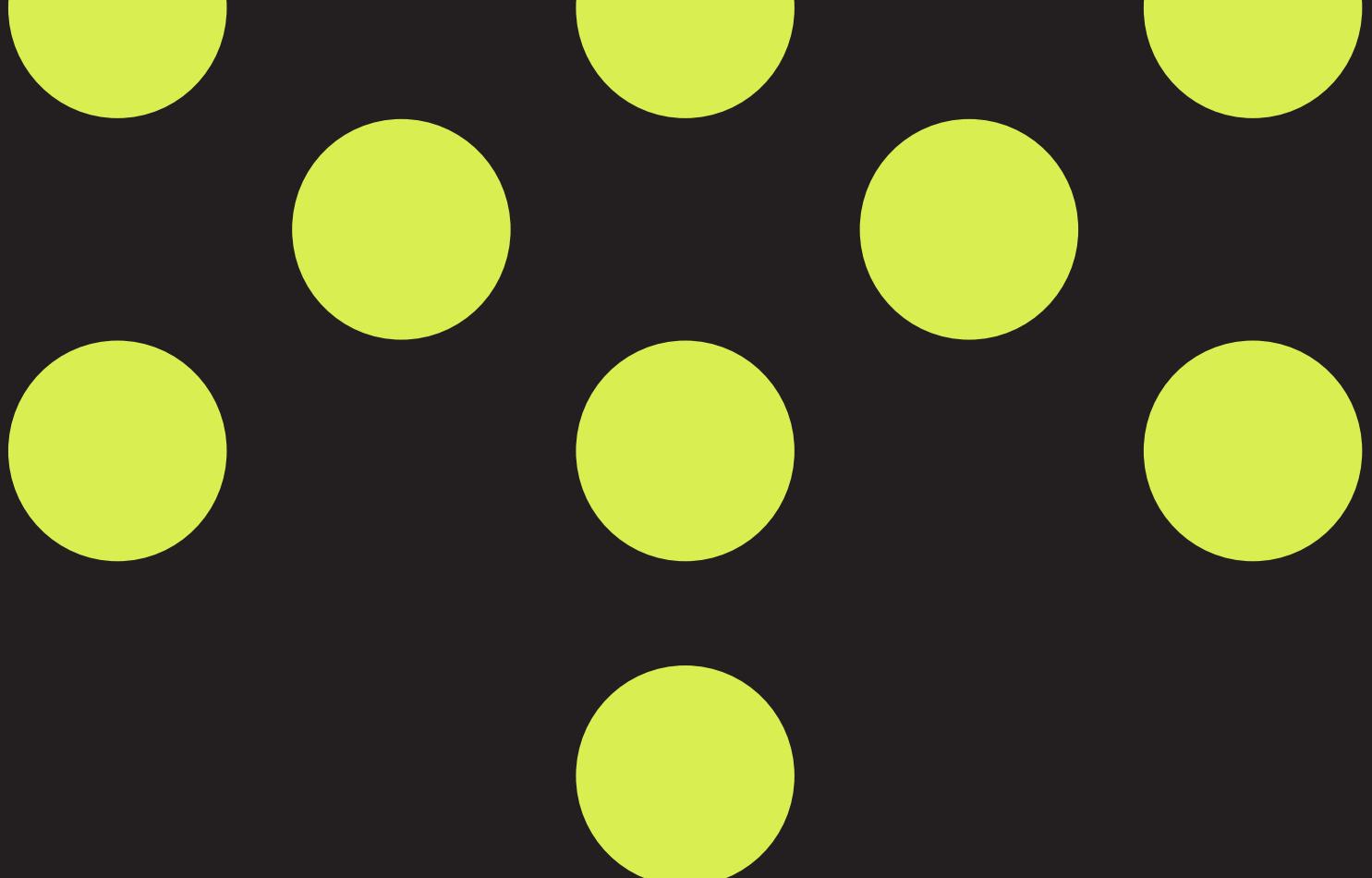




# Build vs Buy

---

How to choose the right  
voice AI solution  
for your contact center



## CONTENTS

Introduction.....	3
The voice AI tech stack .....	4
<b>Choosing the right solution for your conversational AI project .....</b>	<b>8</b>
DIY with modular conversational AI .....	10
Enterprise cloud providers.....	12
Contact center as a service (CCaaS) providers.....	14
Omnichannel platforms .....	16
Specialist voice platforms .....	18
PolyAI.....	20
Final considerations .....	22

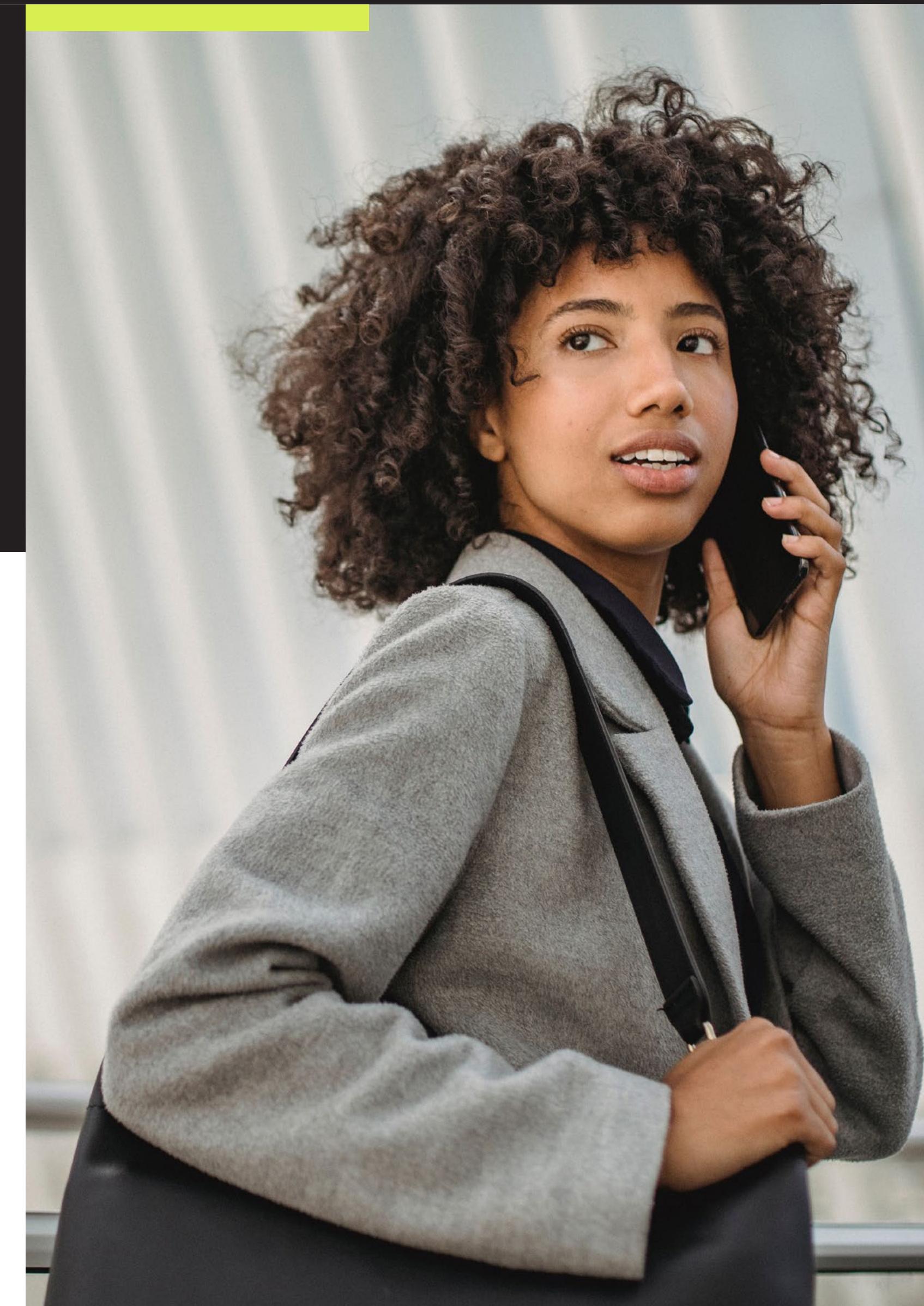
## INTRODUCTION

# The efficiency vs personalized service dilemma

Contact centers with limited staff often face a tradeoff between two priorities: serving every customer quickly or dedicating time and attention to individual customer needs. In other words, they must decide whether to prioritize efficiency or personalized service.

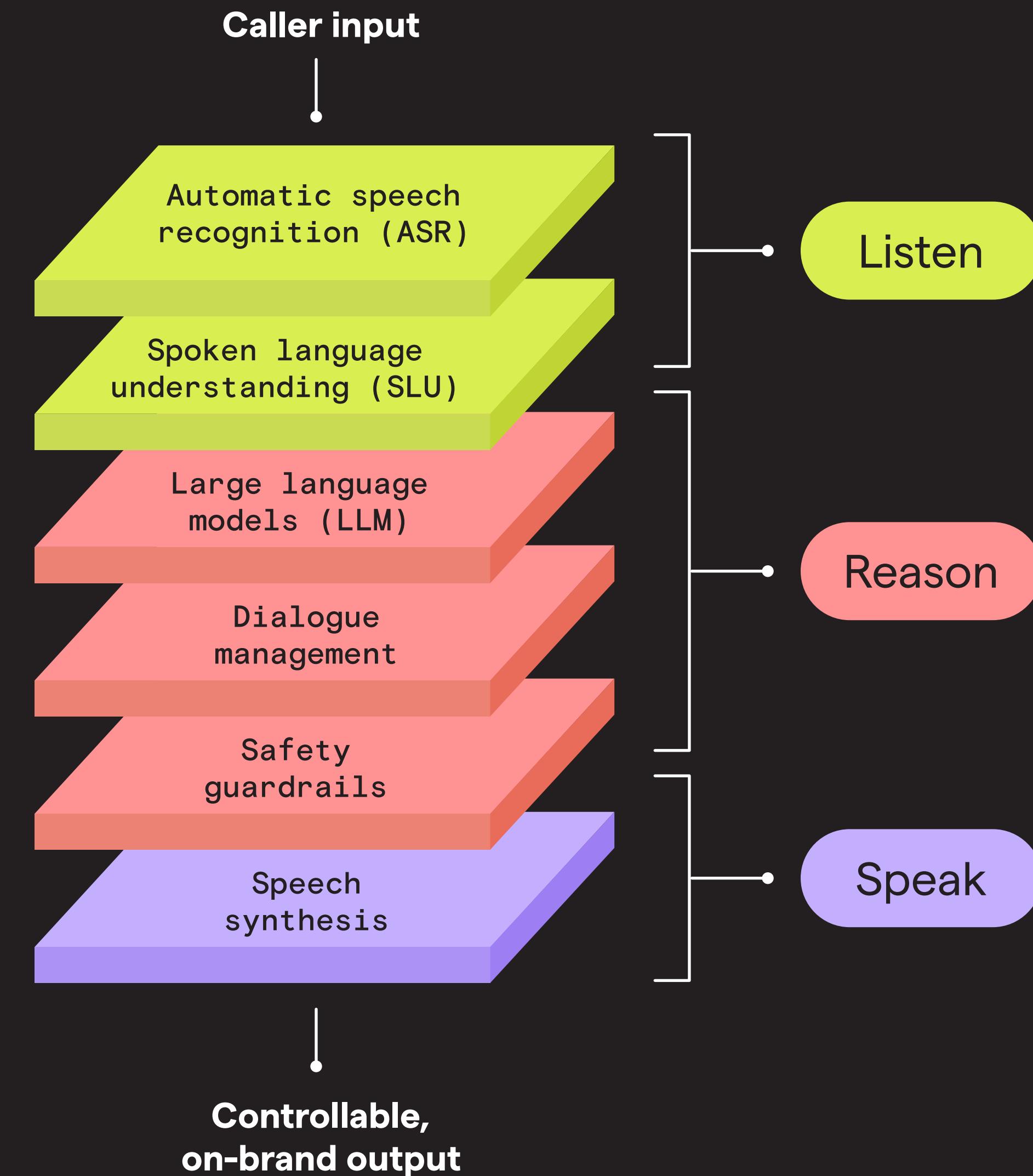
A new generation of voice assistants promises to solve this dilemma, offering great CX at scale. But with thousands of vendors touting similar offerings, deciding how to buy voice AI can be challenging.

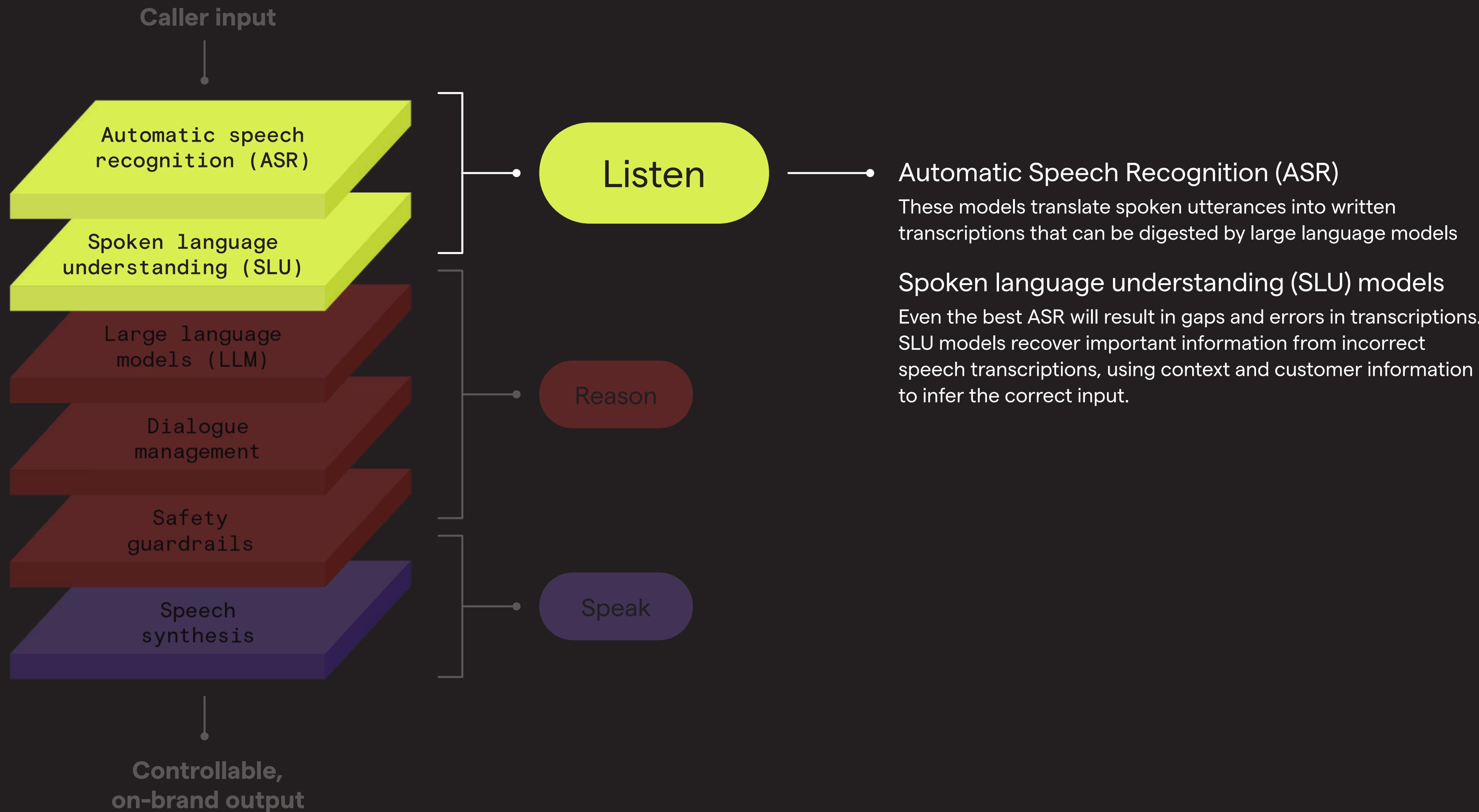
We've created this guide to give you a straightforward view of the voice AI landscape, the technology and resources required, and the options you have to buy voice AI. This will allow you to focus on what matters most to your business: delivering exceptional customer experiences.

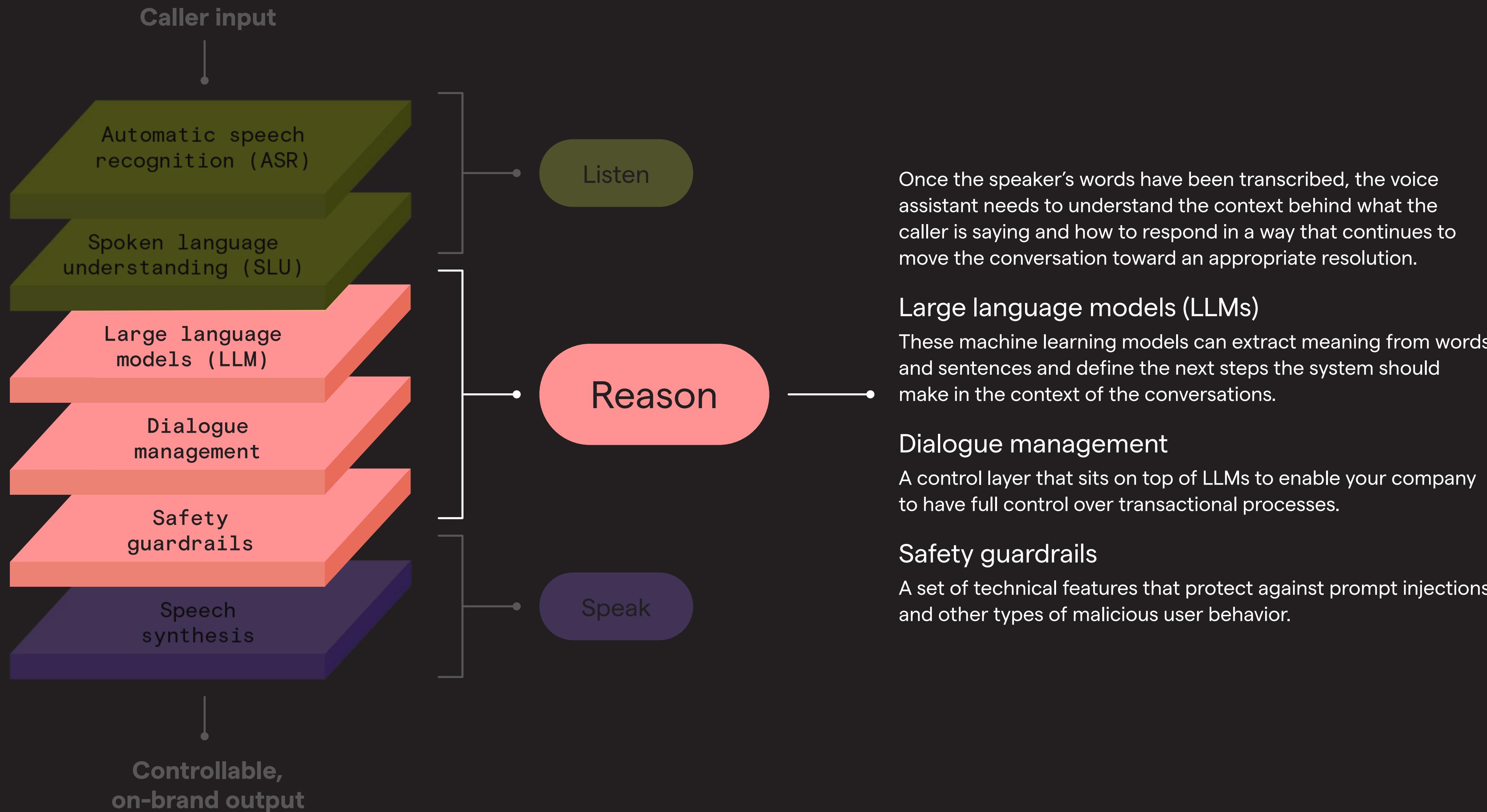


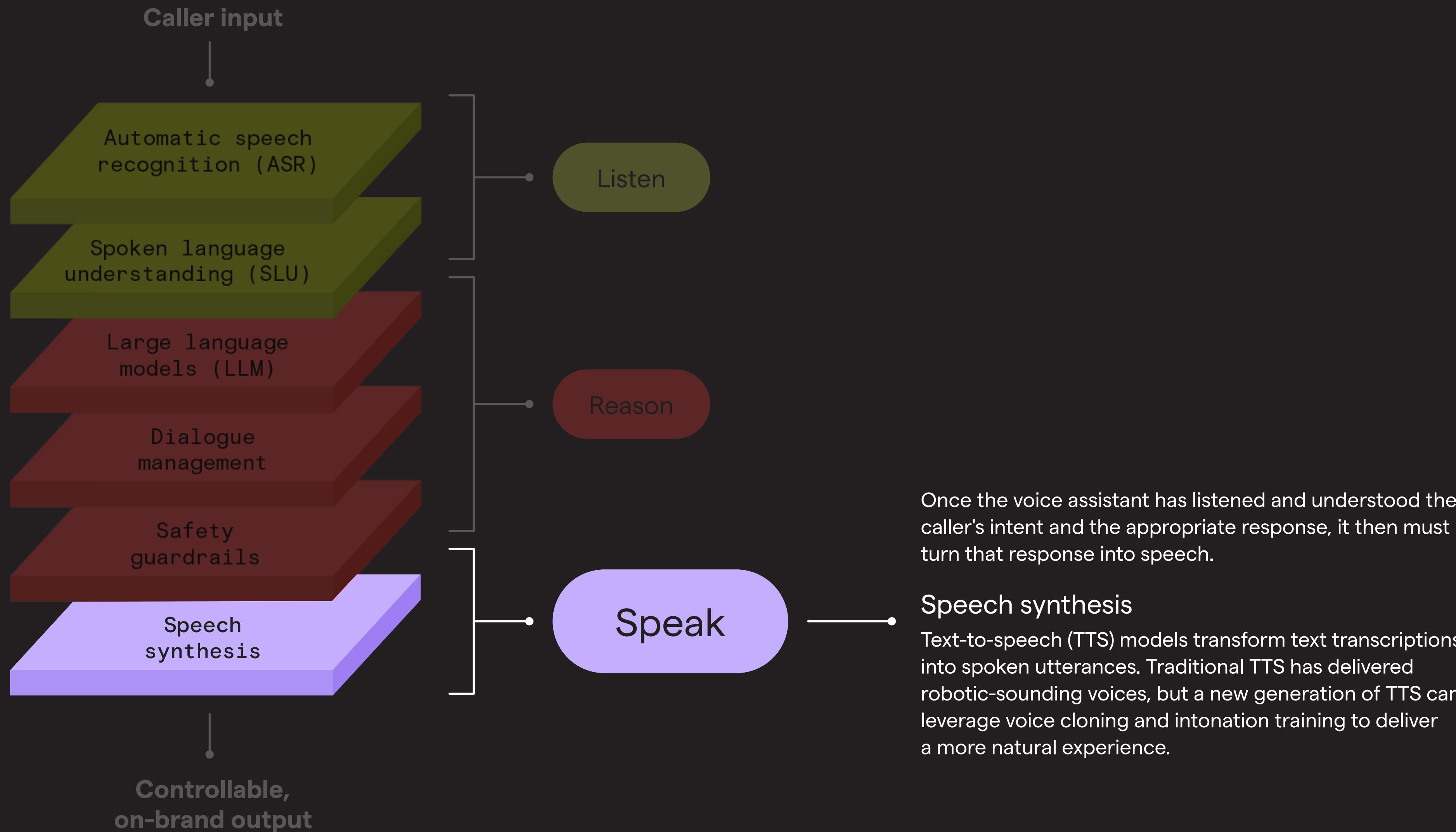
# The voice AI tech stack

To deliver great customer experience over the phone, a voice assistant must successfully listen, understand, and respond to what the customer is saying. This requires a specific tech stack to handle the nuances of natural conversation.



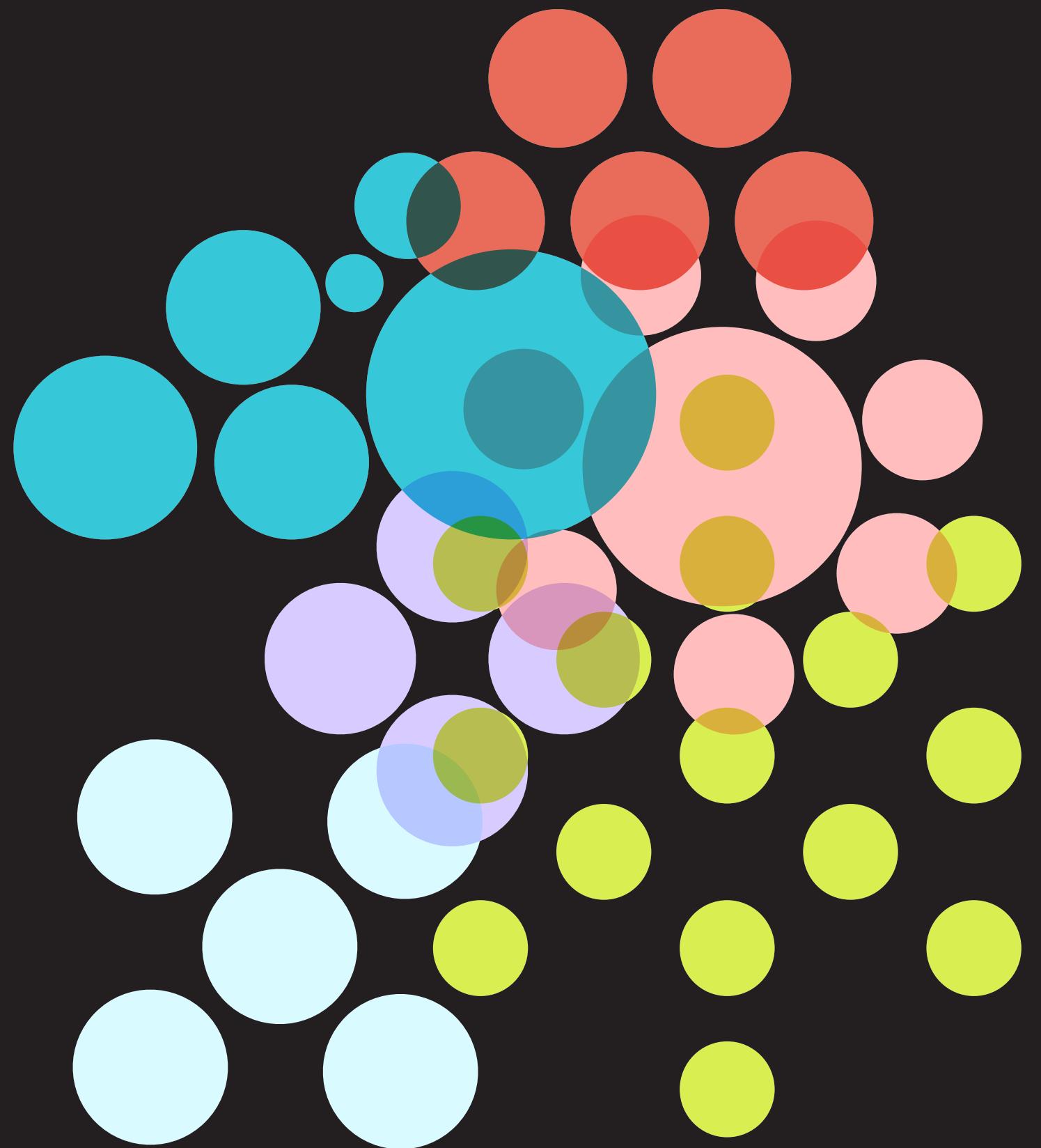






# Choosing the right solution for your conversational AI project

Finding the right solution in a crowded voice AI marketplace while managing the day-to-day running of your contact center is no easy task. There are thousands of vendors offering conversational AI solutions. Who you choose will make or break your project.



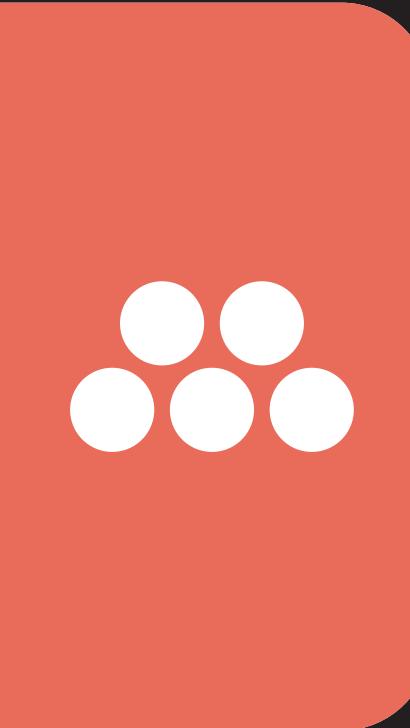
## We've divided the voice AI marketplace into the following categories:

**DIY WITH MODULAR CONVERSATIONAL AI**

---

These companies provide the building blocks for a more pure DIY approach that relies less on builder platforms.

---

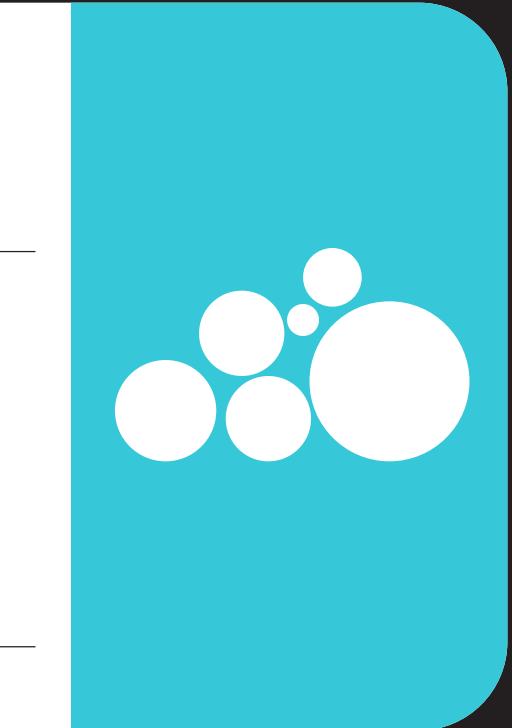


**ENTERPRISE CLOUD PROVIDERS**

---

Providers like Google and Amazon offer their own low-code tools for building conversational assistants.

---

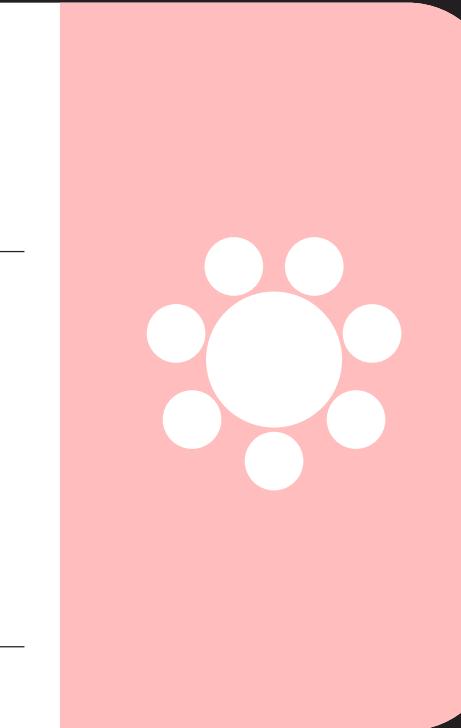


**CCAAS PROVIDERS**

---

Offer a complete suite of software for contact centers. In recent years, they've added automation to their product offerings.

---

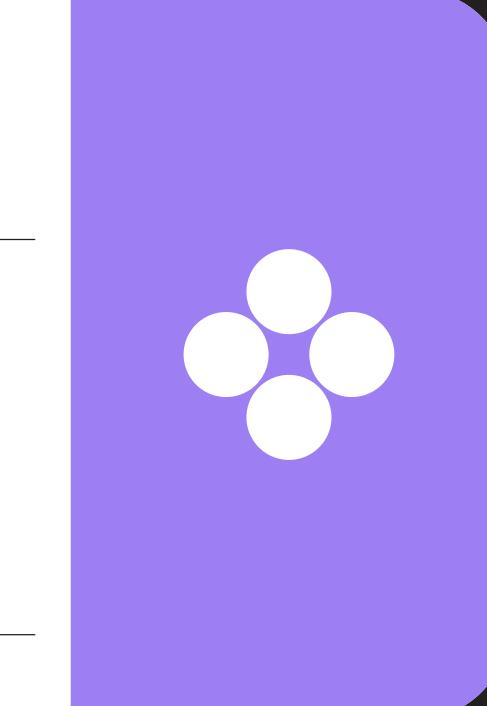


**OMNICHANNEL PLATFORMS**

---

These platforms offer a wide suite of tools for contact center automation.

---

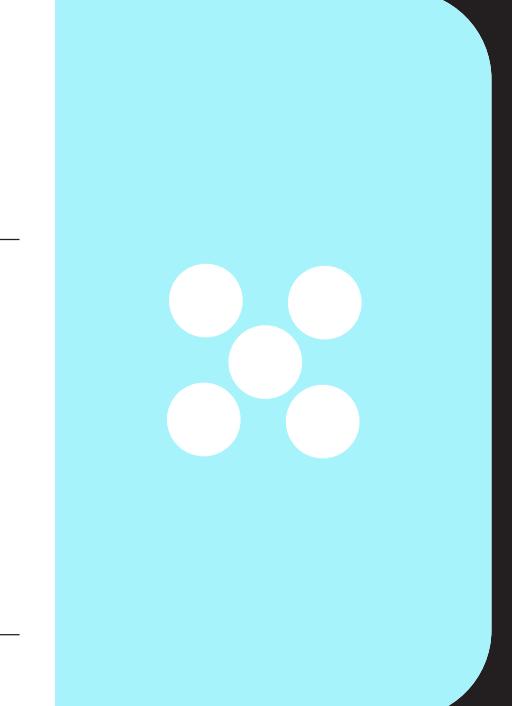


**SPECIALIST VOICE PLATFORMS**

---

These companies focus specifically on building conversational experiences for voice use cases.

---

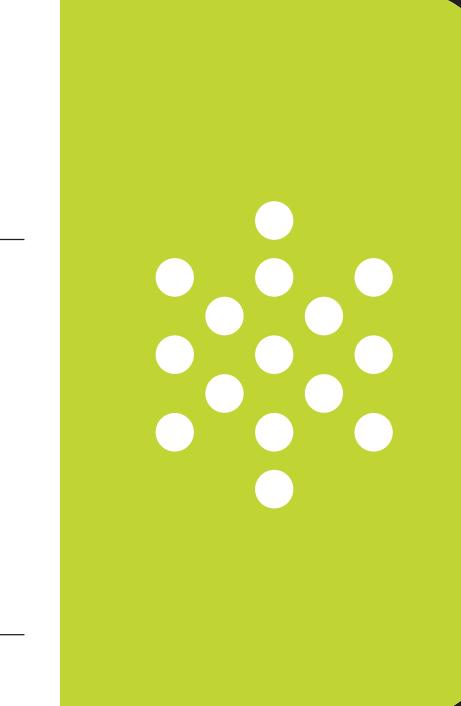


**POLYAI**

---

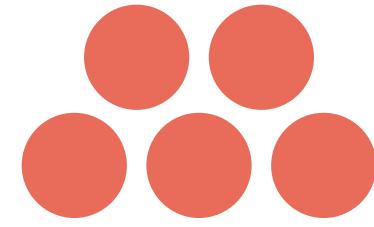
PolyAI offers the world's most lifelike voice AI for enterprise customer service.

---



Some routes to buying voice AI are more resource-intensive than others and will impact cost, quality, and the speed of deployment.

Let's take a look at the different types of solutions and how they approach voice AI deployment.



# DIY with modular conversational AI

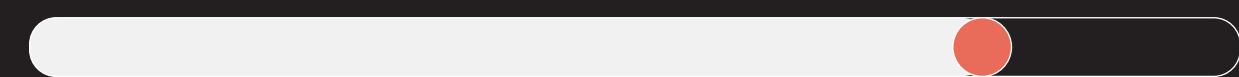
This approach enables your company to build a custom voice AI solution by selecting and managing the technologies yourself. You'll need to include key components like text-to-speech (TTS), speech-to-text (STT), and large language models (LLMs).

There are several critical considerations if you choose to deploy voice AI using this approach.

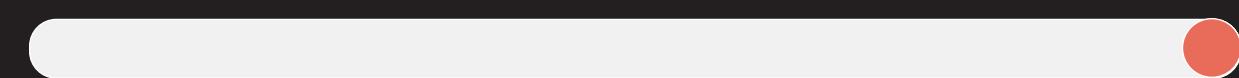
## DIY WITH MODULAR CONVERSATIONAL AI



### QUALITY



### REQUIRED EXPERTISE



### TIME-TO-VALUE



### SUPPORT



### RISK



### COST



## 🏆 Quality

High-quality experiences start with high-quality components, each requiring fine-tuning for optimal performance. Out-of-the-box text-to-speech and ASR solutions offer moderate accuracy, especially over the phone. SLU techniques like ASR biasing, N-best selection, and phonetic matching can improve results. Speech-to-text models may sound robotic, so choose solutions with customization options to reflect your brand voice. While advancements in LLMs make it easier than ever to automate conversations, staying on-brand and on track requires strong safety guardrails and dialogue management frameworks.

## ❤️ Support

You may receive support from vendors on each component, but it will be your responsibility to make the components work together. You'll need to evaluate the availability and quality of support from the providers of your TTS, STT, and LLM components, as well as the strength of your in-house support for ongoing maintenance.

## 💡 Required expertise

Your team needs strong technical expertise in AI, machine learning, and software integration. Given that this is a voice application, they'll also need a good telephony background and experience fine-tuning text-to-speech and speech-to-text solutions to deliver optimized experiences at every turn of the conversation.

## ❗ Risk

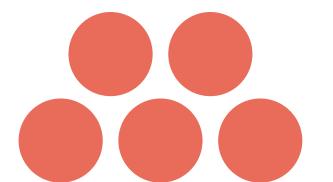
The technology and expertise required to deploy an effective voice assistant explain why only 11% of enterprise applications reach production. While solutions may perform well in testing, scaling introduces complexities like speech recognition errors and background noise, which can cause them to deliver a poor customer experience. There's a significant risk that what you build will not reach deployment-level quality, resulting in sunk cost and effort.

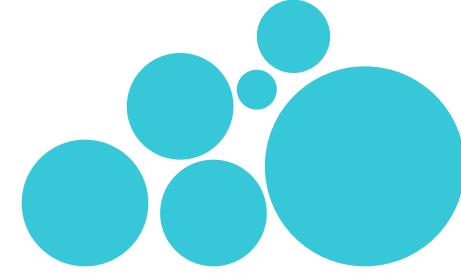
## ⌚ Time-to-value

This approach is an extremely time-consuming process that requires a full team of experts. It can take over a year to carry out the initial build.

## ฿ Cost

You'll need to consider telephony costs and licensing fees for each component. You'll also need a team of 5+ experienced machine learning and conversational AI experts at competitive salaries.





# Enterprise cloud providers

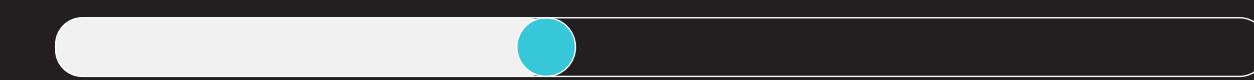
This approach enables you to build voice assistants on top of platforms developed by enterprise cloud providers. These platforms handle the underlying natural language processing (NLP) for you and typically provide a variety of capabilities.

You might take this approach to keep your AI efforts contained in your existing cloud platform.

## ENTERPRISE CLOUD PROVIDERS



### QUALITY



### REQUIRED EXPERTISE



### TIME-TO-VALUE



### SUPPORT



### RISK



### COST



## 🏆 Quality

These platforms mainly work with chat use cases, they only support limited voice-centric features. Their models are not built with the voice experience in mind, meaning it is not very good at handling accents, dialects, silences, and short responses.

While many cloud providers 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, which prevents you from choosing the best-in-class technology in each area.

## ❤️ Support

Support typically comes at a cost and is often provided by professional services firms rather than the cloud providers. This can include setup assistance, custom development, and ongoing support.

## 💡 Required expertise

While the platforms are relatively simple to use, especially for basic tasks, achieving sophisticated conversational AI requires expertise in conversational design and machine learning. Implementation is typically handled by professional services consultancies at an additional cost.

## ❗ Risk

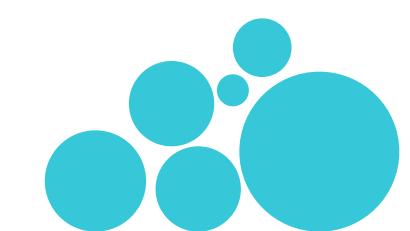
These systems work best for text-based chat rather than spoken conversations over the phone, which means there is a significant risk that the initial build may never reach deployment-level quality.

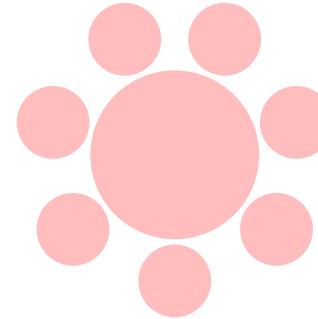
## ⌚ Time-to-value

Expect an initial build to take 6-12 months, depending on the complexity of the project and your team's skill level. This timeline includes the design, development, testing, and initial deployment phases.

## 💲 Cost

While the platforms themselves are relatively low cost, expect additional expenses for professional services, telephony integration, support, and maintenance. You'll also need to hire 3-5 additional full-time employees with conversational AI expertise to manage and optimize the system.





# Contact center as a service (CCaaS) providers

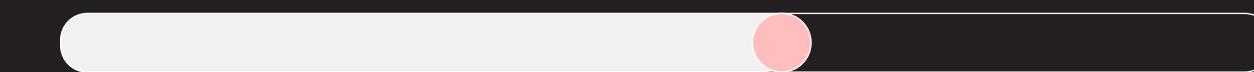
CCaaS providers offer a complete suite of software solutions for contact centers, including call routing, customer relationship management, and analytics. In recent years, these providers have added automation and AI capabilities, such as conversational AI and chatbots, into their product offerings.

## CCaaS PROVIDERS

### QUALITY



### REQUIRED EXPERTISE



### TIME-TO-VALUE



### SUPPORT



### RISK



### COST



## 🏆 Quality

CCaaS providers offer conversational AI quality similar to big cloud providers. They have low-code tools for building simple voice apps like conversational IVRs, where callers state a keyword to trigger a handoff to an agent. These systems struggle with complex transactional call flows, so if you need a voice assistant that handles calls fully, they may not be the best option. If a robotic IVR-like experience is for you, it's worth considering, but adding advanced features later is unlikely. For more complex needs, expect to start building again from scratch.

## ❤️ Support

CCaaS providers typically offer support for implementation, integrations, testing, fine-tuning, and ongoing support, but it comes at a cost. The level of support you'll need is difficult to estimate upfront and will vary depending on your use case. For smaller businesses with limited budgets, the cost of professional services may be prohibitive.

## 💡 Required expertise

Although CCaaS platforms are designed to be user-friendly, implementing and optimizing voice AI requires specialized skills. You'll need a team with expertise across conversational design and machine learning.

## ❗ Risk

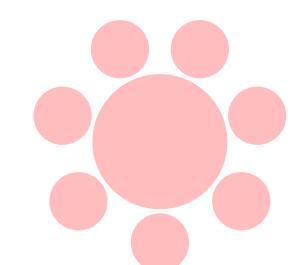
CCaaS providers offer voice capabilities as part of a wider service offering and don't offer best-in-class technology to solve the uniquely difficult problem of automating customer service interactions over voice. Therefore there's a risk of investing heavily in a solution that ultimately proves to be unusable.

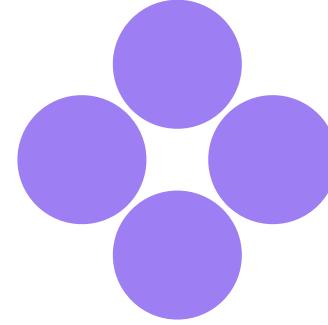
## ⌚ Time-to-value

Deployment time can vary based on the complexity of your requirements and the capabilities of the CCaaS provider. Expect an initial build time of 3-9 months.

## ฿ Cost

You will need to account for ongoing costs, including professional services, support fees, and potentially 3-5 additional full-time hires for AI and machine learning expertise.





# Omnichannel platforms

Omnichannel platforms offer a wide suite of tools for contact center automation across web, mobile, messaging apps, email, social media, and SMS. These platforms enable you to build, deploy, and manage chatbots and voice assistants.

## OMNICHANNEL PLATFORMS

 QUALITY

 REQUIRED EXPERTISE

 TIME-TO-VALUE

 SUPPORT

 RISK

 COST



## 🏆 Quality

Most omnichannel platforms offer decent text-based chat. But as these platforms try to meet a wide range of requirements, voice AI-specific capabilities are often weaker. This forces developers to cram voice functionality into a system primarily designed for text, where the flow must work for both channels. However, phone conversations lack a visual interface for scrolling through documents like T&Cs or buttons for ways to respond. The result is an unnatural, robotic voice experience that drives customers to insist on speaking to an agent.

## 💡 Required expertise

These platforms are useful for building simple proofs of concept but getting something truly enterprise-grade is dependent on professional services, so few in-house skills are required outside of IT support for building out API integrations.

## ⌚ Time-to-value

Initial deployment timelines can vary, typically ranging from 3 to 6 months, depending on the complexity of customization and integration with existing systems.

## ❤️ Support

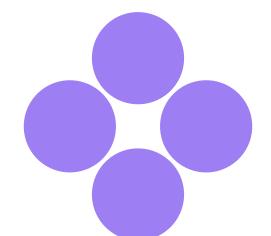
While these platforms are often marketed as DIY, achieving optimal performance usually involves significant professional services for customization, ongoing maintenance, and support.

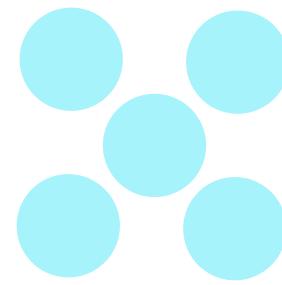
## ❗ Risk

Conversations with a voice assistant should feel natural to create an engaging user experience. The less a user notices the voice quality, the more likely they are to stay focussed on what they want to do and stay on the call. When the voice experience is poor quality, customers are likely to insist on speaking to an agent, meaning you end up paying for the same call twice.

## ฿ Cost

Costs include professional services for setup and customization, fees per minute of usage, charges for updates and maintenance, and ongoing operational costs and support.





# Specialist voice platforms

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

## SPECIALIST VOICE PLATFORMS

### QUALITY



### REQUIRED EXPERTISE



### TIME-TO-VALUE



### SUPPORT



### RISK



### COST



## 🏆 Quality

The quality of voice AI can vary significantly depending on the platform. Leading platforms with proven success in similar deployments can deliver near-human level conversations, thanks to advanced NLP, speech synthesis, and fine-tuned LLMs.

The best voice platforms don't require users to change how they speak to fit the technology's limitations.

## 💡 Required expertise

Because of the high level of customization needed to deliver customer-worthy experiences, these platforms typically offer professional services to build a truly enterprise-grade voice assistant. Low code platforms should be considered as 'playgrounds' to try the technology, enabling configurability controls after the voice assistant has been built and deployed.

## ⌚ Time-to-value

The initial build time varies by platform. Some may take 3-9 months. Those with a proven track record of live deployments can complete it in 1-3 months, depending on the complexity of the use case.

## ❤️ Support

Ongoing support should be part of the service agreement. Your specialist voice provider should monitor the performance of your voice assistant, analyze caller interactions and make any necessary adjustments.

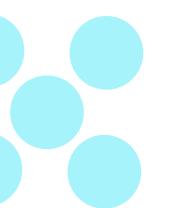
Continuous improvements will be essential as you learn more from each customer interaction. Experienced voice specialists will monitor key metrics, edit prompts, and make tweaks to enhance customer experience and efficiency.

## ❗ Risk

Working with a specialist voice platform should come with limited risk, so you need to choose a vendor with a proven track record of successful live deployments. Assess vendors based on their technology stack, use cases, and customer success stories. You should be able to call into live deployments for real brands to try the solution as a consumer.

## 💲 Cost

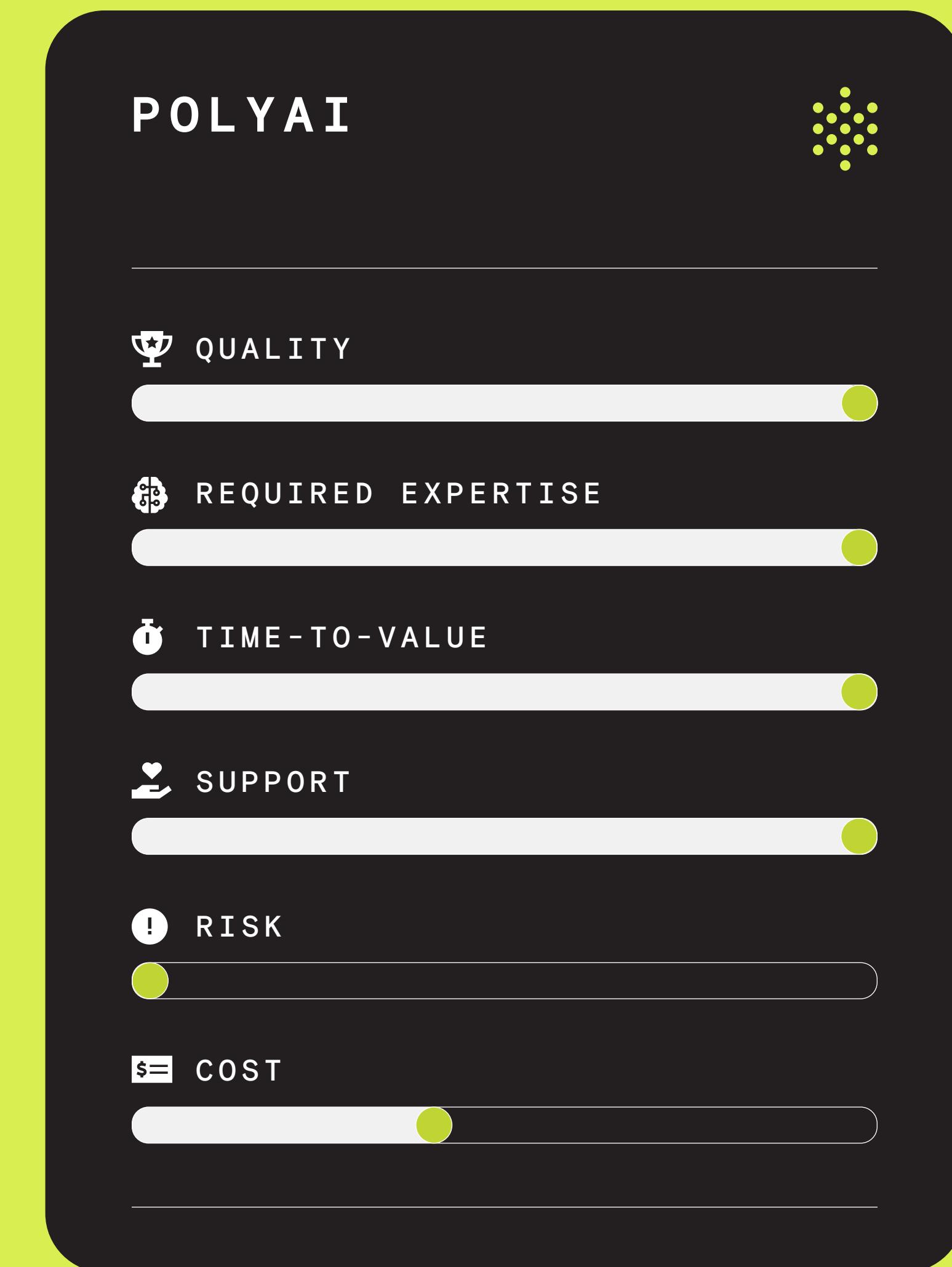
If you take this approach, you'll need to consider costs for professional services, cost per minute, cost per update, and ongoing maintenance fees.





# PolyAI

PolyAI offers the world's most lifelike voice AI for enterprise customer service. With voice AI that feels like talking to a real person, customers can expect to see a 50% reduction in call volume, dramatically decreasing call abandonment rates, improving employee experience, and empowering every customer to get the support they need when they need it.



## 🏆 Quality

Our industry-leading voice platform is engineered specifically for spoken conversations over the phone. It enables our voice assistants to understand what callers say, however they speak, guide callers through business logic to complete transactions, and respond appropriately in an engaging tone.

A combination of ASR models supercharged with SLU features delivers human-level understanding, enabling callers to speak naturally. Dialogue management frameworks ensure LLMs stay focused on the task while allowing free-flowing conversation in a custom-branded voice tailored to your business.

## ❤️ Support

PolyAI joins your internal project team to truly collaborate on your voice assistant throughout the implementation process. Voice AI requires close monitoring to ensure the highest quality of performance and to make adjustments to improve based on your business goals

## 💡 Required expertise

Our industry-leading experts in voice AI, dialogue design, and telephony collaborate closely with your internal project team to deliver a voice assistant that feels almost indistinguishable from speaking with a real person.

## ❗ Risk

We have experience serving over 50 enterprises with voice assistants that can handle use cases across a variety of industries. PolyAI has deployed a number of generative-AI powered voice assistants that handle customer service calls for some of the world's best-known brands. We have robust safety frameworks in place that are fine-tuned for each customer's requirements and tested rigorously to remove risk before deployment.

## ⌚ Time-to-value

A simple telephony integration is all that's required to route calls between your voice assistant and your team. This is virtually the same for every voice assistant, which means you can start taking customer calls in just 8 weeks.

## 💲 Cost

We charge an up front fee for the initial build of your voice assistant and cost per minute for calls taken so you only pay for what you use.



## CONCLUSION

# Final considerations

Deploying a voice assistant in your contact center can greatly enhance customer experience and operational efficiency, but the path you choose and the solution you work with will significantly affect the outcome.

The right voice AI solution will help your contact center handle routine calls efficiently, allowing your agents to focus on more complex and valuable interactions. By carefully evaluating your options and aligning them with your business goals, you can enhance customer satisfaction and deliver exceptional customer experiences.





PolyAI is  
the world's  
most lifelike  
voice AI.

PolyAI offers the world's most lifelike voice AI for enterprise customer service and delivers high-quality, human-like conversations optimized for customer engagement at scale.

PolyAI helps enterprises be the best versions of themselves in every customer call by consistently delivering the best brand experience, achieving accurate resolution, and uncovering data-driven business opportunities.

Sign up for our monthly demo, to find out more about how PolyAI can help you answer every call immediately, improve loyalty, resolve over 50% of calls, and deliver effortless CX at scale.

**Expect to learn more about:**

- The ROI of voice AI
- The best way to design voice experiences for customer engagement
- How other companies have successfully deployed voice AI
- How to begin your voice AI journey

[Register now](#)