State of Voice AI | 20 25

The Rise of Enterprise Voice AI Agents

Presented by

Deepgram + **opusresearch**

Report by

Table of contents

2025 is the year of human-like voice AI agents

Introduction

Past surveys indicated that 2022 was the year of voice technology, while 2023 was the year of AI.

Now, in 2025, these themes have converged– **we are in the year of *voice AI technology*, specifically voice agents.**

Advancements in Large Language Models (LLMs) have significantly enhanced the capabilities of voice agents. These models now offer improved natural language understanding, enabling seamless, context-aware conversations with users across various domains. More specifically, according to a16z, "Advancements in model development have streamlined the infrastructure "stack," resulting in voice agents with lower latency and improved performance. This improvement has largely materialized in the last six months with new conversational models."

Innovations in speech recognition and synthesis, alongside real-time processing, now allow for near-instantaneous, personalized, and emotionally aware interactions. The integration of auditory capabilities and a focus on privacy and ethical considerations further solidify voice AI's transformative role in user experiences. As a result, 2025 marks a key turning point for the widespread adoption and sophistication of voice technology.

With these insights in mind, we decided to examine the voice tech landscape first-hand. Deepgram's 2025 State of Voice AI Survey, conducted in partnership with Opus Research, explores the applications and key features of voiceAI across more than a dozen industries, based on insights from 400 business leaders.

Now, let's see why 2025 is shaping up to be one of the most thrilling years for voice AI and human-like voice AI agents.

Whom we surveyed

To generate a representative dataset, we surveyed 400 business leaders across North America. The vast majority—82 percent—were based in the U.S. Among the respondents, the majority (83%) came from large enterprises with over $100M in revenue. In fact, 36 percent of respondents came from organizations with over $1B in revenue.

Even more specifically, 42 percent of respondents are key decision-makers, including C-suite executives, (senior) vice presidents, and heads of business units in their respective organizations. See Figures 2 and 3 below for more details.

Additionally, 29 percent of respondents are technically-inclined, working in the Engineering, Data Science, or Artificial Intelligence departments. Another 21 percent reported working in Product Management, a role that often requires technical expertise. Thus, the number of respondents who are familiar with the intricacies of coding and AI architecture could be as high as 50 percent.

Whom we surveyed

Department of Business Unit

Department	%
Engineering	15%
Data Science	10%
Business Analytics	9%
Contact Center Operations	13%
Customer Experience	14%
Marketing Research & Innovation	14%
Product Management	21%
Artificial Intelligence / Generative AI	4%

Current Job Title

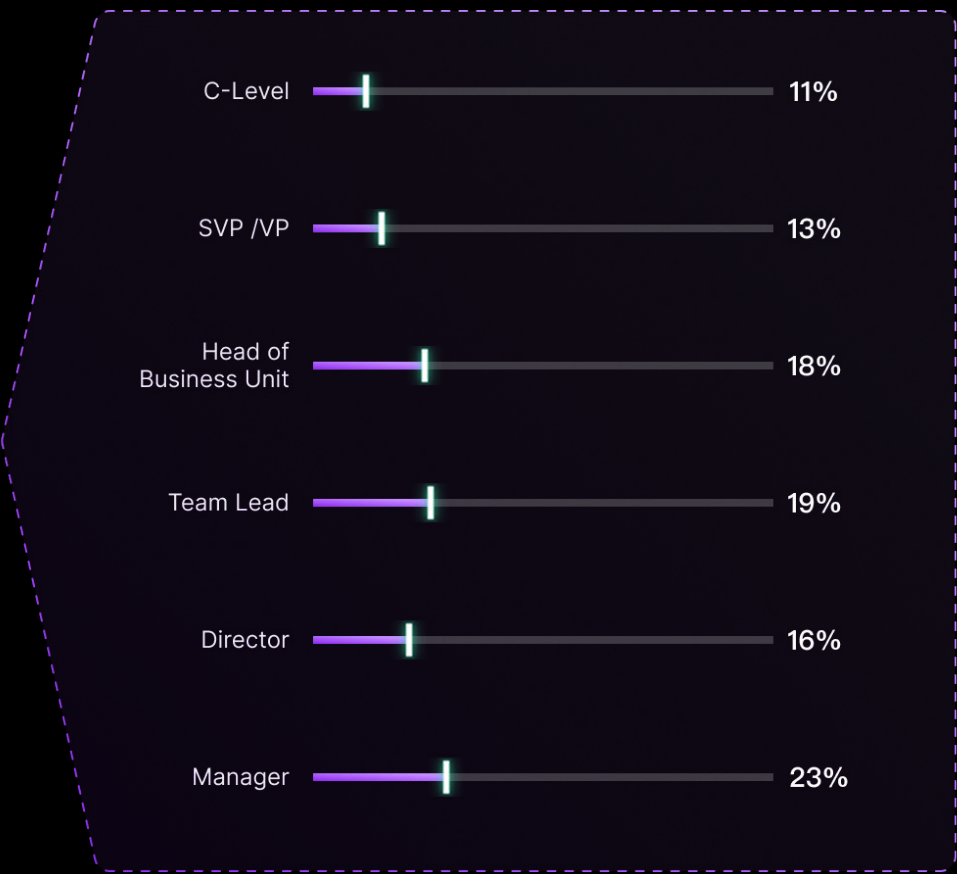| Title	%
C-Level	11%
SVP /VP	13%
Head of Business Unit	18%
Team Lead	19%
Director	16%
Manager	23%
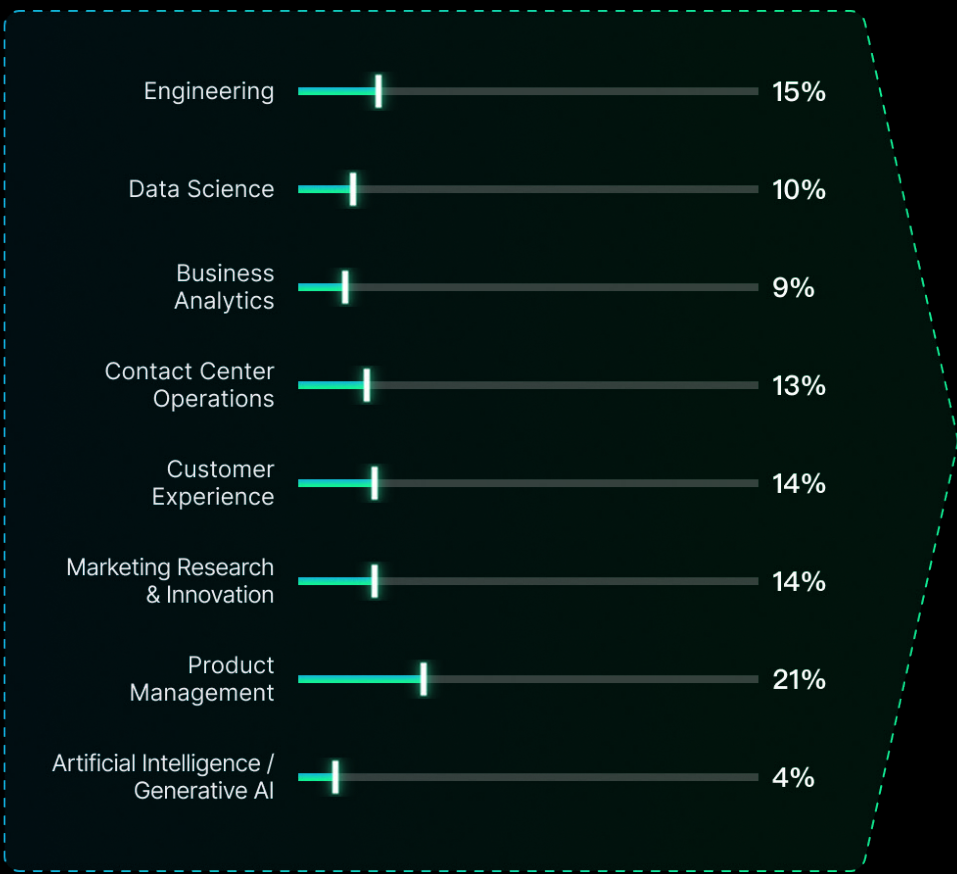
Base: All respondents (n=400)

We ensured a diverse representation of industries in our survey. Retail/E-commerce remains the best-represented industry (as it did in 2023), accounting for 16 percent of the respondents. Following closely in second place is the Restaurant/QSR (quick service restaurant) industry at 14 percent. Other industries represented include healthcare, telecom, finance/banking, insurance, and travel/hospitality.

Annual Revenue (USD)

5%	12%	18%	12%	17%	22%	14%
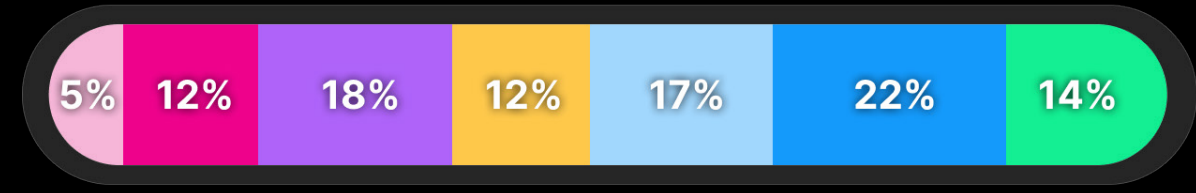
- < Under $50M
- $50M to $99M
- $100M to $249M
- $250M to $499M
- $500M to $999M
- $1B to $5B
- More than $5B

Decision-maker

7%	36%	51%	4%	2%
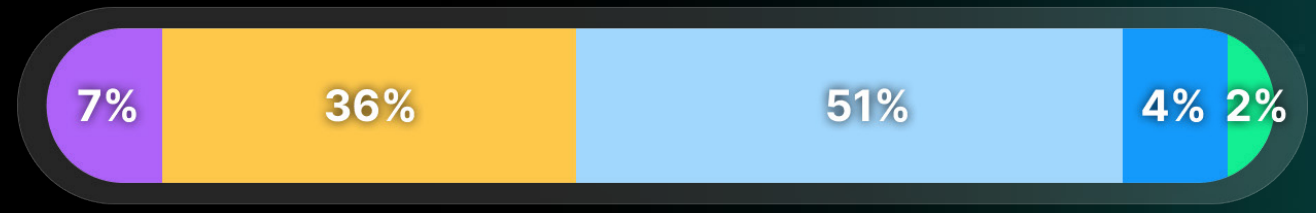
- CEO
- CTO
- Product Management
- Developer
- Engineering/Data Science

The chart below provides a more complete breakdown of industries covered in this survey.

Primary Industries Surveyed

13% Healthcare
11% Telecom
8% Contact Center / BPO
13% Finance / Banking
13% Insurance
14% Restaurant / QSR
16% Retail / E-commerce
12% Travel / Hospitality
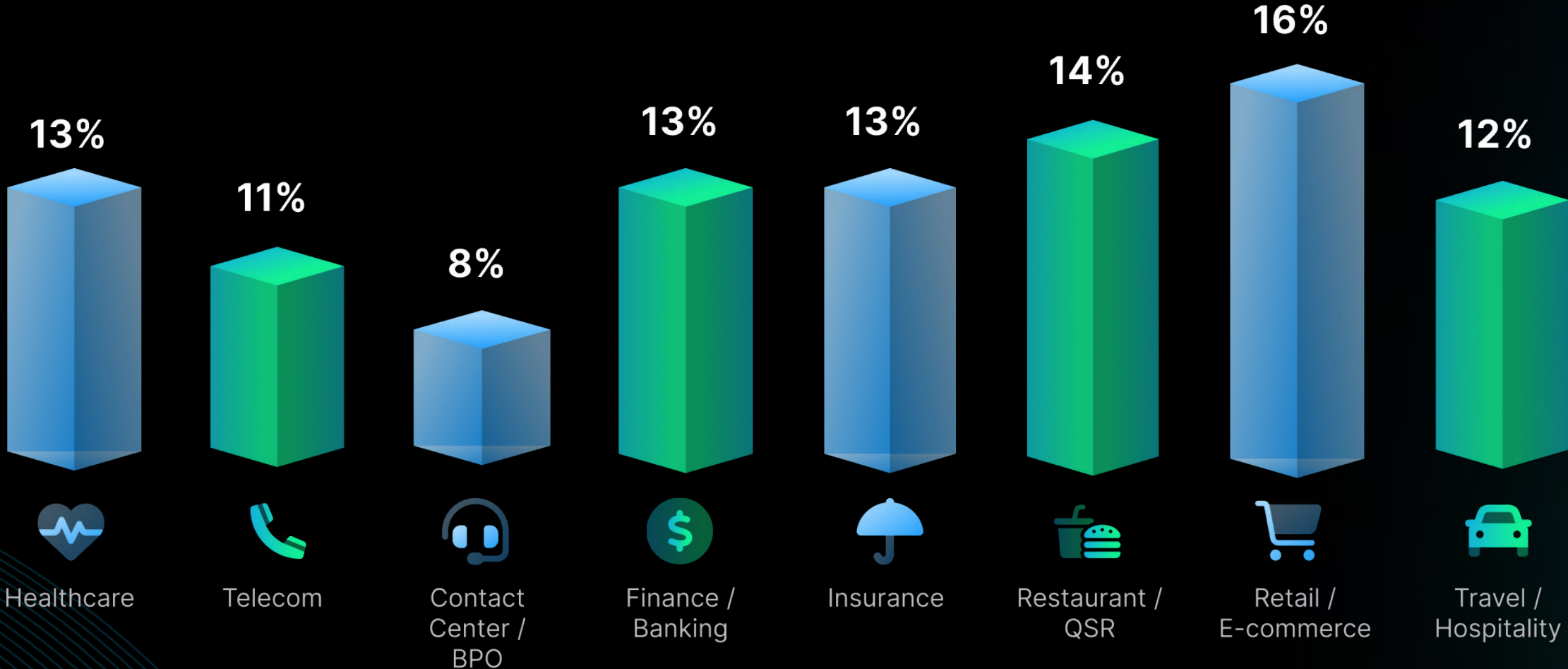
Regardless of industry, most companies share a similar set of internal departments. Whether you're working at a bank, a hospital, a restaurant, or an e-commerce store, you will need to figure out business analytics, engage in customer service, and maintain some type of data science functionality. In fact, for organizations operating at scale, certain departments are essential to remain competitive.

As we'll see below, voice AI remains a key tool for businesses looking to thrive.

Executive Summary:

Voice AI: No longer optional, now foundational for AI innovation

1 **Transcription: From Feature to Fundamental**

92% of respondents capture their speech data, and **56%** transcribe more than half of their interactions.

% audio transcribed

% of respondents

75-99% **2%**

50-74% **54%**

25-50% **42%**

10-25% **2%**
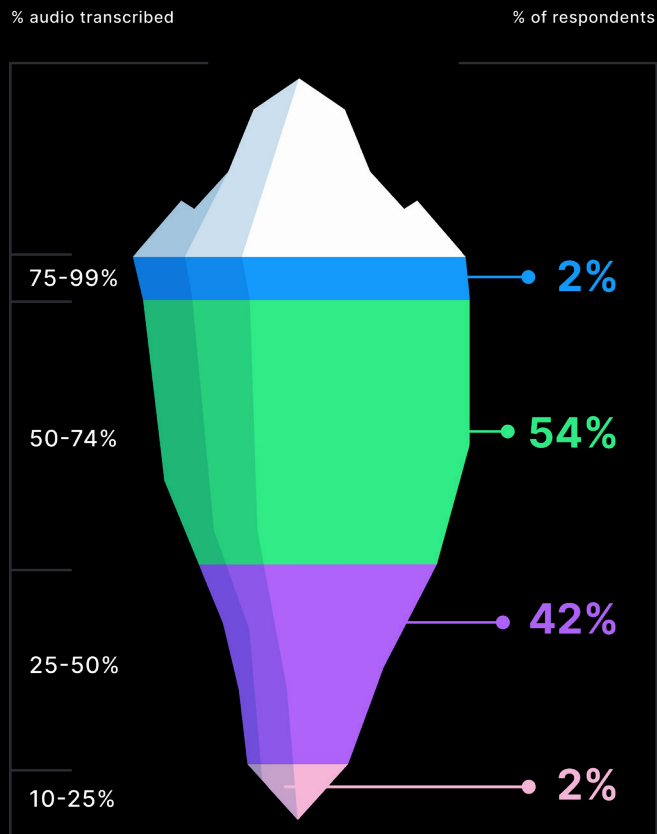
2 **Voice AI is Now Foundational**

67%

of organizations consider voice AI core to their product and business strategy

3 **The voice AI Agent Revolution Awaits**

80%

of organizations use traditional voice agent systems, yet only **21%** are very satisfied.

21%
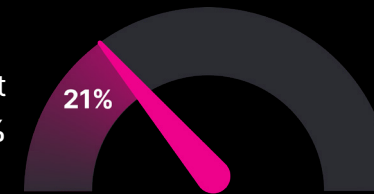
4 **Investment in Voice Technology is Rising**

84%

plan to increase budgets in the next 12 months

5 **Customer Service Automation: The Launch Pad for Voice AI Agents**

50%

use traditional voice agents for task/ service automation and consider it the most compelling use case for voice AI agents.

task/service automation

50%

50%

other
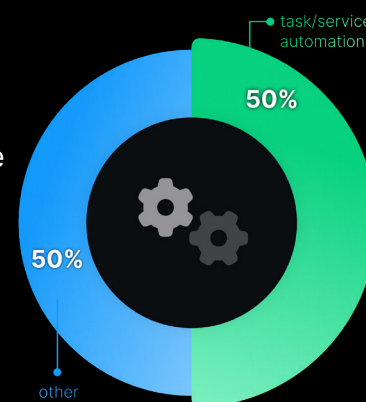
6 **Model Customization Drives Voice AI Adoption**

46%

cite fine-tuning models as a key to greater adoption

7 **Racing Towards Human-like Voice AI Agents**

15% of organizations are already actively developing voice AI agents.

98% of these plan to deploy within 12 months.

Key Takeaways

1 Voice Technology is Expanding and Becoming Foundational

Voice AI is rapidly shifting from an emerging technology to a foundational pillar of business strategy, with organizations increasingly investing in AI-driven voice technology to automate customer interactions, streamline workflows, and drive compliance and accessibility. The survey data reveals that transcription is becoming table stakes and a gateway into developing voice AI applications.

97% of respondents are currently using voice technology of some kind—automated speech recognition, legacy voice agents, text-to-speech, and speech analytics.

92% of respondents capture and analyze their speech data, with a majority (56%) transcribing more than half of their conversational interactions.

67% of respondents view voice implementation as **"foundational"** to their products and strategy.

2 The Future of Voice AI is Bright

Companies are making space in their budgets—sometimes even raising their budgets—for the explicit purpose of integrating voice technology into their tech stacks. In fact, more than half of organizations now expect to realize substantial benefits from weaving voice technology across the customer journey. This signals a shift to viewing voice AI as a transformative layer that enhances customer interactions at every stage. The growing emphasis on compliance, security, and accessibility is also driving organizations to invest in cutting-edge voice AI solutions.

84% of respondents plan to increase their budgets for voice technology over the next twelve months.

70% have growing expectations of realizing benefits from integrating voice technology across multiple customer touchpoints.

3 Voice AI Agents: The New Frontier of Customer Service Automation

Classic IVR systems are giving way to more advanced technology. From customer service to sales support, automated calls are no longer limited to simple multiple-choice queries. Rather, Voice AI Agents allow businesses to automate more complex tasks like answering dynamically customer FAQs and even taking detailed orders. As a result, the vast majority of organizations are integrating voice agents into their current customer service systems.

Soon enough, adopting Voice AI Agents will no longer be seen as being "ahead of the curve." Rather, it will be the norm. Organizations that fail to implement them will risk falling behind as human-like AI becomes the expected standard for automating customer interactions.

Despite **80%** of surveyed organizations using some form of voice agent—ranging from traditional IVR systems to AI-powered solutions—only **21%** are "very satisfied" with existing technology, highlighting the need for more advanced AI capabilities that deliver human-like responsiveness.

15% of organizations are already actively developing voice AI agents, and a majority of them (**98%**) plan to have them in production within the next year.

Over **50%** of organizations already use traditional voice agents for customer service and task automation (i.e. answering FAQs, sales support, and order-taking) and believe it's the most transformative use case for voice AI agents.

4 Flexibility Will Drive Voice AI Adoption

The market for voice AI technology teems with options from various tech companies. When an organization decides to adopt voice AI, they must therefore choose which model(s) to adopt. Unsurprisingly, organizations care more about being able to fine-tune and customize their AI, signaling the need to adapt voice AI models to specific industries or terminologies for higher accuracy across enterprise use cases. Likewise, they prefer to purchase AI models that are compatible with their existing systems.

46% of respondents said the ability to fine-tune speech models would lead to greater voice AI adoption.

72% cite performance quality (voice quality, conversational flow, etc.) as the most critical barrier for deploying voice AI agents.

65% cited compatibility with existing AI systems as a key factor in vendor selection.

5 Compliance & Accessibility Remain Key Drivers for Use of Voice AI

Both compliance and accessibility are primary motivators for voice AI adoption. Voice interfaces naturally extend accessibility to individuals who cannot easily use digital tools or who favor speaking in non-native languages over typing. Voice AI simultaneously broadens customer reach and diversifies potential workforce talent by removing traditional interaction barriers. As accessibility standards and compliance frameworks evolve, many organizations view voice technology as both a competitive differentiator and a compliance necessity, positioning it at the intersection of innovation and regulatory adherence.

More than half of respondents see voice AI compliance as a primary motivator behind implementation.

86% see voice AI as a key driver for more accessible and inclusive customer experiences.

The Voice Landscape

The breadth of voice AI technology spans widely—ranging from text-to-speech (TTS) and speech-to-text (STT) to cutting-edge human-like voice AI agents that react and respond to users in real-time. Additionally, voice AI ought to work well in noisy environments, have multilingual capabilities, and scale with the size and demands of various user bases.

The goal of this updated annual report is to provide a clearer understanding of the current capabilities of voice AI and the rate at which it's evolving. In this section, we'll explore how organizations are incorporating this technology into their operations.

How Widespread is Voice AI?

A vast majority—97% of respondents—actively use some form of voice technology, such as legacy voice agents, speech-to-text (transcription), and text-to-speech (voice synthesis), in their operations.

The remaining 3% of respondents stated that they have budget constraints, lack of understanding of voice technology, or limited personnel resources.

Among the organizations who do use voice tech, an impressive 87 percent have built at least some of their solutions in-house. The remaining 13 percent relied entirely on out-of-the-box technology from external vendors.

Thus, we conclude that the vast majority of businesses across various domains experience the benefits of voice AI so clearly that they are willing to spend both money and personnel resources to implement and integrate such solutions.

Current use of voice technology

97%   3%

Currently use
voice technology

■ Yes   ■ No   ■ I don't know
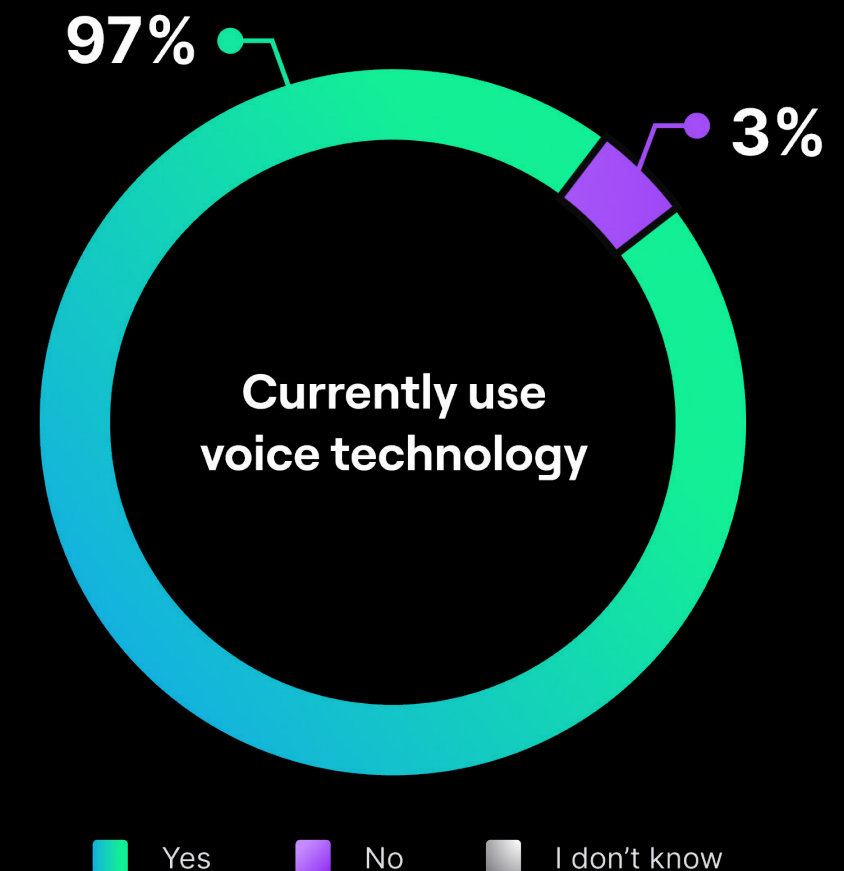
If yes, did you build your own voice
solution or buy one?

Built parts of our solution in-house
and purchased parts of our
solution from external vendors — 46%

Built our solution fully in-house — 41%

Purchased all voice-related
solutions from external vendor(s) — 13%

Current Uses for Voice Tech

As companies have become more accustomed to voice technology over the past few years, clear frontrunner applications have emerged. Summarizing meetings and automating customer service remain neck-and-neck as the top two most common ways that businesses believe voice tech will be transformative.

Transcription is now table stakes, with more than 40% of respondents noting that employee coaching, compliance monitoring, and customer experience analysis are transformative use cases.

Meanwhile, more agentic use cases are emerging as the next frontier. The most significant impact? Automating customer interactions. 52% of respondents believe this is voice AI's most transformative application. One-in-three respondents believe voice AI for recruiting efforts is a transformative use case, and 15% of respondents see AI team members as representing the future of voice technology applications.

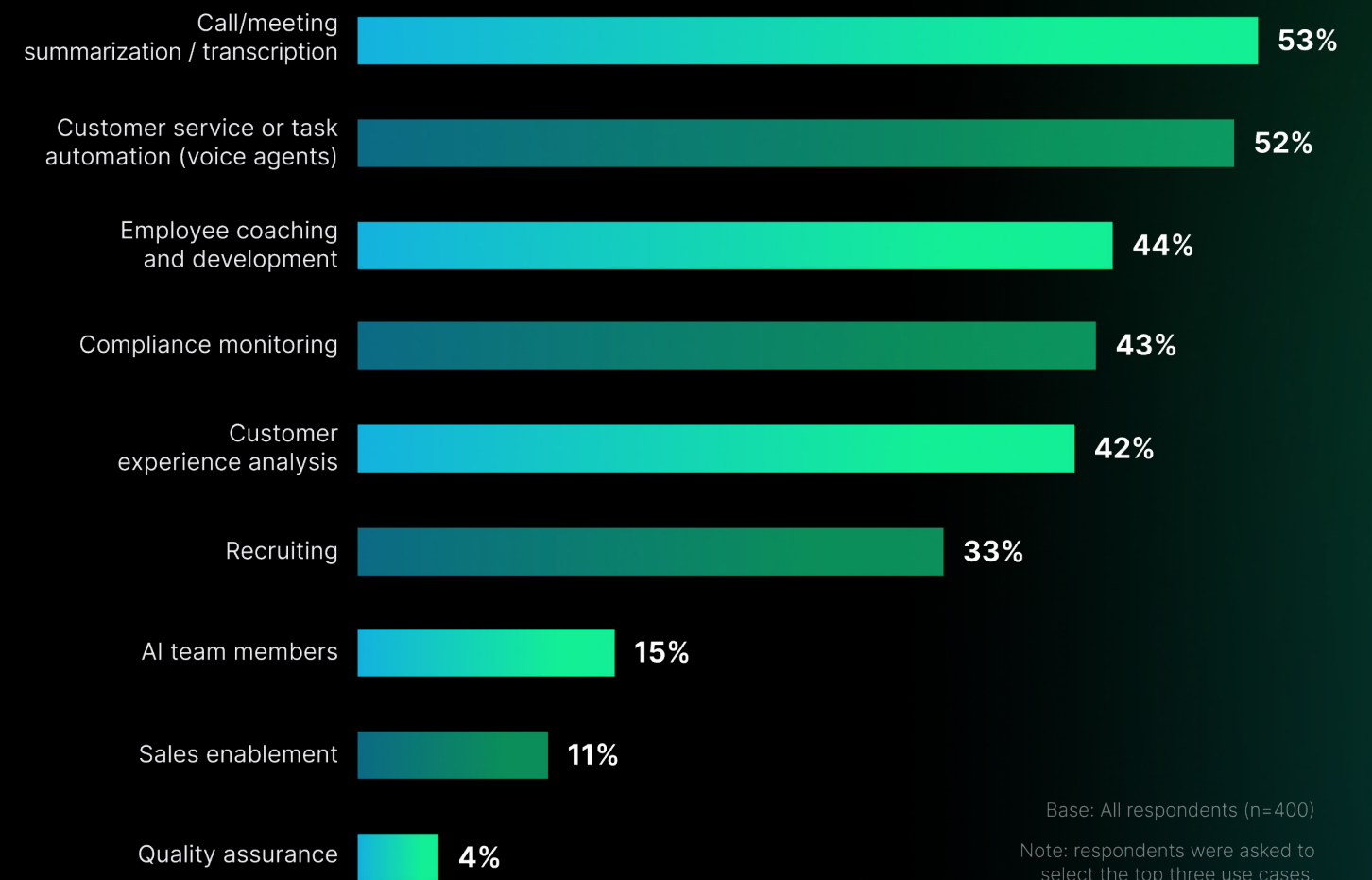# From IVR to IVA: The Importance of Automating Customer Service

A majority of organizations (52%) believe "customer service or task automation" to be the most transformative use case for voice technology. But there's more to this insight than meets the eye.

Respondents see tasks that are already automated (or automatable) through more traditional means like IVRs as compelling reasons to implement a voice agent: 61% cite complete transactions or checkout, 59% point to answering FAQs, and 48% to appointment scheduling (referenced further down in Figure 12).

- **The short-term win:** In essence, these organizations aim to shift from an IVR-like experience to more of an IVA experience. The fact that these use-cases have long been candidates for automation makes them low-hanging fruit, quick wins even.

- **The long-term win:** Agentic-powered interactions that can potentially drive resolution of unclear or thorny problems are the long-term opportunity.

Nevertheless, in regards to the tougher "long-term win," there are some organizations that consider such use cases fair game today: 30% of survey respondents saw initiating/resolving a service request as a compelling reason to dive into the world of voice AI agents.

Most Transformative Use Cases for Voice Technology

Call/meeting summarization / transcription — **53%**
Customer service or task automation (voice agents) — **52%**
Employee coaching and development — **44%**
Compliance monitoring — **43%**
Customer experience analysis — **42%**
Recruiting — **33%**
AI team members — **15%**
Sales enablement — **11%**
Quality assurance — **4%**

Base: All respondents (n=400)

Note: respondents were asked to select the top three use cases.

As usual, the breadth of these responses highlights the diverse needs that voice AI addresses. The fact that the top six use cases are utilized by between 1-in-3 and 1-in-2 respondents showcases just how much voice technology satisfies the varied demands of businesses across multiple domains.

Adoption of Voice AI Technology

Organizations adopt voice AI for a variety of reasons, from reducing costs to improving efficiency. As expected, nearly every single one of these businesses achieved their operational goals in a satisfactory way. These stats highlight the current effectiveness of voice AI. Since such technology is still evolving, we can only imagine how adopting it will benefit businesses in the future.

Let's now examine the motivations behind implementing voice technology and the level of satisfaction with its impact.

Why People Implement Voice AI

Simply put, organizations implement voice AI because it adds value to their businesses in multiple ways. Two-thirds of organizations say that **voice AI is foundational to their products,** while one-third say that this technology reduces costs.

Furthermore, 45% of respondents use voice AI to promote greater operational efficiency, while 35% use it to promote productivity among salespeople and employees.

Top Reasons for Implementing Voice-enabled Solutions

Reason	Percentage
It is foundational to our product(s)	67%
Compliance with regulatory mandates	56%
Promote operational efficiencies	45%
Reduce costs	36%
Improve salesperson / employee productivity	35%
Increase revenues	32%
Identify new business opportunities	22%
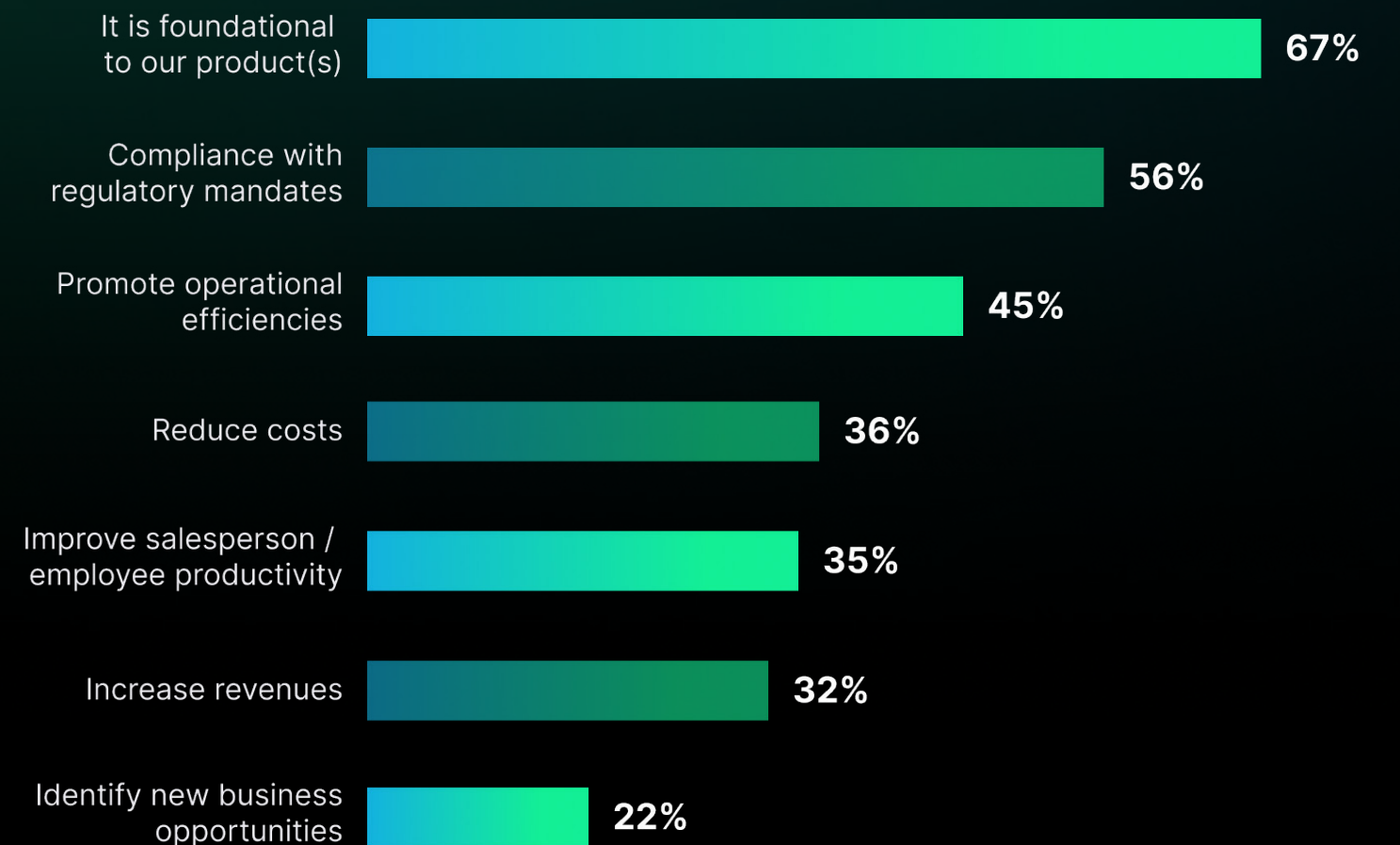
Base: Respondents who selected "Yes" at Q1 (n=389)
Note: respondents were asked to select the top three reasons.

Regulatory Compliance and Accessibility Fuel Voice AI Adoption

Accessibility—which is often directly linked to regulatory compliance—is now a strong driver for voice AI. Intuitively, this focus on accessibility makes sense, as the voice modality opens up interactions to those who struggle to use traditional digital tools. Furthermore, voice AI agents improve accessibility for those who are more comfortable speaking in a non-native language rather than writing in it. Consequently, improved accessibility and inclusion has emerged as the biggest customer experience (CX) improvement brands expect from voice AI.

All that said, accessibility is about much more than CX improvements—there are actual top- and bottom-line improvements that come from improved accessibility. On the revenue-generating side, brands with accessible experiences:

- **Expand their customer reach –** Accessibility ensures that more people, including those with disabilities, can engage with a brand's products, services, and content, broadening the potential customer base. The "curb cut effect" shows that it is not just those with disabilities that see improved experiences from greater accessibility.

- **Increase revenue opportunities –** By removing barriers, companies make it easier for all consumers to engage, purchase, and advocate for their products, driving higher conversion rates and revenue growth.

On the cost-savings side, accessible experiences drive:

- **More efficient development –** Designing with accessibility in mind from the start prevents costly retrofits and reduces technical debt associated with patching inaccessible features later.

- **Lower support costs –** More accessible experiences lead to fewer customer frustrations, reducing the need for live agent support and lowering contact center and support ticket volumes.

On the compliance front, **56% of respondents cite compliance with regulatory mandates as a primary driver for voice AI implementation**—ranking it the second most important factor.

Companies are exploring compliant voice AI solutions to simultaneously improve accessibility, enhance customer experience, and create operational efficiencies that deliver measurable ROI beyond mere regulatory checkbox-ticking.

For example, healthcare providers are already leveraging voice-enabled applications to automatically transcribe and document patient interactions while maintaining HIPAA compliance, streamlining clinical workflows and improving patient experiences.

This approach transforms voice AI from a compliance necessity into a strategic advantage with tangible benefits across the enterprise. By starting with regulatory requirements and expanding to broader applications, organizations are finding that compliance-driven voice AI becomes a foundation for wider digital transformation initiatives.

Based on the graph above, we can conclude that the use of voice technology is expanding and becoming foundational. This continued momentum points to the growing importance of leveraging voice technologies across both employee and customer experiences.

The Expected ROI of Voice AI adoption

Beyond use cases, it's important to consider ROI. What benefits do these companies expect to see from implementing voice technology? What would drive them to further integrate voice AI into their workflows and products?

First, let's examine what businesses expect the ROI to be from the customer side. The chart below indicates that the majority of businesses expect voice AI to yield 24/7 availability for customers, enhanced engagement, and improved accessibility and inclusion. Other factors such as improved customer insights and cost efficiency are expected as well.
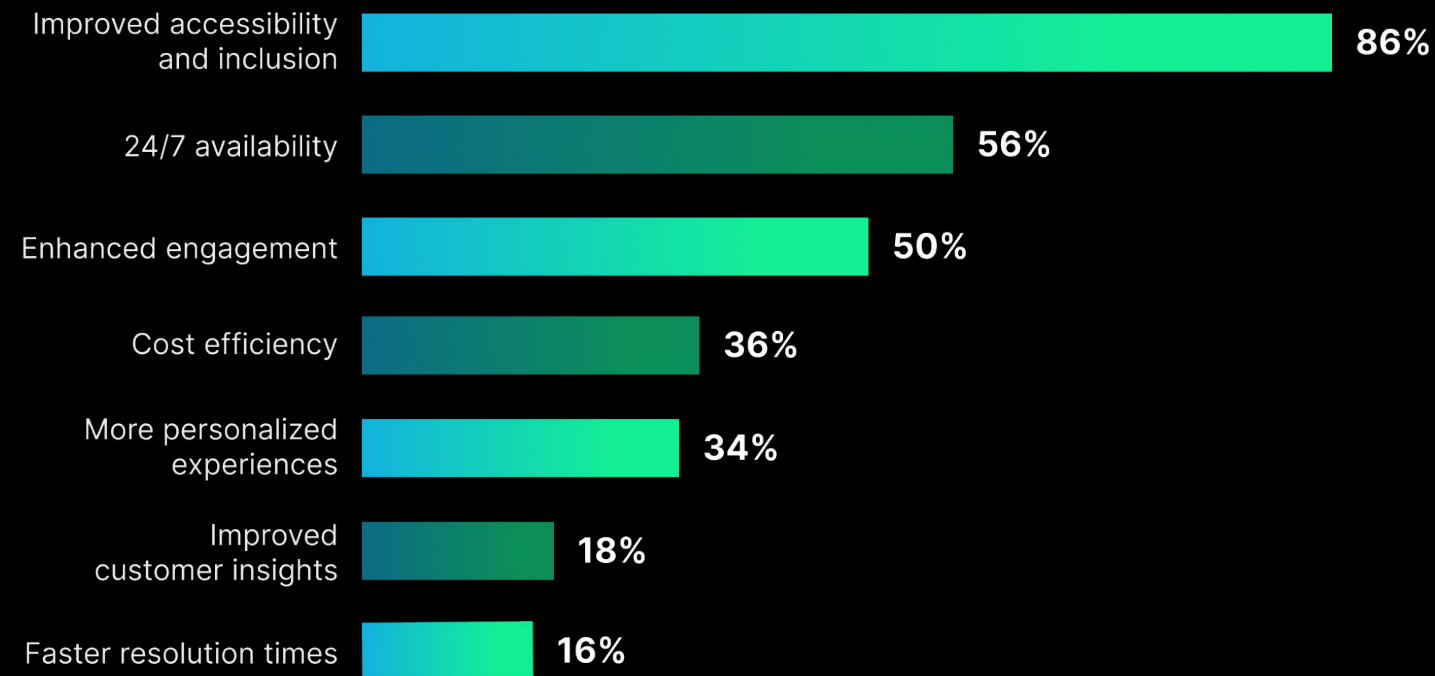
Let's now examine what businesses expect ROI to be from the employee side. Again, the majority of businesses expect increased accessibility. Likewise, nearly three-fourths of respondents expect more streamlined workflows, while nearly half of these organizations expect boosted productivity and enhanced training.

Lastly, roughly a quarter of survey respondents believe voice AI will reduce workload and turnover in their organization. This underscores the broad potential voice AI holds for improving various aspects of workplace experience.

Top Ways Voice Technology is Expected to Enhance CX

Improved accessibility and inclusion	86%
24/7 availability	56%
Enhanced engagement	50%
Cost efficiency	36%
More personalized experiences	34%
Improved customer insights	18%
Faster resolution times	16%

Base: All respondents (n=400)
Note: respondents were asked to select the top three answers.

Top Ways Voice Technology is Expected to Enhance Employee Experience

Streamlined workflows	74%
Increased accessibility	50%
Boosted productivity	48%
Enhanced training and development	47%
Improved collaboration	32%
Reduced workload	25%
Reduced turnover	22%
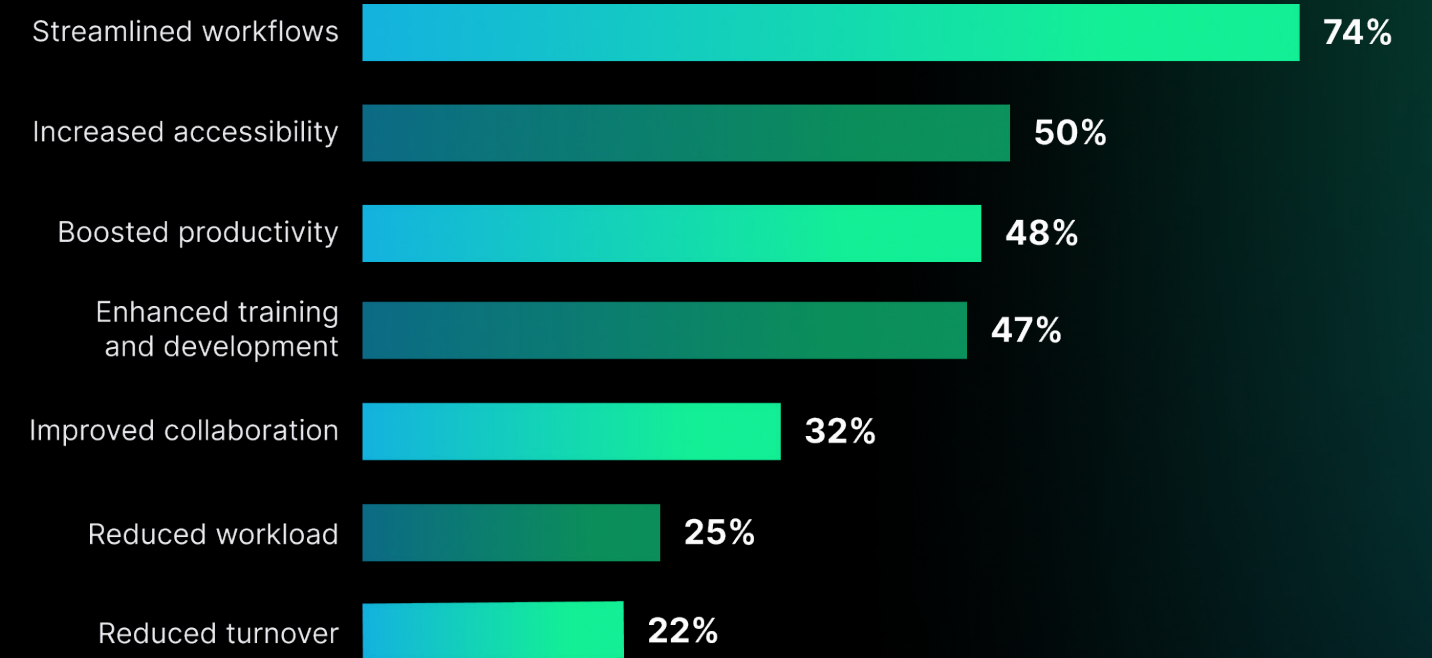
Base: All respondents (n=400)
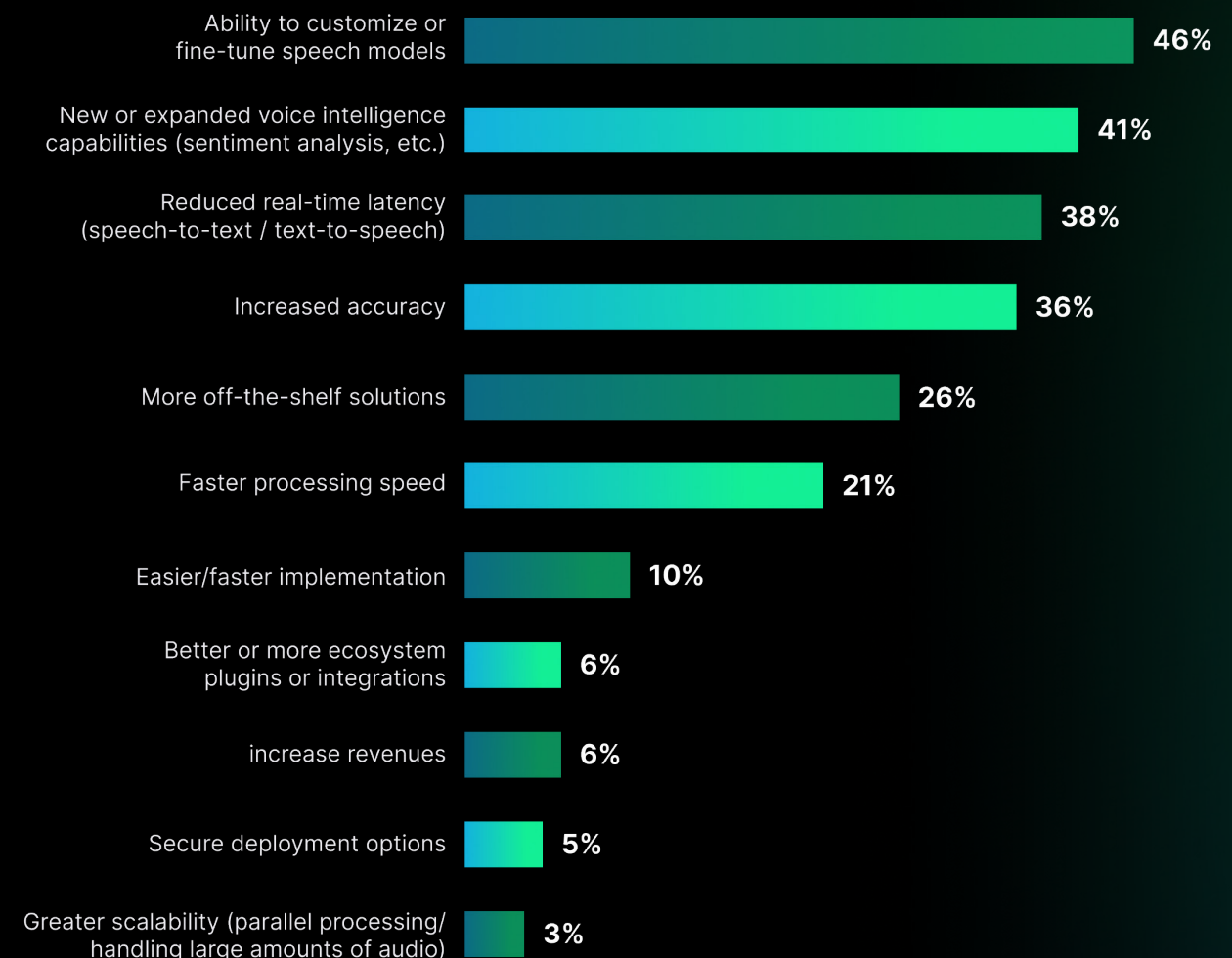Note: respondents were asked to select the top three answers.

From General LLMs to Fine-Tuned Voice AI

The voice AI market is rapidly shifting away from reliance on a few dominant platforms toward a more open, customizable landscape. The rise of low-cost, capable open source language models, like DeepSeek, has reinforced the idea that companies can "own their own" AI solutions rather than being locked into a handful of super-platforms.

Consequently, we are now living in an LLM-agnostic world where businesses are prioritizing flexibility, cost-efficiency, and control over their AI models. Instead of being forced into a one-size-fits-all approach, organizations now have the freedom to build and fine-tune voice AI solutions that meet their specific needs using a variety of best–in-class building blocks.

Our survey highlights why this shift is happening. When asked what would encourage greater adoption of voice AI, 46% of respondents pointed to the ability to fine-tune speech models—a clear sign that many businesses require voice technology that adapts to domain-specific terminology for use cases like healthcare and finance. Furthermore, 41% prioritized expanding voice intelligence capabilities (e.g. sentiment analysis and topic detection), highlighting a growing demand for automating insight extraction from voice data. This trend reveals organizations' eagerness to leverage AI-powered voice analytics to uncover actionable patterns and automate processes–particularly crucial for contact centers where voice interactions generate massive untapped data repositories.

Primary Motivation for Implementing Voice Technology or Switching Providers

Motivation	Percentage
Ability to customize or fine-tune speech models	46%
New or expanded voice intelligence capabilities (sentiment analysis, etc.)	41%
Reduced real-time latency (speech-to-text / text-to-speech)	38%
Increased accuracy	36%
More off-the-shelf solutions	26%
Faster processing speed	21%
Easier/faster implementation	10%
Better or more ecosystem plugins or integrations	6%
increase revenues	6%
Secure deployment options	5%
Greater scalability (parallel processing/ handling large amounts of audio)	3%

Base: All respondents (n=400)
Note: respondents were asked to select the top three reason

If organizations can truly realize benefits such as lower technology costs, increased revenue, faster processing speeds, and secure deployment options from AI models that they can customize or fine-tune themselves, the question shifts from whether they will adopt the technology to how soon.

The Year of the Voice AI Agents

While the graphs above mainly focus on voice technology as a whole, it's important to narrow our scope to identify which aspects of voice technology will become the most prominent in the near future. The data collected strongly suggest that voice AI agents are gaining momentum and here to stay.

A voice AI agent is a software system powered by artificial intelligence that can understand and respond to spoken language. It is designed to automate communication processes, such as customer service or virtual assistance, enhancing efficiency and user experience. By leveraging natural language processing (NLP) and machine learning it can engage in dynamic, responsive, and context-aware conversations with users.

Let's explore why.

Enterprises Are Eager to Move from Legacy IVR to AI-Powered Voice Agents

Eighty percent of surveyed organizations have deployed some form of voice agent, many relying on traditional IVR systems. However, these rigid, robotic-sounding solutions are quickly becoming obsolete. Companies are shifting toward AI-powered voice agents that deliver human-like responsiveness and dynamically handle complex interactions.

This increased investment highlights a growing focus on improving voice technology. With only 21% of respondents reporting they are "very satisfied" with their current voice solutions, organizations are looking beyond traditional approaches, signaling rising interest in AI-powered voice agents.

As a result, companies are allocating more resources—both financial and personnel—to replace legacy systems with AI-driven solutions, develop voice AI agents from the ground up, or scale the AI they already have in place. The case for implementing voice AI agents is compelling: voice AI technology has matured from simple voice recognition to sophisticated agents powered by low-latency transcription, high-fidelity text-to-speech (TTS), and advanced large language models (LLMs). These voice AI agents now deliver human-like conversations to personalize customer experiences, while providing the scalability and consistency that traditional IVRs cannot match. Voice AI agents allow organizations to maintain high service standards even during peak periods without staffing challenges and quality inconsistencies, simultaneously reducing the burden on human teams and driving significant cost savings.

In fact, 15% of organizations are already actively developing voice AI agents, and a majority of them (98%) plan to have them in production within the next year (see chart).

These accelerated adoption timelines, combined with growing investment trends in voice AI, underscore why 2025 is the year enterprises embrace voice AI agents.

Current Adoption of Voice Agents, including Legacy IVR

80% **15%** **5%**
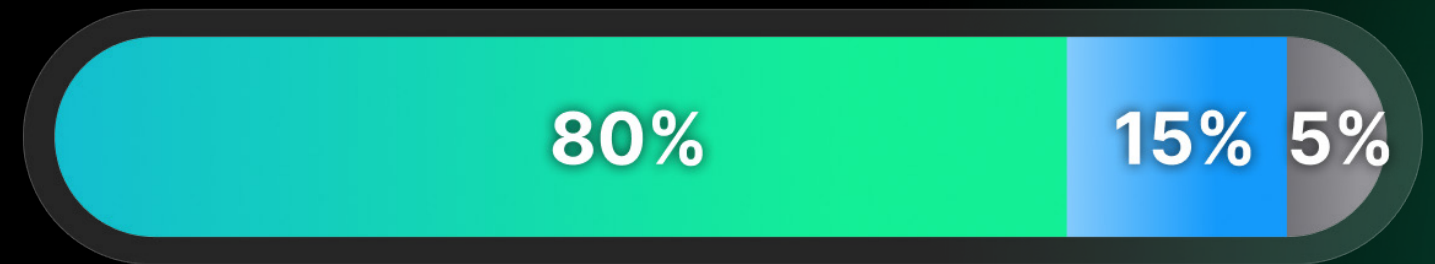
- 🟢 Already using traditional voice agent
- 🔵 Currently developing voice AI agents
- ⬜ Haven't started developing voice AI agents yet
- 🟣 Don't have plans to develop voice AI agents
- 🔴 Voice AI agents aren't applicable to us

Projected Timeline for Deploying New Voice AI Agents

7% **65%** **26%** **2%**
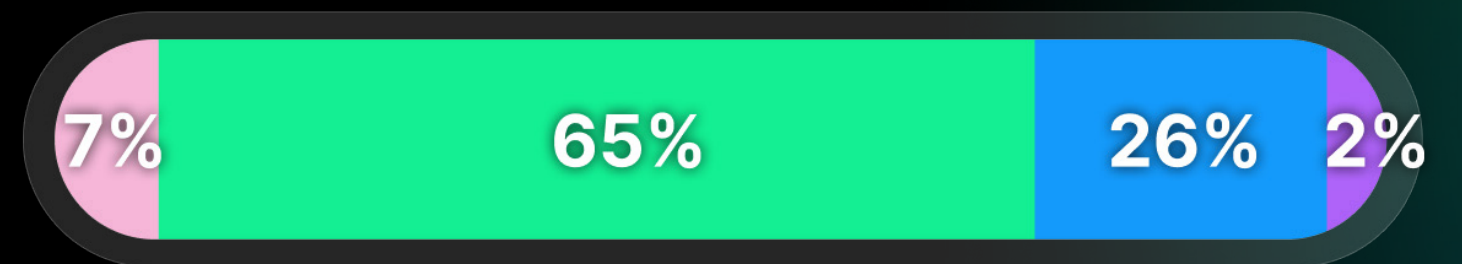
- 🌸 0-3 months
- 🟢 3-6 months
- 🔵 6-12 months
- 🟣 A year or more
- 🔴 N/A

Base: All Respondents (n=400)

How Satisfied Are Businesses with Voice AI Agent Technology?

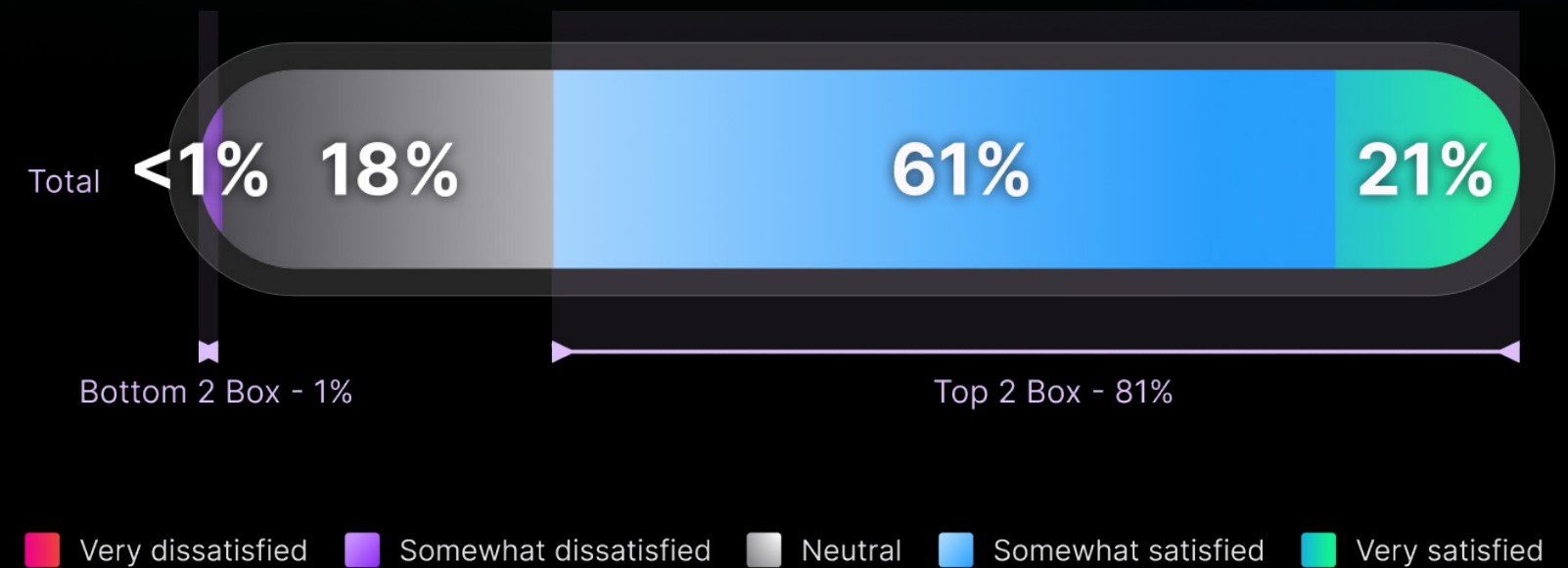The survey data reveals a mixed sentiment: while 81% of organizations express some level of satisfaction with their current voice agent technology, the majority (61%) are only "somewhat satisfied," indicating clear room for improvement.

This chart highlights this nuance—only 21% of organizations report being "very satisfied," while dissatisfaction remains low. This suggests that while current voice agents are functional, they fall short of expectations in key areas.

At the same time, substantial voice AI budget expansions suggest that companies aren't content with the status quo. Rather than maintaining legacy IVR and basic conversational AI systems, they are investing in next-generation capabilities to address critical gaps in conversational quality (e.g., latency, voice clarity) and integration with broader AI ecosystems.

Satisfaction-level with Current Voice Agent Technology

5 point scale: (1=Very Dissatisfied and 5=Very Satisfied)

Total **<1% 18% 61% 21%**

Bottom 2 Box - 1%   Top 2 Box - 81%

■ Very dissatisfied   ■ Somewhat dissatisfied   ■ Neutral   ■ Somewhat satisfied   ■ Very satisfied

Base: Respondents (n=389)
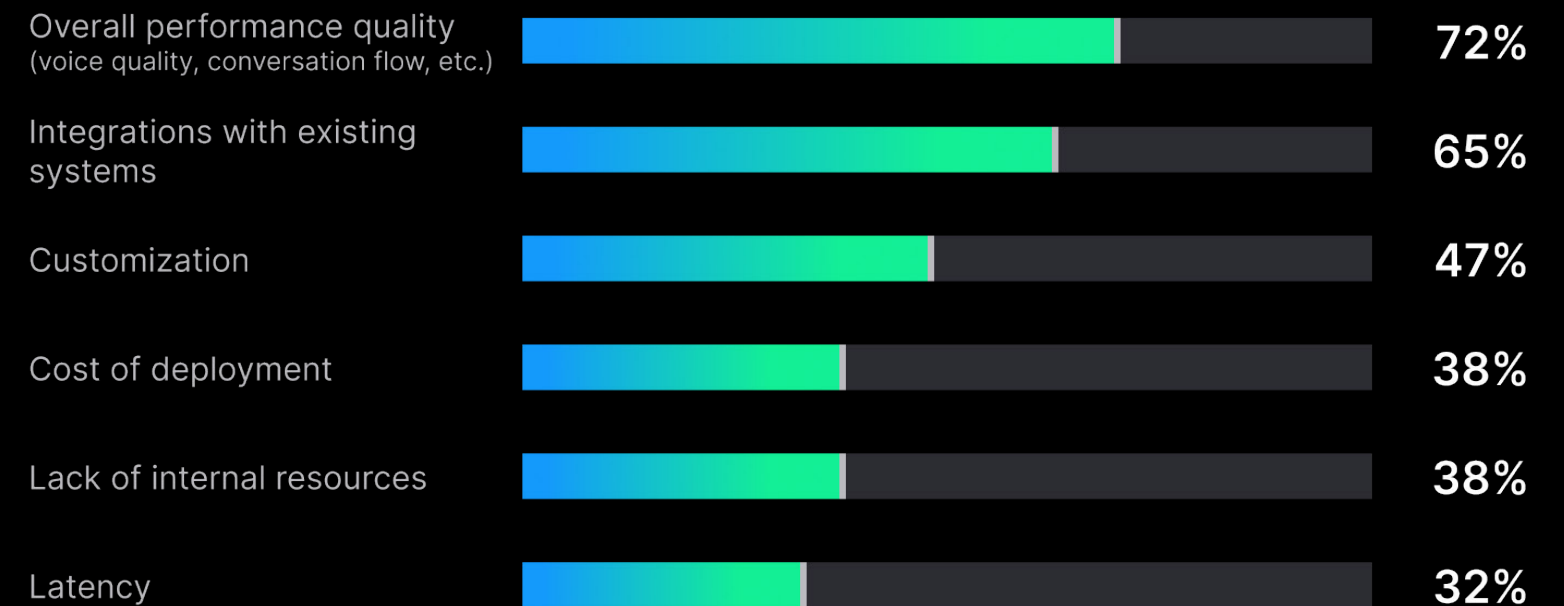
Barriers to Entry

While the technology in the AI world is cutting-edge and incredibly capable, it is not perfect. And these imperfections often transform into hurdles that both the AI's developers and their enterprise customers must overcome.

The graph below unveils the fact that performance and integrations slow down implementations. More specifically:

Solution quality—which includes **voice clarity, conversational flow, and overall performance—**was identified by 72% of respondents as a major hurdle. This finding highlights a significant level of dissatisfaction among enterprise customers with their current voice technology solutions. Many remain dependent on outdated systems that fall far short of the performance benchmarks now expected in the GenAI era, where human-like, responsive voice agents are becoming the
new standard.

65% of respondents emphasized **compatibility with existing AI systems** as a barrier to voice agent development. Furthermore, 60% cited **integration** as one of the biggest challenges to voice AI adoption. This suggests that once organizations take on the real work of implementing a voice solution, integration obstacles (e.g., data transformation, security constraints, internal stakeholder buy-in) can suddenly loom large. Solutions that ease the pain of deployment and integrations can help accelerate adoption.

Common Barriers to Developing and Implementing Voice AI Agents

Overall performance quality
(voice quality, conversation flow, etc.)
72%

Integrations with existing systems
65%

Customization
47%

Cost of deployment
38%

Lack of internal resources
38%

Latency
32%

Note: respondents were asked to select the top three answers.

Interestingly, cost does not seem to be a strong barrier to initial entry into the technology, as only 38% feel it is a blocker to initial deployment. We interpret this finding to mean that organizations are prioritizing quality over financial costs when it comes to voice AI. That is, these companies will not necessarily just go for the cheapest option—they're willing to pay higher prices if the technology guarantees their desired results.

Voice AI Agent Use Cases

Voice AI agents are interactive systems that process and respond to human speech in real time, enabling natural and dynamic human-machine interactions.

Behind every effective voice AI agent is a sophisticated technical architecture. Modern implementations rely on highly accurate Automatic Speech Recognition (ASR) to convert spoken words into text. The system's intelligence—its cognitive architecture—comes from a complex integration of LLMs, RAG systems and Knowledge Graphs, which enable understanding and response generation. Finally, Text-to-Speech (TTS) converts these responses into natural-sounding voice output.

To perform effectively in real-world scenarios, these systems must maintain contextual awareness across conversations, support robust multilingual capabilities, and gracefully handle interruptions and background noise. For many implementations, telephony integration ensures universal access without requiring specialized apps or interfaces.
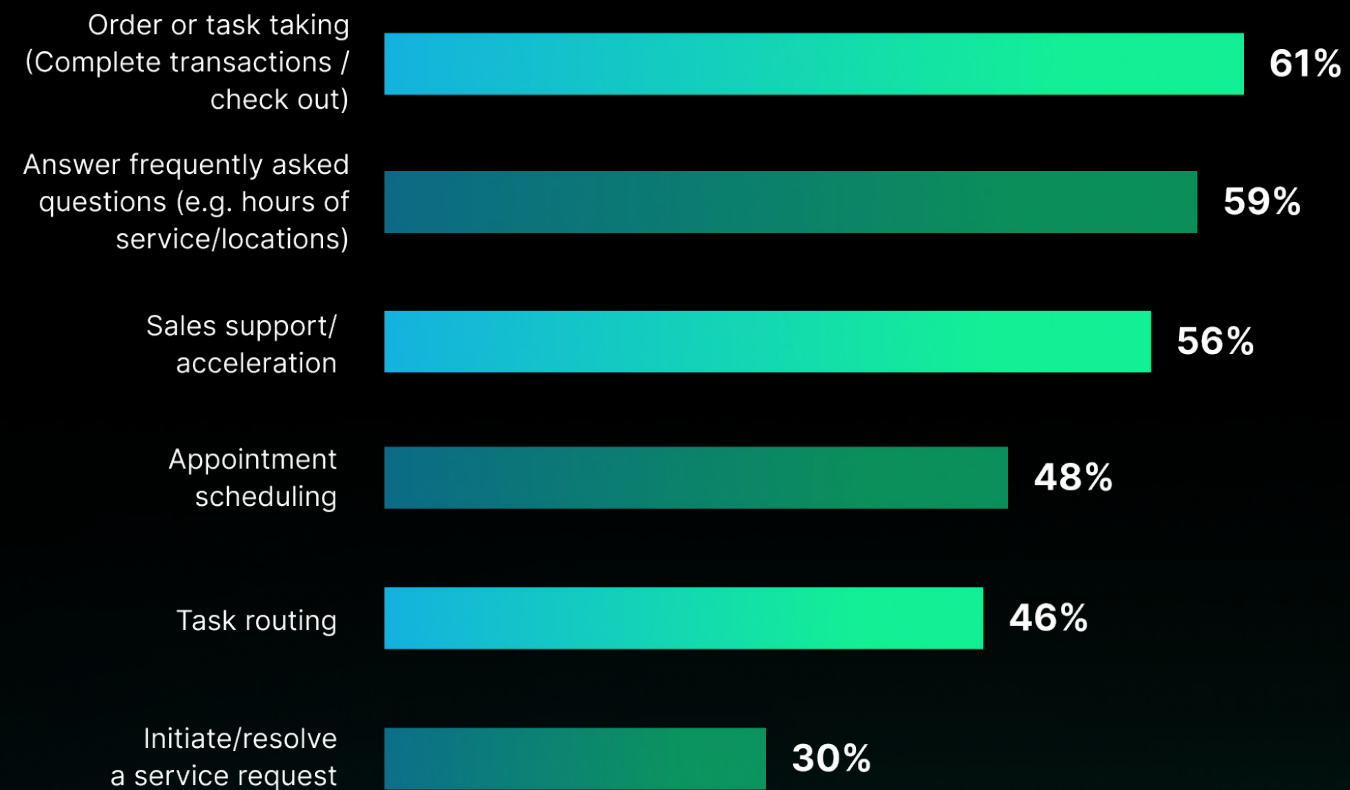
The evolution of voice AI has transformed these systems from basic command-and-response interactions into sophisticated agents capable of understanding context and managing nuanced customer interactions without human intervention. They support a diverse range of use cases and perform various tasks across industries, from handling customer inquiries and scheduling appointments to providing detailed product information and initiating service requests.

Respondents identified the most compelling use cases for implementing voice AI agents as order and task management, answering FAQs, and accelerating sales and support (see chart below). The breadth of these applications highlights the versatility of voice AI agents. Below, we showcase several examples.

Examples of voice AI agent applications across industries include:

- **Order taking for QSRs (quick-service restaurants):** Quickly process food orders at restaurant drive-thrus, leading to faster service times and improved order accuracy.

- **Retail FAQ handling:** Provide instant answers about product information, availability, warranty terms, and return policies, offering 24/7 support and improving customer satisfaction.

- **Policy quoting for insurance providers:** Collect customer requirements, explain coverage options, provide preliminary quotes and qualify leads before connecting prospects with human agents.

- **Medical appointment scheduling:** Book, reschedule, and remind patients of upcoming appointments, while verifying insurance coverage, reducing no-show rates and improving overall patient experience.

Most Compelling Use Cases for Implementing Voice AI Agents

Order or task taking
(Complete transactions /
check out) — **61%**

Answer frequently asked
questions (e.g. hours of
service/locations) — **59%**

Sales support/
acceleration — **56%**

Appointment
scheduling — **48%**

Task routing — **46%**

Initiate/resolve
a service request — **30%**

Primary Voice Agent Use Cases Across Organizations

Total

| 39% | 35% | 16% | 10% |

- Customer service / support automation
- Sales or lead generation
- Appointment scheduling
- Order or task taking
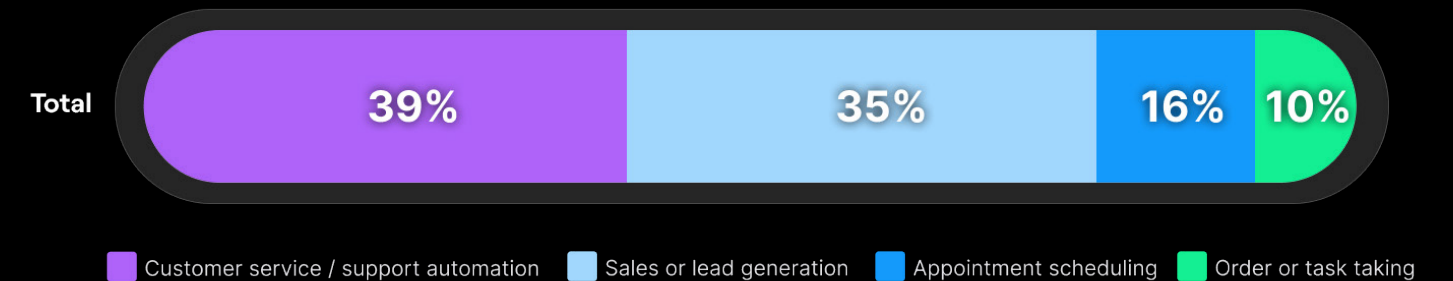
Furthermore, we see that the two most relevant use cases for voice agents—utilized by 39 percent of respondents—are "Customer service/support automation" as well as "Sales or Lead Generation. In fact, the most transformative use cases for voice AI agents included completing transactions, conducting customer check-outs, answering FAQs, supporting the sales team, and scheduling appointments.

In other words, the majority of organizations surveyed are using some form of voice agents (legacy IVR or other) to support their employees' experiences, to optimize their customers' interactions, or both.

The Voice AI Agent Features that Organizations Care Most About

Voice AI agents offer a variety of features, and companies developing them differentiate themselves by optimizing distinct aspects of these capabilities to stand out in the market.
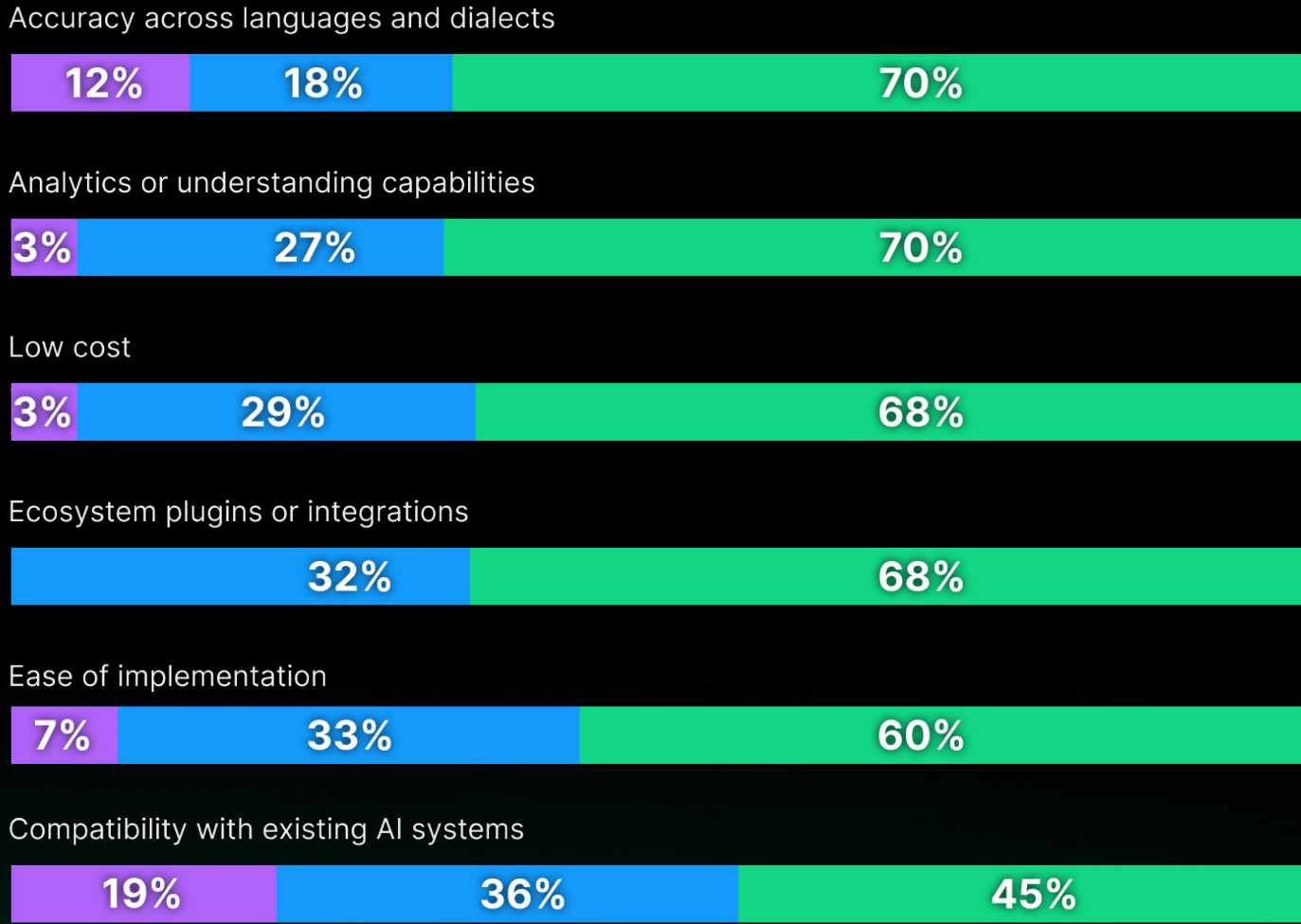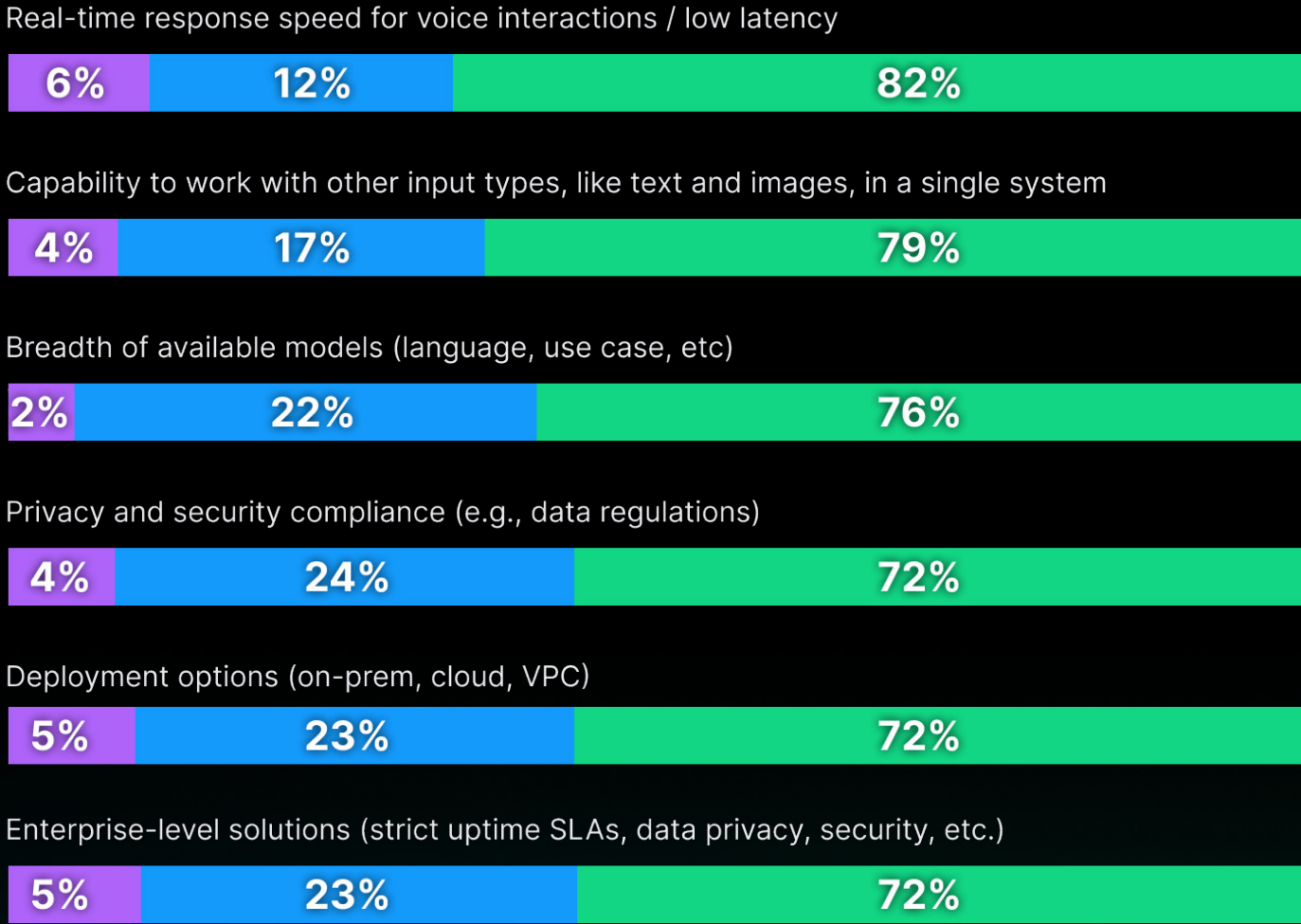
For example, some companies focus primarily on providing their customers with the widest breadth of multilingual models. Meanwhile, others try to optimize their model's latency such that they respond faster than the rest of the AI on the market. When evaluating voice AI agents, organizations may prioritize specific capabilities essential to their implementation strategy, yet the results from this survey reveal a clear hierarchy of needs.

Deploying Human-Like Voice Agents: Features Organizations Should Optimize

- **Low latency** is crucial for building human-like voice agents, and organizations agree. More than four out of every five voice AI agent adopters say that real-time response speed for voice interactions is either "Important" or "Very Important. This overwhelming preference for conversational fluidity outranks all other considerations.

- **Human-like voice quality** is another critical feature. When a voice agent sounds robotic rather than natural, it can lead to customer frustration. Therefore, optimizing for natural-sounding voices directly contributes to improving customer satisfaction.

- **Enterprise-level features** such as data privacy compliance, strong security, and strict uptime SLAs are crucial for organizations using voice agents to support their sales and customer service teams. Respondents share this priority, with 72% rating these capabilities as either "important" or "very important."

The charts on the next page highlight how the perception of "human-like" interaction depends primarily on response timing.

24

Most Important Features When Evaluating Voice AI

Real-time response speed for voice interactions / low latency

| 6% | 12% | 82% |

Capability to work with other input types, like text and images, in a single system

| 4% | 17% | 79% |

Breadth of available models (language, use case, etc)

| 2% | 22% | 76% |

Privacy and security compliance (e.g., data regulations)

| 4% | 24% | 72% |

Deployment options (on-prem, cloud, VPC)

| 5% | 23% | 72% |

Enterprise-level solutions (strict uptime SLAs, data privacy, security, etc.)

| 5% | 23% | 72% |

Accuracy across languages and dialects

| 12% | 18% | 70% |

Analytics or understanding capabilities

| 3% | 27% | 70% |

Low cost

| 3% | 29% | 68% |

Ecosystem plugins or integrations

| 32% | 68% |

Ease of implementation

| 7% | 33% | 60% |

Compatibility with existing AI systems

| 19% | 36% | 45% |

■ Bottom 2 Box   ■ Middle Box   ■ Top 2 Box

5 point scale: (1=Not Important and 5=Very Important)

Base: All respondents (n=400)

25

The Future of Voice AI

So far, we've established three key points: (1) organizations are widely adopting voice AI technology, (2) voice AI agents provide multiple ROI benefits while enhancing accessibility, and (3) companies that are only "somewhat satisfied" with their current voice technology are increasing their budgets to further invest in voice AI.

But where do we go from here?

The future of voice AI looks promising. However, there are a few barriers to entry. In this section, we'll explore not only the positive outlook of Voice AI technology but also what potential challenges lie ahead.

Organizations Plan to Expand Voice AI Usage

84% of respondents plan to increase their budgets for voice technology over the next twelve months, with nearly half indicating a significant rise in spending. This strong investment signals that enterprises not only see value in voice AI today but also anticipate even greater potential in the near future.

Budget Expectations Over the Next 12 Months

Total   **<1%**   **16%**   **38%**   **46%**

Bottom 2 Box - 1%

Top 2 Box - 84%
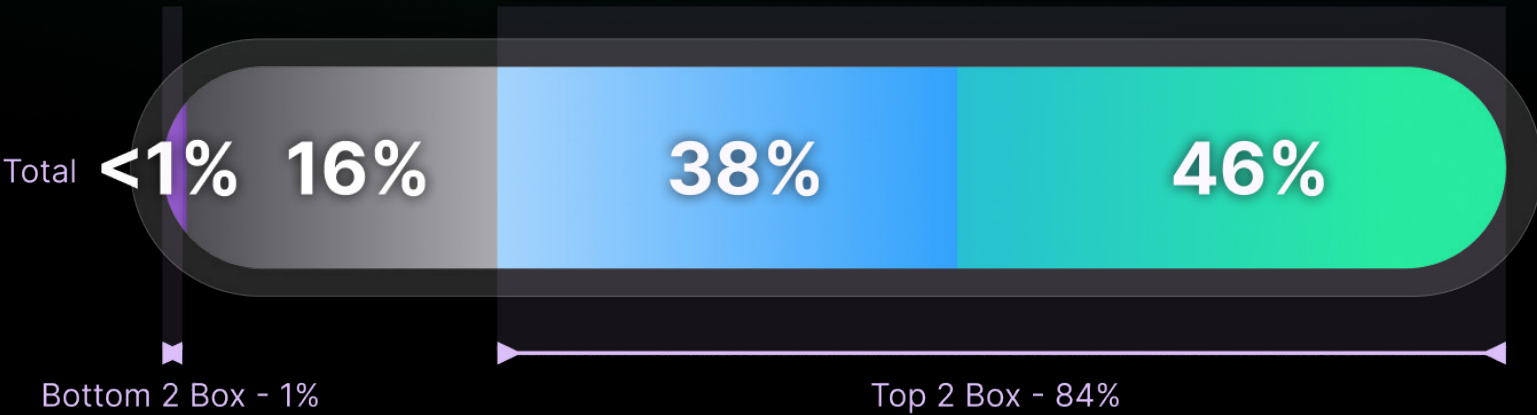
- Decrease significantly
- Decrease slightly
- No changes
- Increase slightly
- Increase significantly

Base: Respondents (n=389)

5 point scale: (1=Decrease Significantly and 5=Increase Significantly)

As businesses increasingly recognize the power of language models to extract insights from conversations—whether between employees or with customers—the demand for accurate, high-quality voice data is expanding. Companies are shifting beyond basic automation to more advanced applications, such as real-time transcription, sentiment analysis, contextual understanding, and AI-driven coaching for customer service and sales teams.

In fact, this graph reveals that 84% of organizations responded that voice technology is important or even fundamental to their overall customer experience (CX) strategies.

The growing emphasis on compliance, security, and accessibility is also driving organizations to invest in solutions that ensure high-fidelity voice capture, reliable transcription, and seamless integration into existing workflows.

Ultimately, this increased investment underscores a broader trend: voice AI is becoming a mission-critical component of enterprise technology stacks–not just for automation but for gaining strategic insights, enhancing customer experiences, and driving revenue growth.

Importance of Voice Technology in Overall CX Strategy

| 14% | 70% | 16% |
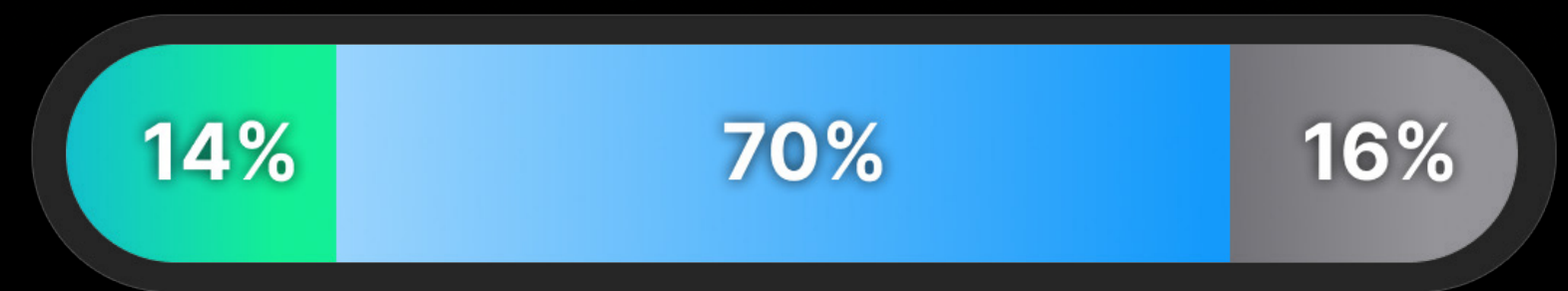
- 🟩 Fundamental (voice-first strategy)
- 🟦 Important (integrated into multiple customer touchpoints)
- ⬜ Supporting specific areas
- 🟪 Minimal (only used occasionally)
- 🟥 Not used

Base: All Respondents (n=400)

Where There's Room to Grow

The 2020s brought us the first effective large language models (LLMs). These models began with a very general knowledge base and a text-only interface. Today, voice AI agents can be trained or fine-tuned on very specific information about an organization's customer base, internal workflows, and more. Furthermore, instead of interacting with these AI models through typing, we can now communicate with them via speech—a faster, more intuitive interface.

In this section, we will use past evidence and trends as a means to predict the future. Synthesizing these technological milestones with the data we received in our survey will give us a clear prediction of the state of voice AI in the coming years.

Key Trends

- **Improved Natural Language Understanding:** LLMs have progressively become better at understanding context, nuances, and user intent in text-based communication, making them more capable when applied to voice interactions.

- **Enhanced Conversational Abilities:** With advancements in LLMs, voice AI agents are becoming increasingly proficient in maintaining multi-turn conversations with users, handling interruptions, follow-ups, and complex queries more naturally.

- **Context Awareness and Personalization:** Voice AI agents are evolving to use user data to offer personalized, contextually relevant responses. This has been greatly enabled by more sophisticated training and fine-tuning techniques in LLMs.

- **Real-time Processing and Responsiveness:** Voice AI agents are increasingly designed to offer near-instant responses, creating a seamless experience for users. This has been achieved through better optimization of LLMs for speed and efficiency.

- **Speech Recognition and Synthesis Improvements:** There has been a significant leap in speech recognition and text-to-speech capabilities, allowing voice AI agents to understand a wider range of accents, languages, and speech patterns while sounding more human-like.

- **Use of Emotion and Tone in Responses:** Newer voice AI agents can modulate tone, pitch, and delivery style based on context, improving user engagement by creating more emotionally aware responses.

- **Cross-domain Knowledge Application:** Voice AI agents are becoming proficient at handling tasks across a broad range of domains, from customer service to healthcare, finance, and entertainment, due to the versatility and scalability of advanced LLMs.

- **Privacy and Ethical Considerations:** As voice AI systems interact more with users in personal contexts, there's a heightened focus on ethical guidelines, data privacy, and transparency regarding how voice interactions are processed and stored.

- **Self-improving Models:** Some voice AI agents are starting to leverage continuous learning, adapting based on real-time user interactions and feedback, improving performance and accuracy over time.

Desired Areas of Growth

For the most part, organizations perceive sales use cases as lagging behind service capabilities.

While 35% of brands identify sales and lead generation as their primary use case (as mentioned above), only 11% see sales enablement as a transformative application of voice AI (also mentioned above). This gap suggests that, while some brands are investing in this area, many have yet to fully grasp voice AI's potential to drive sales and revenue growth. These data present a clear opportunity to educate enterprises on how voice technology can be leveraged more effectively for sales enablement.

Still, when asked directly how voice technology could contribute to revenue growth, a majority of respondents selected upsell and cross-sell opportunities (55%). Interestingly, "data-driven insights" also garnered majority support (55%), highlighting a glaring need for brands and enterprises to share insights across departments. Sharing data-driven insights includes business intelligence use cases that take the form of a conversation using natural language to query and access tools that aggregate and analyze data.

Expected Revenue Growth Contributions from Voice Technology

Upsell and cross-sell opportunities	55%
Data-driven insights	55%
Enhanced operational efficiency	45%
Improved customer retention	42%
Optimized lead conversion	41%
24/7 availability	31%
Access to new customers egments	28%

Note: respondents were asked to select the top three answers.

30

Deepgram's Vision for Voice AI

Deepgram is the leading voice AI platform for developers building speech-to-text (STT), text-to-speech (TTS), and full speech-to-speech (STS) offerings. Over 200,000 developers build with Deepgram's voice-native foundational models— accessed through cloud APIs or as self-hosted / on-premises APIs—due to our unmatched accuracy, low latency, and pricing.

Deepgram CEO Scott Stephenson lays out the landscape as follows:

> *"2024 was a stellar year for Deepgram, as our traction is accelerating and our long-term vision of empowering developers to build voice AI with human-like accuracy, human-like expressivity, and human-like latency is materializing."*

Stephenson continues:

> *"Our product strategy from founding has been to focus on deep-tech first."*

His statement rings true as Deepgram has worked on building 3-factor automated model adaptation and hosting models with efficient and unrestricted hot-swapping. Furthermore, Deepgram has achieved extreme compression on latent space models (LSMs) as well as symmetrical delivery across public cloud, private cloud, or on-premises.

These technological milestones uniquely position Deepgram to

> *"succeed in the $50B market for voice AI agents in demanding environments requiring exceptional accuracy, lowest COGS, highest model adaptability, and lowest latency,"*
> – Scott Stephenson, CEO & Co-founder, Deepgram

Through continued innovation, Deepgram expects to end 2025 as the industry's only end-to-end speech-to-speech solution built to solve the four critical challenges of enterprise-ready voice AI: (1) Accuracy/Audio Perception, (2) COGS at scale, (3) Minimal latency, and (4) Understanding Context.

Developer Tools
Cloud or self-hosted APIs + SDKs (STT, TTS, Voice Agent API)

Real-time Voice AI Models

Foundation Models:
Nova (STT), Aura-2 (TTS), Neuroplex (STS)
Industry/Customer-Specific Models:
Nova-3-[customer], Aura-2-[customer], Nova-3-medical
Task-Specific Models:
Diarization, End-of-Thought, Agent, Language Detection
Open-Source Models:
Managed Whisper API

Deepgram Enterprise Runtime
Adaptability, automation, synthetic data generation, data curation, model hot-swapping, integrations
Built for cost-efficiency at scale / low customer COGS

Deepgram Voice AI Platform

Speech-to-Speech is the Future

Deepgram has reached a [pivotal checkpoint](#) on the path to delivering enterprise-grade speech-to-speech architecture—a foundational technology that underpins voice AI agents. This robust technology meets the enduring demand for innovative, reliable voice solutions, as our ongoing research shows that the demand for voice agents is growing, with 84% of organizations expanding their budgets for this transformative technology.

The new architecture promises to deliver the following benefits:

- **Breakthrough in Speech-to-Speech Technology:** Deepgram has achieved a major milestone by developing a speech-to-speech model that operates end-to-end, without relying on text conversion. This advancement enables more natural, real-time voice interactions while preserving important elements like intonation, emotional tone, and contextual nuances.

- **Improved Latency and Naturalness:** Unlike traditional speech-to-speech systems that rely on multiple stages of processing (e.g., speech-to-text and text-to-speech), Deepgram's new architecture eliminates the need for text, improving latency and ensuring more emotionally-aware, fluid interactions for enterprises.

- **Advanced Developer Tools:** The new speech-to-speech architecture includes built-in debuggability, allowing developers to inspect, refine, and troubleshoot each stage of the AI's processing. This is vital for businesses that need control, transparency, and the ability to fine-tune their voice AI systems.

- **Efficient, Scalable Model Training:** Leveraging transfer learning and latent space embeddings, Deepgram reduces the data and costs typically required for training large-scale models, offering high-accuracy results with far less computational overhead compared to traditional methods.

Deepgram's platform delivers an enterprise-ready speech AI stack, which integrates seamlessly with existing infrastructure. This includes high-performance, cost-efficient deployment options that scale with the needs of businesses, ensuring adaptability and innovation for long-term success.

Deepgram's breakthrough marks the beginning of a new era for speech AI. As the technology evolves, enterprises will be able to deploy more sophisticated, end-to-end systems that deliver a fully immersive, responsive, and human-like conversational experience — opening the door for more advanced voice applications across industries.

Deepgram

Since its inception in 2015, Deepgram has been in the business of discovering, developing, and implementing deep learning models that don't merely raise the bar with incremental improvement—they advance the state-of-the art in voice AI performance for enterprises worldwide.

This survey reveals that businesses are looking for reliable, scalable, and human-like voice agents powered by cutting-edge speech-to-speech technology, regardless of the cost. Deepgram aims to provide these enterprises with voice AI they can trust in every possible scenario.

We have transcribed over 2 trillion words—or 50,000 years worth of audio—for hundreds of customers, and we're just getting started.

To start building with Deepgram's APIs for speech-to-text, text-to-speech, and voice agents, sign up and claim $200 in free credit. To learn more visit deepgram.com or contact us.

About the Report

[Opus Research](#) fields an annual survey of 400 decision-makers seeking to identify, evaluate and quantify emerging trends and business use cases for voice technologies and related resources. The goal is to understand the motivation for, and impact of voice technologies built on voice AI to drive efficiencies and productivity for organizations.

The 400 respondents, based in North America representing eight vertical industries, with decision-making and purchasing influence roles across varied business units to understand how voice and speech technologies are being integrated into built for purpose, conversational solutions.

About **opus**research

Opus Research is a diversified advisory and analysis firm providing critical insight on software and services that support digital transformation. They are focused on the merging of intelligent assistance, NLU, machine learning, conversational AI, conversational intelligence, intelligent authentication, service automation and digital commerce. To learn more visit [OpusResearch.net](#)

About **Deepgram**

Deepgram is the leading voice AI platform for enterprise use cases, offering speech-to-text, text-to-speech, or full speech-to-speech capabilities. Over 200,000 developers build with Deepgram's voice-native foundational models - accessed through cloud APIs or as self-hosted / on-premises APIs – due to our unmatched accuracy, latency and pricing. With 50,000+ years of audio processed and over 1 trillion words transcribed, no one understands voice better than Deepgram. Learn more at [deepgram.com](#).