



GUIDE

Adding voice AI to your call center tech stack



Contents

- Overview 3
- Applications..... 4
- Technology5
- Deployment 9
- Use cases11
- Benefits.....12
- Get started15





What is call center voice AI?

Call center voice AI is revolutionizing the contact center experience and communication between customers and machines.

Using a unique stack of technologies, these AI systems are designed to handle and respond to interactions in the call center. These systems leverage a combination of natural language processing (NLP) and speech recognition to understand and interpret spoken language, allowing customers to resolve inquiries through natural conversation.

Call centers can implement voice AI to automate and streamline customer interactions to create a more cost-effective and efficient way to handle routine queries, requests, and tasks without the need for human intervention.





Call center voice AI applications

There are three main ways that contact centers can apply AI to voice calls.



Voicebots

Voicebots are automated systems that use AI technology to interact with callers using spoken language. These bots can handle various tasks, such as answering FAQs, taking payments, routing callers, managing orders, and guiding customers through authentication processes — all through spoken conversation

Previous generations of voicebots have been technology-led. They’ve required callers to change their behavior to suit the limitations of technology.

However, a new generation of voice assistants are offering customer-led experiences. That means customers can say whatever they want, however they want to say it.



Agent assist

Agent assist uses AI to help customer service representatives during live calls. Agent assist tools analyze dialogue in real-time, understanding customer needs and suggesting relevant information.

By plugging into companies’ knowledge bases, agent assist tools provide agents with quick answers to common queries and customer information to offer more personalized experiences.



Speech analytics

Speech analytics software analyzes and interprets spoken language in audio recordings of customer interactions within a call center. This software uses NLP techniques to extract valuable insights, patterns, and information from conversations to create more efficient operations, improve customer satisfaction, and enhance agent productivity.



Elements of a call center voice AI tech stack

To deliver great customer experiences over the phone, call center voice AI solutions must be successful at listening, understanding, and responding to what the customer is saying.

To deliver in these key areas, the following technical capabilities are required.



Listening technologies for call center voice AI



Automatic speech recognition (ASR)

ASR uses machine learning to convert spoken language into written text. Advances in deep learning, particularly using neural networks, have significantly improved the accuracy and performance of ASR systems.



Spoken language understanding (SLU)

While ASR models have vastly improved over the last decade, they are still often insufficient in accurately transcribing spoken language over the phone due to patchy connections and differing accents and pronunciations.

SLU techniques and technologies can be used to fine-tune ASR systems to allow enterprises to overcome the challenges of the voice channel, including background noise, poor call quality, and various dialects and accents, and effectively automate customer interactions over the phone.



Alphanumeric capture

Alphanumeric capture is the process by which call center voice AI tools extract and understand both letters and numbers from spoken language during a conversation between a person and an automated voice system.

This could include account numbers, zip codes, or order IDs that enable the voice assistant to identify customers and complete transactions.

Previous applications of call center voice AI have relied on keypad input for alphanumeric capture due to limitations in speech recognition technology.

With sufficient SLU, accurate alphanumeric capture can be achieved through spoken language.



Phoneme matching

Phonemes are the smallest units of sound in a language that can change the meaning of a word. Phoneme matching is used to identify and understand the specific sounds produced by a person during a conversation and allows call center voice AI to identify callers by name by extracting information at a phonetic level.



Understanding technologies for call center voice AI



Natural language understanding (NLU)

Natural Language Understanding (NLU) is the process by which technology is able to understand natural human language.



Large language models (LLMs)

LLMs are used to deliver high-performance NLU. LLMs have been trained on vast language datasets and can accurately understand the meaning and relationships behind different words and phrases. They are particularly useful in spoken conversations where synonyms, idioms, slang and turns-of-phrase are used in place of keywords.

There are two main types of large language models that are used in call center voice AI.

1. Intent-based models for call center voice AI

Intent-based models use natural language understanding to extract the 'intent' or meaning behind a spoken utterance. They then match the intent with an appropriate response from a pool of predefined utterances.

2. Generative AI models for call center voice AI

Generative AI models are not reliant on pre-defined 'intents.' Instead, they leverage their training data to match spoken inputs with relevant outputs that are generated in real-time.



Responding technologies for call center voice AI



Text-to-speech (TTS)

Text-to-speech models transform text transcriptions into spoken utterances. Traditional TTS has delivered robotic-sounding voices, but a new generation of TTS offers a more natural experience.



Speech synthesis

Natural-sounding voices that feel like speaking to real people ensure callers engage with the system, increasing call containment rates.



How to deploy call center voice AI

Once you're ready to get started with call center voice AI, you'll find that there are a number of different ways to design, build, and deploy in your contact center.

Here are the four different approaches to deploying voice assistants:

DIY from scratch

This approach allows for purpose-built technologies your company selects, deploys, and maintains directly.

Building from scratch requires a unique stack of foundational technologies, including systems for Automatic Speech Recognition (ASR), Natural Language Understanding (NLU), Natural Language Processing, alphanumeric capture, dialogue management, and speech synthesis. This approach also requires telephony and integration engineers, business analysts, and a dedicated product owner.

It's important to consider that this approach is an extremely expensive and time-consuming process that requires a full team of experts and often takes over a year to deploy, making it accessible only to very large companies, and often has a very slow time to impact on your customers and your bottom line.

Using off-the-shelf components

Another approach to deploy call center voice AI is to leverage existing technologies from a third-party provider that come pre-built and tested. This approach allows your developers to focus on assembling and configuring the components instead of having to build everything from scratch.

With this approach, you depend on multiple external providers to access the tech stack required to build a voice assistant and for ongoing updates, development, and technical support.



Using a DIY platform

A more practical DIY approach is to build off voice platforms such as Google Dialogflow and Amazon Lex. These platforms handle the underlying NLP and speech algorithms for you.

These platforms typically provide a variety of capabilities and are often favored by organizations that prefer to keep development in-house.

The no-code environment is designed to make it easy for non-technical teams to develop and maintain chatbots and voicebots. However, while many DIY platforms handle the underlying NLP and speech algorithms, this usually comes at the cost of control. You'll become locked into using a single supplier for every piece of the tech stack, preventing you from choosing the best-in-class technology in each area.

Voice-first vendors

Voice-first vendors provide a unique technology stack tuned exclusively for voice conversations to increase accuracy and account for context throughout an entire conversation.

Where many conversational AI vendors offer voice capabilities as part of a wider service offering, voice-first vendors provide a unique technology stack tuned exclusively for voice conversations to increase accuracy and account for context throughout an entire conversation.

Building, deploying, and maintaining a voice assistant requires a team of project managers, voice user interface designers, API developers, implementation designers, and testers (plus many more!). By working with a voice-first vendor, organizations can access a team of specialists without hiring anyone in-house.





Call center voice AI use cases



Account management

Chatbots and voice assistants can be used to help customers change their details, upgrade their accounts, and learn more about account services.

Voice deployments will require advanced technologies such as accurate ASR and phoneme matching to ensure that details like addresses and phone numbers are recorded accurately.



Answering FAQs

Automating FAQs with call center voice AI, such as voice assistants or chatbots enables businesses to give customers answers to their questions quickly.

By handling repetitive questions, agents can focus on more complex tasks. The most effective call center voice AI solutions enable customers to ask questions naturally without needing to use keywords to get the answer they need.



Authentication

Agents can spend a significant amount of time per call authenticating customers. Call center voice AI can be used to check callers are who they say they are by running callers through a series of security questions in the same way an agent would.

By automating this process, agents get time back to focus on more complex customer queries.



Billing & payments

Call center voice AI can be used to guide customers through billing and payments, just like a customer service representative. Voicebots can take payments, check refund statuses, balances, and more, all while identifying and offering upselling opportunities.



Bookings and reservations

Call center voice AI can integrate into existing booking platforms and processes to take new bookings, payments, or deposits, negotiate a new time or process cancellations, and contact customers with reminders and booking confirmations.



Call routing

Call center voice AI streamlines call routing using advanced speech recognition and natural language processing. When a customer calls, they interact with an automated system that prompts them to state the purpose of their call.

The AI then transcribes and analyzes the caller's response, identifying the intent based on keywords and phrases. Using this information, the system makes a quick decision on the most appropriate department or agent to handle the call, such as billing or technical support.



Order management

Call center voice AI frees up agents to add more value by empowering customers to track orders, report issues, and reschedule deliveries on their own terms without speaking to an agent. By integrating with CCaaS, CRMs, and order management systems, call center voice AI solutions can create seamless, personalized experiences for each customer.





The benefits of call center voice AI

As companies continue to search for new ways to retain staff, streamline processes and improve the customer experience, many are turning to call center voice AI to offer 24/7 support to their customers, handling routine and complex queries to deliver exceptional customer service.



The benefits of call center voice AI



Reduced call volume

Call center voice AI can help support teams by resolving up to 50% of customer calls, eliminating the need for human intervention. This not only reduces call volume but returns valuable time and resources back to the call center. These resources can be used to cut costs, address labor shortages, and enhance customer experience.



Mitigate seasonality and peaks in demand

Industries often anticipate and prepare for sudden spikes in call volume, such as back to school periods in retail and enrollment season in insurance. But some peaks in call volume are unexpected.

Spikes in volume lead to substantial hiring and training costs during peak seasons, budget shortfalls required to recruit enough workforce for higher volumes, and lower CX metrics due to long wait times.

With the ability to scale up as volumes increase, call center voice AI plays a crucial role in leveling out these peaks and minimizing the need for hiring seasonal staff.

Voice AI has helped organizations achieve:

50%

reduction in seasonal hiring

\$840k

saved in seasonal hiring costs

10x

improvement in self-service levels

20–30%

reduction in AHT





Customer insights from structured conversational data

Call centers typically have access to unstructured data, like call recordings and transcripts, which contain valuable customer information but lack organization.

Structured data, on the other hand, is categorized and analyzable, offering insights into customer intents, values (e.g., reservation details), and specific conversational turns.

Companies that use call center voice AI solutions can remove manual processes and automatically structure data during calls, gaining insights into customer behavior, peaks in specific queries, and product/service demand, enabling contact centers to enhance call routing accuracy, staff teams more effectively, address issues promptly, and optimize the overall customer experience while reducing operating costs.



Better CX

Call center voice AI can improve customer experience across your entire contact center program.

When agents are required to handle less low-value calls, wait times are reduced, and agents can spend more time with the people who need them most.

Call center voice AI solutions like voice assistants allow customers to get support anytime, all year round. They can also help connect customers with digital resources, bridging the gap between digital self-service and phone support.



Agent retention

High attrition rates, often within the first four months, are a chronic challenge in contact centers. Many experience annual attrition rates of 100%, but that doesn't necessarily mean replacing 100% of the staff; they typically replace the lower 25% of their staff up to four times.

By reducing the number of agents required and removing many of the mundane calls, retailers are able to hold onto their more qualified agents for longer and avoid the need for excessive entry-level hires.



Are you ready to implement call center voice AI?

We're on a mission to empower companies to be the best version of themselves for their customers. The notion that people can talk to computers has been around for a long time. Most of what's out there isn't very good.

Why? Because understanding the nearly infinite ways that something can be said, replying helpfully, and carrying on the conversation for as long as it takes to fulfill the customer's request — all by a robot — is a really hard problem to solve.

While most of our competitors have focused on chat because it's cheap, we are voice first. Customers will always call. So everything we do here at PolyAI — our products, services, research, and development — focuses on helping call centers deliver a best-in-class voice experience.



Speak to our team today

PolyAI is the inventor of customer-led voice assistants.

To find out more about how PolyAI can help
you put your customers in the driver's seat,
improve loyalty and reduce cost-to-serve,
[schedule a demo with our team.](#)

