



APPLICATION NOTE

AudioCodes VoiceAI Connect

Easily connect your AI solution to any voice or telephony system

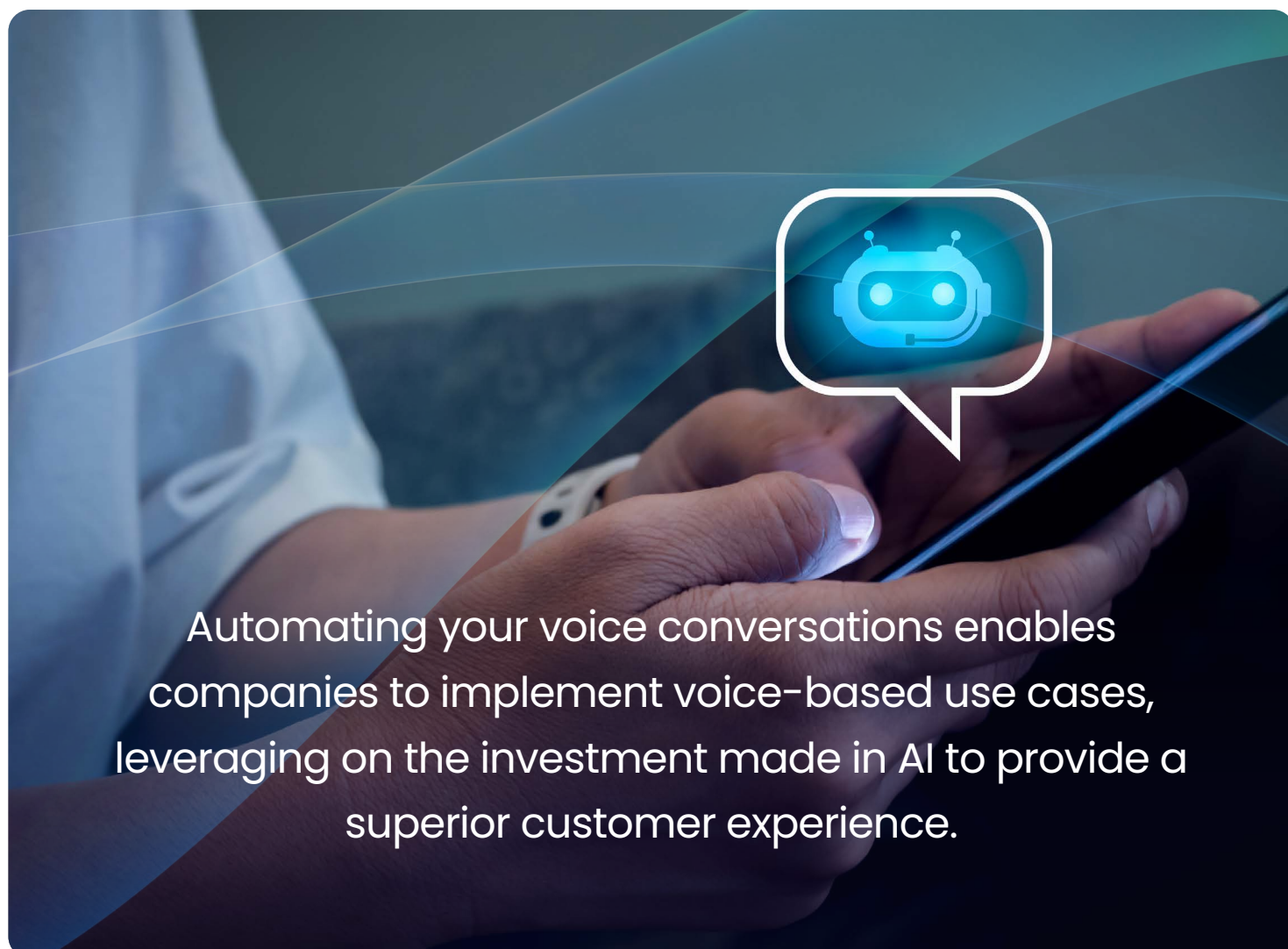


Overview

In the last few years, dramatic leaps forward in machine learning and AI have driven a revolution in the way enterprises boost engagement with their customers. These incredible advances mean that businesses can now utilize conversational AI technologies offered by various providers to automate their customer service departments and train bots to give callers a high level of service whenever they get in touch.

According to Cavell's Q2 2024 CCaaS Market Evolution Report, approximately 90% of companies are either already implementing AI in their contact centers or plan to do so in the coming years. Notably, some industry players anticipate that by 2028, automated systems will resolve 60-80% of customer support tickets, leaving live agents to handle only 20-40%. This reflects a significant increase in automation driven by AI advancements.

Most current bot implementations are text-based. Automating your voice conversations enables companies to implement voice-based use cases, leveraging on the investment made in AI to provide a superior customer experience. They can use voice AI agents to communicate with customers, replace their legacy IVR system with conversational IVR and use bots as live agent assistants to optimize agent productivity, resulting in exceptional customer satisfaction and a significant drop in customer support costs.



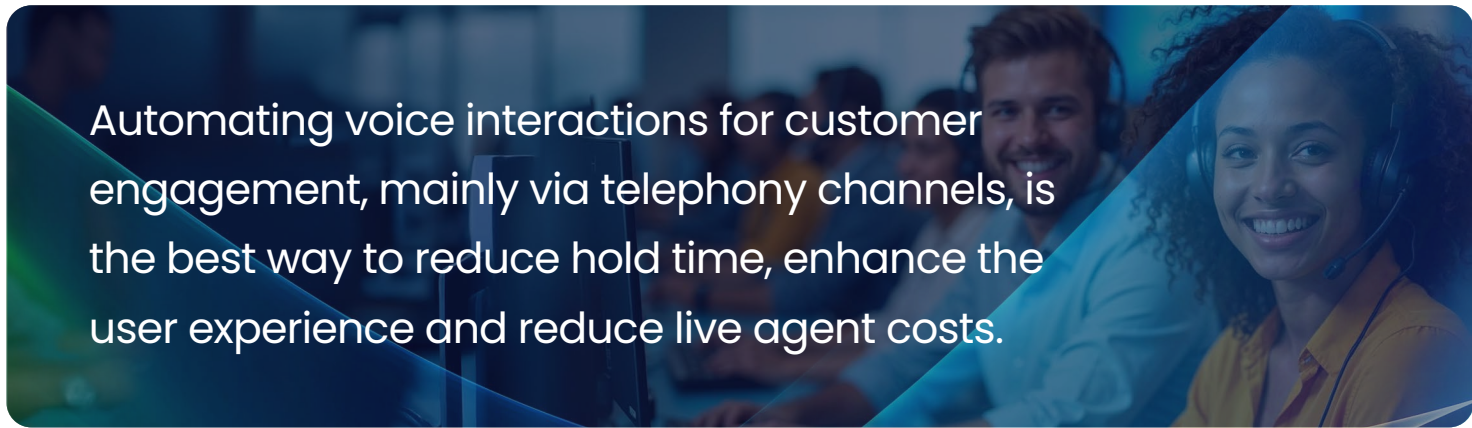
Automating your voice conversations enables companies to implement voice-based use cases, leveraging on the investment made in AI to provide a superior customer experience.

The Importance of Voice and Telephony Channels

Voice is – and probably always will be – the most intuitive form of human communication. The rise and continued success of voice streaming offerings from generative AI providers such as OpenAI, Claude and others are examples of how voice communication is also successful in human-machine interactions. There are multiple ways to perform voice interactions, and the most common one – available everywhere – is based on telephony with mobile and landline phones.

Large volumes of customer contact center interactions are performed using voice and telephony. This high load on the contact center leads to longer customer hold times, and therefore customer dissatisfaction, as well as increasing costs for adding more and more live agents to handle the ever-growing volume.

Automating voice interactions for customer engagement, mainly via telephony channels, is the best way to reduce hold time, enhance the user experience and reduce live agent costs.



Automating voice interactions for customer engagement, mainly via telephony channels, is the best way to reduce hold time, enhance the user experience and reduce live agent costs.

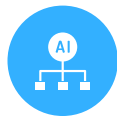
Voice Conversational AI Use Cases

There are three main use cases for voice-bots:



Voice AI Agents

Offload simple and repetitive tasks from live agents to voice-bots. It allows live agents to concentrate on more complex customer interactions, reduces hold time and improves the customer experience without increasing the number of agents. If the interaction cannot be completed by the AI agent, the call will be transferred smoothly, along with the relevant details, to a live agent for completion. This solution is characterized by high scalability, a high return on investment and an exceptional user experience.



Conversational IVR

Legacy IVR allows humans to interact with a computer-operated phone system using DTMF tones inputted via a keypad. It is usually a long and tedious process, which has a detrimental effect on the customers' user experience. Conversational AI-based IVR uses natural language understanding to replace hierarchical menus with a free speech experience. It ascertains the customers' needs and instantly routes the call to a virtual or live agent. The solution saves time, is more accurate and improves the user experience.



Live Agent Assistants

The bot assists live contact center agents by listening unobtrusively to conversations between customers and agents, analyzing the data and sending real-time insights and after-call summaries to the agents or their supervisors. The insights can be utilized to guide the live agents through the process of handling customer enquiries and suggesting relevant answers, while the after-call summaries can help the agent stay focused on the user's issues and prepare for the next call. This optimizes agent productivity and enhances the user experience.



Real-Time Voice Translation

Enables seamless multilingual voice communication by instantly translating voice conversations with high accuracy and natural-sounding speech. This service allows businesses to break language barriers in customer interactions and global operations. By harnessing advanced speech recognition, machine translation and text-to-speech technologies, real-time voice translation ensures a smooth and efficient conversational experience. This solution empowers contact centers, enterprises and service providers to enhance customer engagement, improve accessibility and expand their reach in a multilingual world.

Voice-Bot and Telephony Connectivity Challenges

Most conversational AI platforms offer connectivity to various text engagement channels. But although some platforms offer voice APIs, only a few of them supply native telephony connectivity – and those that do only support connections to their own platform, resulting in vendor lock-in. Therefore, enterprises and other solution providers looking to implement voice-bots or other voice conversational AI use cases with telephony connectivity will likely face the following challenges:

- **Telephony System Integration:** Telephony consists of multiple VoIP protocols and multiple voice codecs that are based on international telephony standards. Each VoIP vendor implements its own flavor of these protocols or codecs, requiring a sufficient level of expertise in them to enable successful integration with each new platform. This is completely different from communications with cognitive services such as speech-to-text, text-to-speech and bot frameworks, since these are based on proprietary HTTPs and are widely used for internet/web communications. This creates a challenge for developers when communicating with the telephony world.

Furthermore, there are multiple types of telephony channels, ranging from public telephony providers through to contact centers and enterprise communication platforms, each with its own uniqueness to address during integration. Telephony IT teams will usually be reluctant to make any changes in their telephony systems, preferring the adaptations to be made on the bot side.

Last but not least, the connectivity with the various telephony channels should be secured, highly available and scalable, matching the standards in the telephony space.

Most bot frameworks offer connectivity to various text engagement channels. But although some offer voice APIs, none of them supply native telephony connectivity.

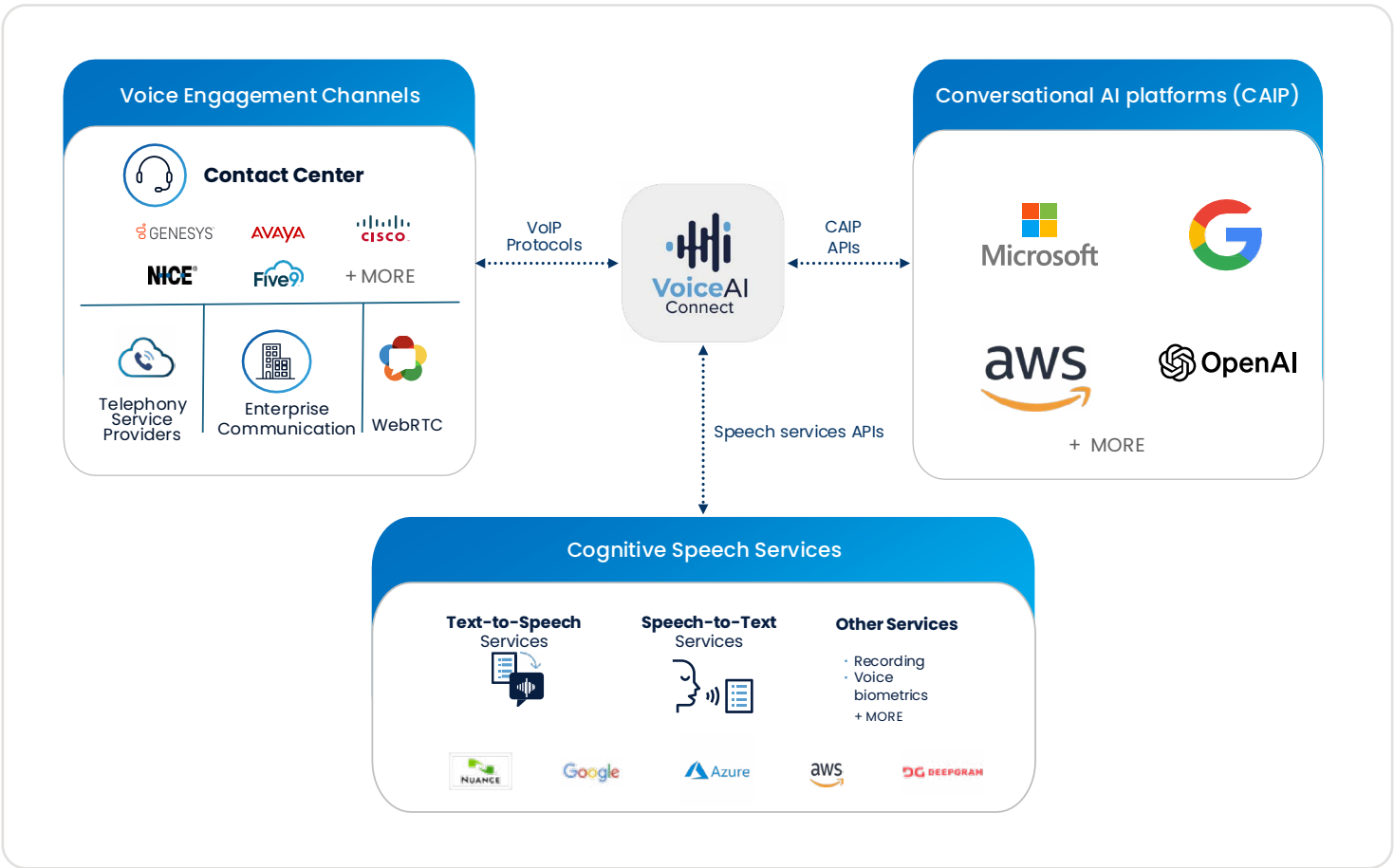
- **Voice Quality and Voice Latency:** Two basic factors of the voice-bot user experience are related to the accuracy of the speech detection and voice latency. This challenge is more complex for telephony connections that rely on low band voice codecs, traverse multiple hops (some of them over the internet) and those that come from noisy environments. This can lead to bad voice quality that affects speech detection accuracy. As a result, the bot's inability to understand the user and respond appropriately can cause perceptible and sometimes unacceptable delays in the human-bot conversation.
- **Multiple Speech Providers and Conversational AI Platforms:** Basic building blocks involved in implementing a voice-bot are voice-to-text, text-to-voice and connectivity to various conversational AI platforms (including LLMs) and middleware. There are multiple providers in this domain, and it is vital to choose the best provider for each of the functions while keeping the costs down as much as possible. Factors such as languages and dialects, speech detection training and optimizing, connecting to home-grown solutions and cloud or on-premises deployment should be considered.
- **Adapting the User Experience to Voice:** The overall user experience is influenced by the bot's capabilities, and the possible interaction the solution enables with the telephony world. There are three main challenges affecting the UX when connecting bots to telephony:
 - Using events and data that come from telephony such as the calling number, caller name, any customer data that was gathered by a contact center in previous interactions, dialed digits (DTMF) and more.
 - Initiating activities that are relevant for telephony by the bot such as escalation to a human agent, adding metadata to the contact center session context and recording the call.
 - Adjusting the bot user experience to the voice medium. On one hand, eliminating inputs that are not relevant to voice, like visual menus, emojis and images; and on the other hand, adjusting the bot to handle situations that do not happen in chat interaction, such as caller interruptions and no user input.
- **The Bot Developer Control:** Bot developers are accustomed to controlling all the application aspects using no-code/low-code coding, flow editors or simple HTTP-based APIs. Therefore, it is advantageous for bot developers to use a solution that provides similar APIs when integrating with the telephony world.
- **Diverse Contact Center Environments:** Many organizations operate on-premises contact centers with future plans to migrate to the cloud, while others manage multi-vendor environments due to acquisitions or global operations. These organizations can benefit from adopting a solution that decouples telephony channels and voice AI capabilities from the underlying contact centers. This abstraction allows developers to create voice AI applications once and then deploy them seamlessly across all contact center platforms, ensuring consistency, scalability and faster innovation.

AudioCodes VoiceAI Connect

AudioCodes, a leading provider of advanced communications software for the digital workplace, directly addresses this key market need with its innovative [VoiceAI Connect](#) solution.

Utilizing AudioCodes’ vast voice expertise, VoiceAI Connect creates a bridge between any conversational AI platform and any telephony system based on a best-of-breed approach that allows you to pick the best provider for each of the voice cognitive services.

Utilizing AudioCodes’ vast voice expertise, VoiceAI Connect creates a bridge between any conversational AI platform and any telephony system based on a best-of-breed approach that allows you to pick the best provider for each of the voice cognitive services.



As illustrated in the diagram above, VoiceAI Connect acts as a bridge between the different telephony channels and cognitive services, enabling secured, highly available and scalable architecture on both sides of the solution. VoiceAI Connect contains two main components:

- **Telephony Engagement Controller:** Based on AudioCodes' market-leading enterprise session border controller (SBC) architecture, the Telephony Engagement Controller integrates with the telephony channels. In this way, VoiceAI Connect can integrate with any telephony channel, public telephony provider, contact center or enterprise communication platform, as well as any platform communicating via WebRTC.
- **Cognitive Services Orchestrator:** This component interacts with the various cloud services and implements the best-of-breed approach by adapting to any cognitive service and orchestrating speech-to-text, text-to-speech and conversational AI platform voice flow. It can connect to various conversational AI platforms such as Microsoft Copilot Studio, Google Dialogflow, Rasa and Amazon Lex. It also exposes public APIs that can be used to connect with any other bot framework or middleware. Many conversational AI platforms have already used these APIs to integrate the solution into their platforms.

VoiceAI Connect Features

VoiceAI Connect's unique architecture enables an improved user experience and provides superior tools and features for the bot developer to use when developing and implementing a voice conversational AI solution:



Smooth integration with various contact centers, supporting advanced use cases such as escalation to human agents, handing metadata to/from the bot, Barge-In, compliance recording and more.



Methods for the bot to **control telephony actions**, such as basic call initiation to/from the bot, playing prompts and ring back tones, call recording and call disconnection.



Send the **bot notifications** arriving from the telephony channel, such as DTMF, answering machine and FAX detection tones as well as no user input.



Dynamic control on speech engine configuration, such as provider, SSML, language, voice name, speech model and contexts, as well as no-user-input notifications and actions.



Mechanisms for cost reduction on STT/TTS consumption, where VoiceAI Connect caches text-to-speech prompts and suppresses silence to avoid excessive costs.

VoiceAI Connect's unique architecture enables an improved user experience and provides superior tools and features for the bot developer.

VoiceAI Connect Benefits

On top of the fundamental capabilities of connectivity to any telephony channel, bot framework and speech services, the solution has additional important benefits that improve any voice-bot solution:

- **Smooth Integration:** The solution easily connects any telephony system to any bot framework. Because a single vendor performs the telephony and cognitive services integrations, successful implementation and optimal performance is much more likely.
- **Quality of Experience:** VoiceAI Connect is the only solution today that relies on SBC core architecture with direct connectivity from the telephony systems to the speech engines and bot frameworks. This minimizes the number of media hops and delivers excellent voice quality and minimal voice latency, as well as robust security.
- **Best-of-Breed Approach:** VoiceAI Connect has a list of certified speech engine providers (text-to-speech and speech-to-text) that bot developers can choose from to meet their specific needs. The solution also provides public APIs that enable it to be integrated with any other speech engine.
- **Reduce Traffic to Speech Services:** Based on the SBC's ability to detect silence and voice activity, and stop and start speech-to-text detection accordingly, VoiceAI Connect can reduce the consumption and cost of the speech-to-text service by up to 40%. For text-to-speech, the solution implements a caching mechanism that reduces both consumption and the cost of this engine and also minimizes voice latency.
- **Multiple Deployment Options:** VoiceAI Connect is offered in two variations to suit any deployment option and to provide the highest possible flexibility for the bot developer.

Deployment Options

VoiceAI Connect has two service deployment options to meet the different needs companies have when enabling voice conversational AI.



Live Hub is a SaaS self-service portal for easily extending bot usability to telephony use cases. It enables the smooth implementation of virtual agent, conversational/smart IVR and live agent assist applications.

Live Hub connects a wide range of bot frameworks to your contact center business flows, public telephony and cognitive speech services. It offers quick and simple self-service onboarding, using a simple wizard, so you can call your bot in just a few clicks. You can bring your own carrier (BYOC) or purchase phone numbers from AudioCodes for use worldwide.

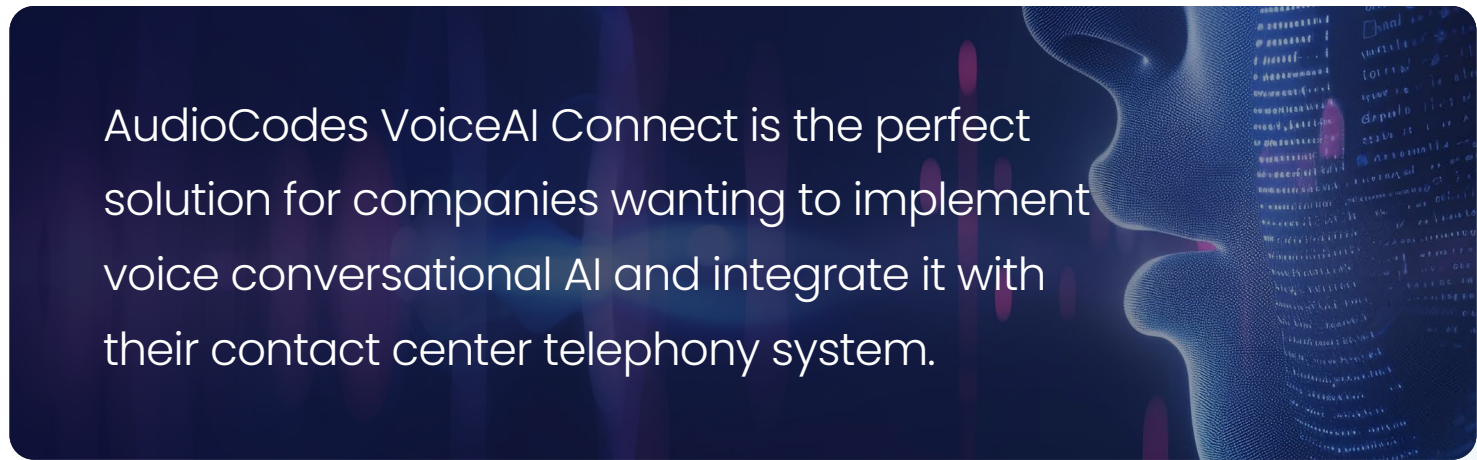
The VoiceAI Connect Enterprise edition is a dedicated setup that can be deployed on the customer's private data center or public cloud account (Azure, AWS or GCP). It is also offered as a dedicated setup hosted on the AudioCodes Cloud.

This type of deployment is designed for organizations with strict privacy and security policies that require a dedicated setup.

Now's the Time to Give Voice to Your Bot

For more than three decades, AudioCodes has specialized in connecting voice solutions and networks. VoiceAI Connect leverages these capabilities and extends them to connect to any voice AI service, allowing bot developers using any conversational AI platform to build voice AI applications and connect them to any voice network with simple APIs.

VoiceAI Connect is the perfect solution for companies wanting to implement voice conversational AI and integrate it with their contact center telephony system. It enables immediate deployment with your chosen voice AI providers and preserves your development efforts and costs during telephony provider migrations.

A banner image with a dark blue background. On the right side, there is a profile of a human head in a light blue, wireframe-like style. The background is filled with faint, glowing blue and purple light effects. The text is white and positioned on the left side of the banner.

AudioCodes VoiceAI Connect is the perfect solution for companies wanting to implement voice conversational AI and integrate it with their contact center telephony system.

CONTACT US

International Headquarters
Naimi Park, 6 Ofra Haza Street,
Or Yehuda, Israel
Tel: +972-3-976-4000

AudioCodes Inc.
80 Kingsbridge Road
Piscataway, NJ 08854
Tel: +1-732-469-0880
Fax: +1-732-469-2298

www.audiocodes.com/contact
www.audiocodes.com

©2025 AudioCodes Ltd. All rights reserved. AudioCodes, AC, HD VoIP, HD VoIP Sounds Better, IPmedia, Mediant, MediaPack, What's Inside Matters, OSN, SmartTAP, User Management Pack, VMAS, VolPerfect, VolPerfectHD, Your Gateway To VoIP, 3GX, AudioCodes One Voice, AudioCodes Meeting Insights, and AudioCodes Room Experience are trademarks or registered trademarks of AudioCodes Limited. All other products or trademarks are property of their respective owners. Product specifications are subject to change without notice.