

Voice Agent Contact Centres Explained

Seamless Support for Modern Businesses

These days, when you reach out to your telecom service provider, bank, or e-tailer customer care over a phone call, you are often connected to an AI voice agent that listens to you and responds to your queries or complaints. Businesses are increasingly relying on AI voice agent contact centres to provide 24x7 support to their customers while reducing overall support costs. AI voice agents provide a natural human-like interaction, aiming to improve your contact centre experience.

While contact centres usually have both human and AI voice agents, purely AI-enabled contact centres with only AI voice agents are also possible. Let's explore a contact centre with only AI voice agents, enabled by the AI capabilities of OpenAI's large language models (LLM) and API.

How an AI Voice Agent Contact Centre Works

Let's see how such a contact centre works, using an example of your interaction with a telecom company XCom's contact centre. Let's assume the contact centre is connected to the XCom Knowledge Base (KB) as primary information source and to OpenAI as a fallback.

Let's consider the following cases:

- a) You want to know whether your plan includes a Netflix subscription.
- b) You want to complain about not being able to send Whatsapp messages.

This is how your interaction with the contact centre takes place, in simpler terms:

1. When you call the contact centre, a voice agent prompts you to share your query or complaint. When you speak about your query or complaint, your speech is converted to text by an Automatic Speech Recognition (ASR) engine (for example, Twilio Speech Recognition or OpenAI's Whisper).
2. Then the text along with your customer information is passed to a Retrieval Engine. Retrieval Engine searches for information relevant to the text in XCom's KB using [semantic search](#). The KB contains detailed information about the telecom plans and bundles the company sells, updates of its latest system maintenance, frequent issues and their resolution.
 - i) In the [case a\)](#), the Retrieval Engine finds the relevant information in the KB regarding your plan and generates a text response to your Netflix subscription query using [Retrieval-Augmented-Generation \(RAG\)](#).

- ii) In the [case b](#)), the Retrieval Engine does not find the relevant information in the KB, and the text is prepared to be passed to the fallback - the OpenAI LLM (for example, GPT-4).

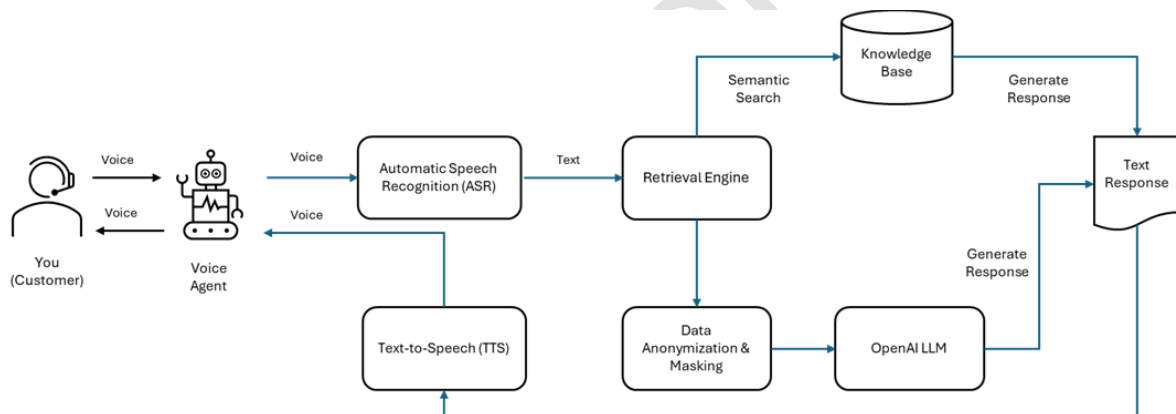
The text is prepared by either removing or anonymizing your and XCom's information (for example, your address, customer ID, or billing information) to insure personal and company data safety. Your customer information too is not passed.

Then OpenAI LLM is prompted with the prepared text as context to generate a text response for your Whatsapp-related complaint. The response is a request to check your device settings and to refer to the Whatsapp documentation.

3. In both cases, the generated text response is passed to a Text-to-Speech (TTS) converter (for example, Amazon Polly or Azure AI Speech) to generate a natural voice response.
4. The voice agent then plays back the response to you.

Here is an illustration of the interaction.

Figure 1 Interaction with AI Voice Agent Contact Centre



Under the Hood

The [previous section](#) presents a simplified version of an AI Voice Agent Contact Centre's working. Actual contact centre interactions involve more components, steps, and human agents as well. Some of the major components are:

- Automatic Call Distributor (ACD) directs your call to human or AI voice agents based on your voice input and customer information. It is a crucial component that identifies cases that cannot be handled by AI (for example, escalations or SIM loss complaint) and directs them to human agents.
- Orchestration Layer (for example, Twilio Functions) coordinates the steps of the interaction.

- Logging, Analytics, and Monitoring components logs transcripts, prompts sent to the LLM, customer voice recordings, and resolution outcomes for QA and analytics. It provides analytics dashboards to monitor key metrics, like call volumes, response times, customer sentiment, and more.
- Security and Compliance components such as firewalls, identity management, and encryption ensure end-to-end security of customer speech input and data along with prompts and generated responses.

These components make the Voice Agent Contact Centres, whether hybrid or pure AI, robust, efficient, and secure.

Reference

- OpenAI Whisper [ASR](#)
- Microsoft Azure [AI Speech](#)
- Twilio's blog on [voice and ChatGPT integration](#)
- Workato's [voice-bot architecture document](#)
- Teneo.ai [RAG approach for voice IVR use cases](#)
- Strac.io [best practices for data privacy](#)