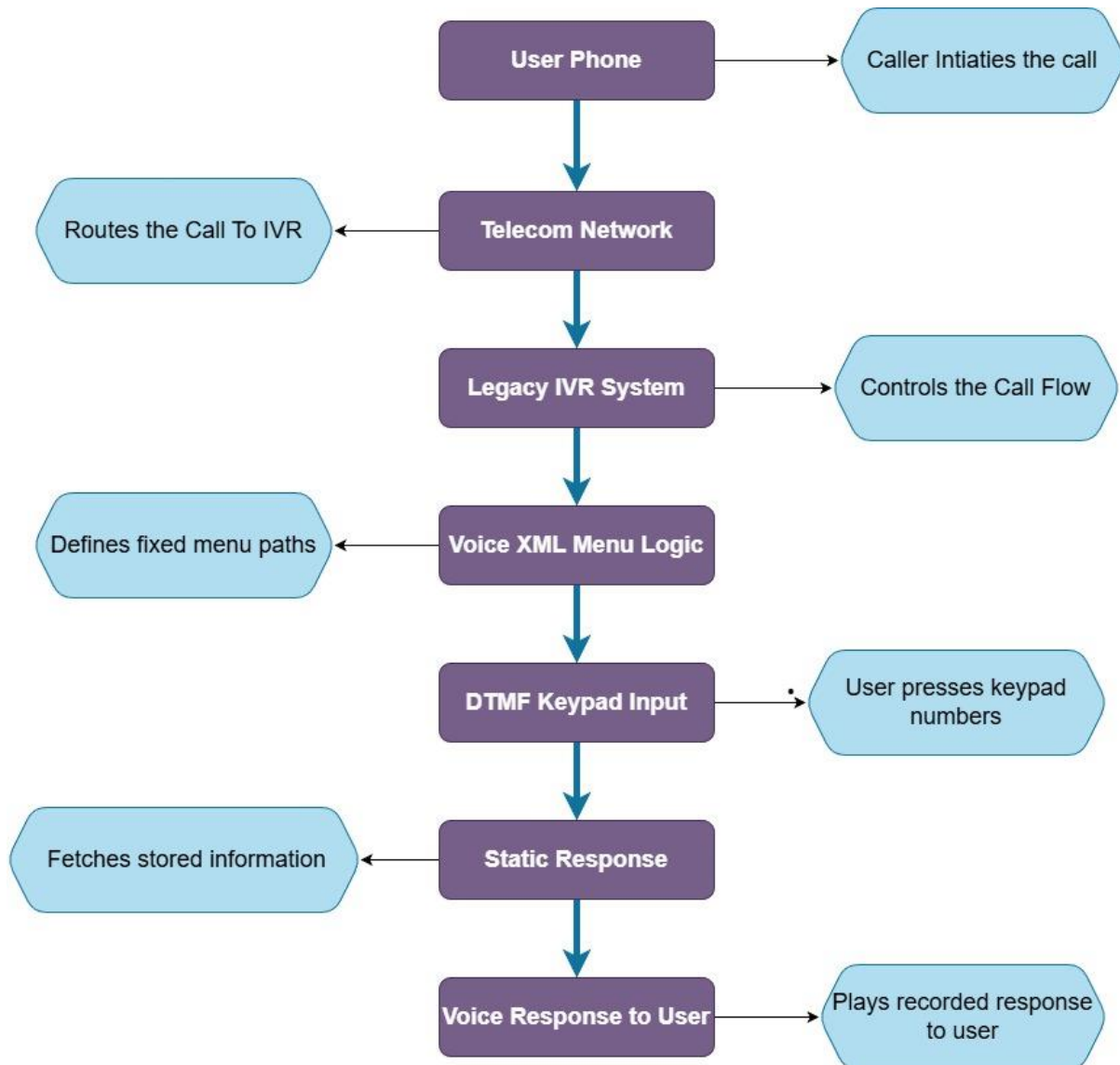
VoiceXML-Based Flow Control

- The entire call flow is implemented in VoiceXML, where VoiceXML is:
- What message to play What's the next step? What happens when a key is pressed

Rigid menu structure

- They must follow the same options in the same sequence. There are no opportunities to bypass steps or speak normally.

Legacy IVR Architecture



Proposed System – Modern Conversational IVR

The proposed system offers Modern Conversational IVR, where users can converse with the system using natural voice commands instead of dialing keypad numbers. The proposed system improves user experience without changing the legacy IVR system.

Instead of being forced to speak options such as “Press 1,” “Press 2,” users can state their requirement. The system will listen to their voice, understand their requirement through Artificial Intelligence, retrieve the required information, and communicate back to them through voice.

This approach does not replace the current IVR system; rather, an integration layer (middleware) is created, bringing together traditional IVR and modern conversational AI services.

How it Works:

User initiates a call

- The customer dials the IVR number.

Voice input captured

- Instead of keypad input, the user speaks naturally, for example:
 - “Check my PNR status”

Speech-to-Text

- Speech recognition converts the spoken voice into text.

AI understands user intent.

- Conversational AI analyzes the text for what the user is requesting.

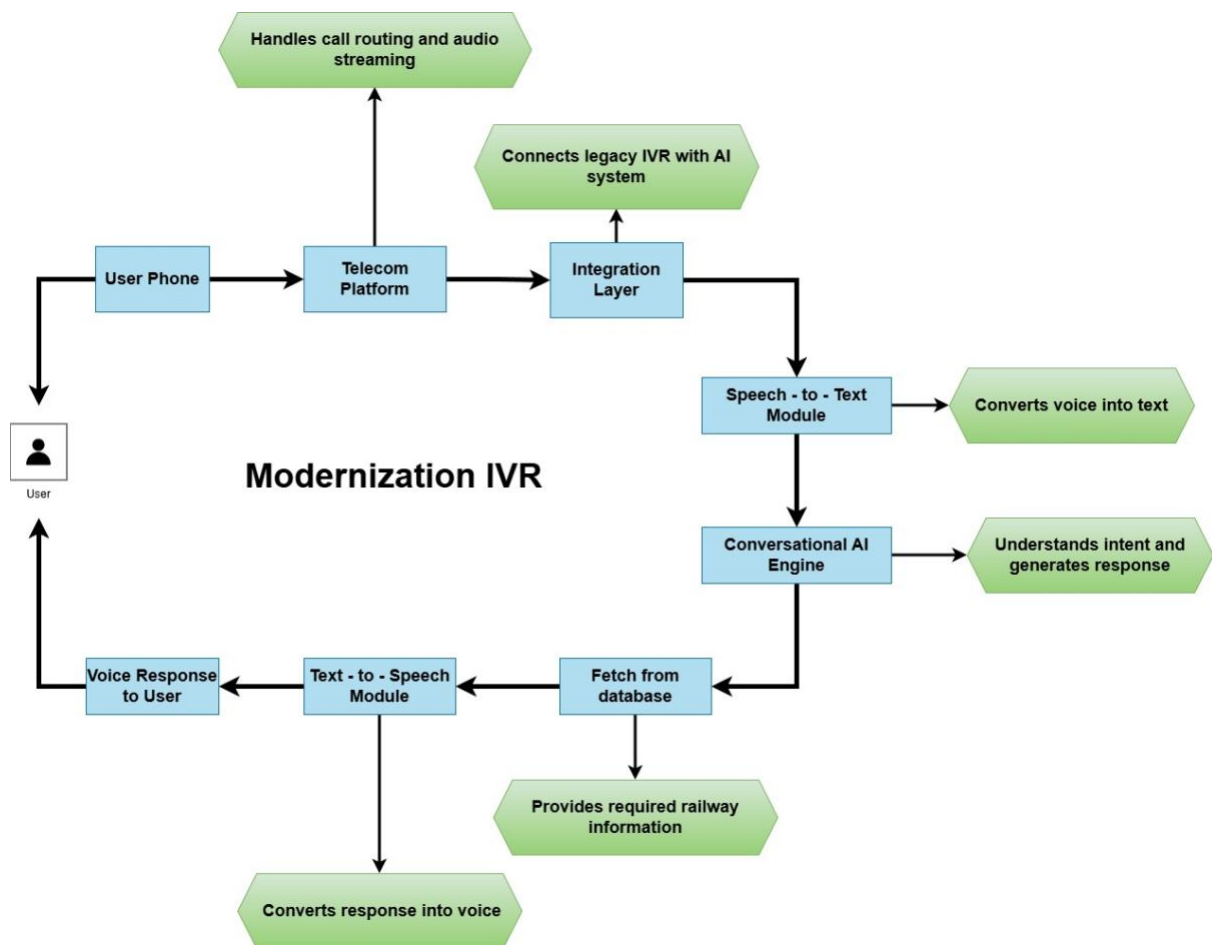
Required data is fetched.

- The system fetches only the relevant information from the database or simulated IRCTC data.

This response is then converted into voice.

- The voice is played back to the user, and response text is converted to voice.

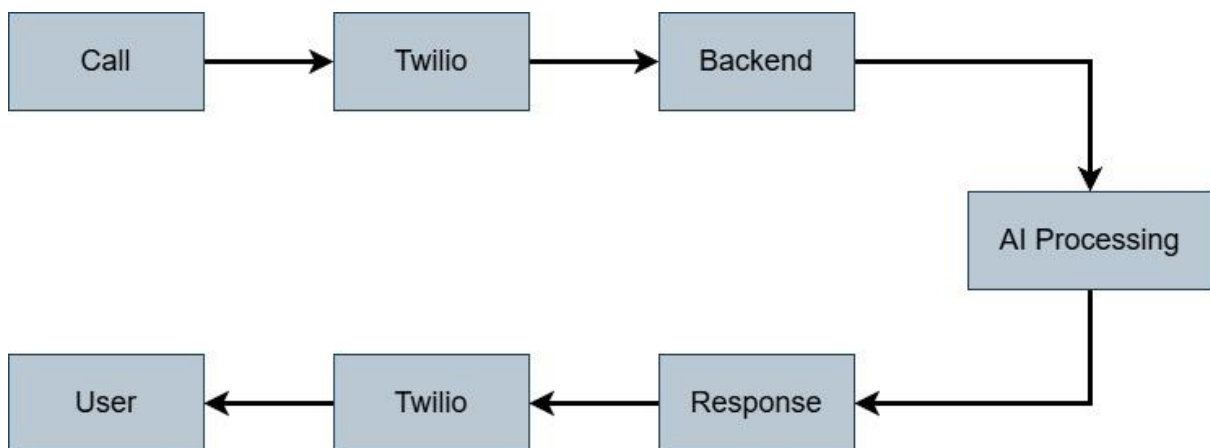
Modernization IVR Architecture:



Technology Stack:

Step	Technology
Communication	Twilio / ACS / BAP
Backend	Python, Flask
AI	OpenAI API
Speech – to – Text	Whisper
Data Storage	Json, other DataBases
Platform	IRCTC

Twilio Overview:



1. User dials a number
2. Call is received by Twilio Cloud
3. Backend provides instructions to Twilio
4. User gives voice input
5. Twilio records and streams audio data
6. Backend processes the request
7. Backend sends response to Twilio
8. Twilio does text-to-voice conversion.
9. Call ends or continues

Technical Challenges:

- **Speech Recognition Accuracy**

There is a possibility of incorrect recognition of the user's speech.
- **Latency in Response Time**

There may be minor delays involved in the processing of AI and API calls.
- **Dependency on Third-Party Services**

The system will depend on Twilio, as well as other AI APIs.
- **Handling Multiple Users**

Handling simultaneous calls in an efficient manner is technologically challenging.
- **Data Privacy and Security**

The user information should be processed and stored properly to avoid misuse.

Constraints:

- **Internet Dependency** – A stable internet is required to run this system.
- **API Dependency** – Here, the system uses external APIs, e.g., Twilio and AI APIs.
- **Simulated IRCTC Data** – This project uses a simulated data approach, not actual integration with IRCTC data.
- **Limited Real-World Deployment** – The system is a demonstrator created for academic purposes and not intended for real railway environments.
- **Cost Constraints** – There may be operation costs involved in continuously using APIs.

Compatibility Gaps:

- **Legacy VoiceXML Integration Gap** – Legacy IVR systems are not originally developed with native support for AI-based conversational flow.
- **Protocol Differences** – There may be variations in the communication protocol used in legacy IVR systems and AI systems.
- **Speech Model Variability** – AI systems may vary in different settings/languages.
- **Hardware Compatibility** – Telecom infrastructure may not be compatible with advanced conversational features.
- **Scalability Differences** – Legacy systems do not offer high concurrent processing of data for Artificial Intelligence.

Conclusion:

This project aims at enhancing conventional IVR systems by integrating modern AI technology. Users can communicate with the system by using their voices, which can retrieve information. A middleware component is used for integrating the traditional IVR system with AI technology. The proposed system can be extended by using it as a model with simulated data, as mentioned in the project. It can be used in the future with other systems.