State of Voice

2023

Presented by

Deepgram +

Report by

opusresearch

Table of contents

Introduction

Let's just call it now:

2023 is the year artificial intelligence really hit the mainstream.

Where to start? There's the ascendance of conversational AI backed by transformer-based large language models (LLMs) that seem capable of explaining or writing basically anything, from prose to Python. There are diffusion models—originally meant to mathematically represent thermodynamic phenomena—that can create visual art, make music, and generate 3D objects, all of which are pulled out of Gaussian background noise.

Foundational AI models deliver a somewhat spooky but nonetheless magical experience to tens of millions of active users around the world, enabling nearly anyone to conjure up whatever they want with little more than a string of words.

Speaking of speech: There's the world of voice AI. End-to-end deep learning (E2EDL) is capable of near-human performance for speech-to-text transcription at virtually limitless scale, and orders of magnitude faster and more affordably than human transcribers. Natural language understanding (NLU) and speech analytics capabilities augment the human intelligence of folks working in fields ranging from customer support to high finance and seemingly everywhere in between.

Deepgram's 2023 State of Voice Survey, conducted by Opus Research, delves into the applications and key features of speech AI across over a dozen industries from the perspective of 400 business leaders surveyed for this report.

2023 promises to be among the most exciting years yet for the field of speech AI. Let's learn why.

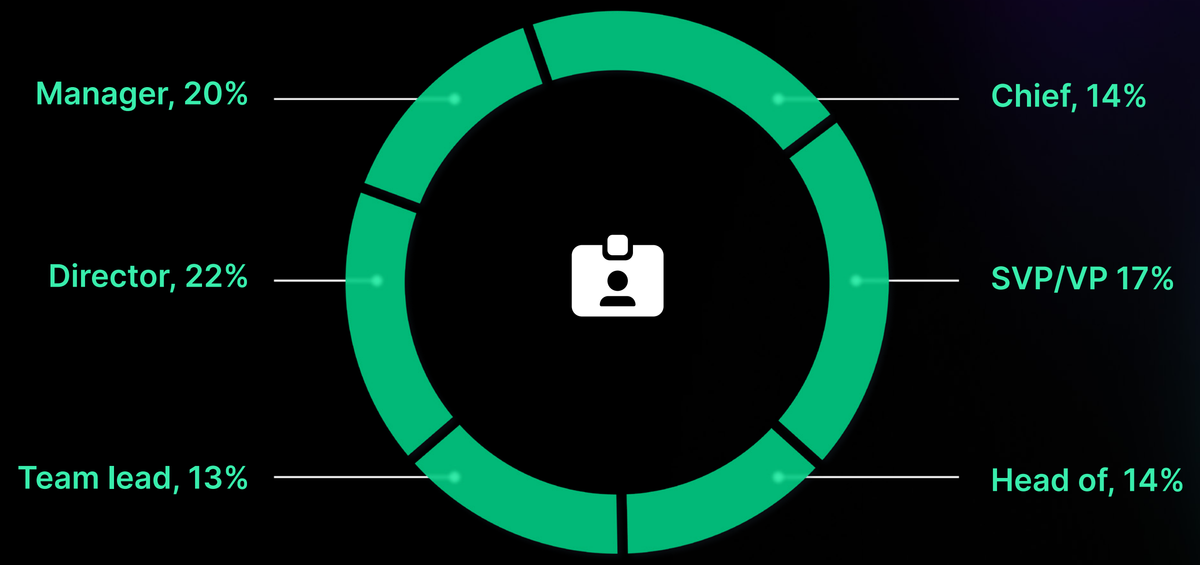
Who we surveyed

To generate a representative dataset, we surveyed 400 business leaders from the United States and Canada. The surpassing majority—88 percent—were based in the U.S.. 45 percent of respondents are either C-suite executives, (senior) vice presidents, or heads of business units in their respective organizations. In other words, these folks are the key decision-makers at their companies.

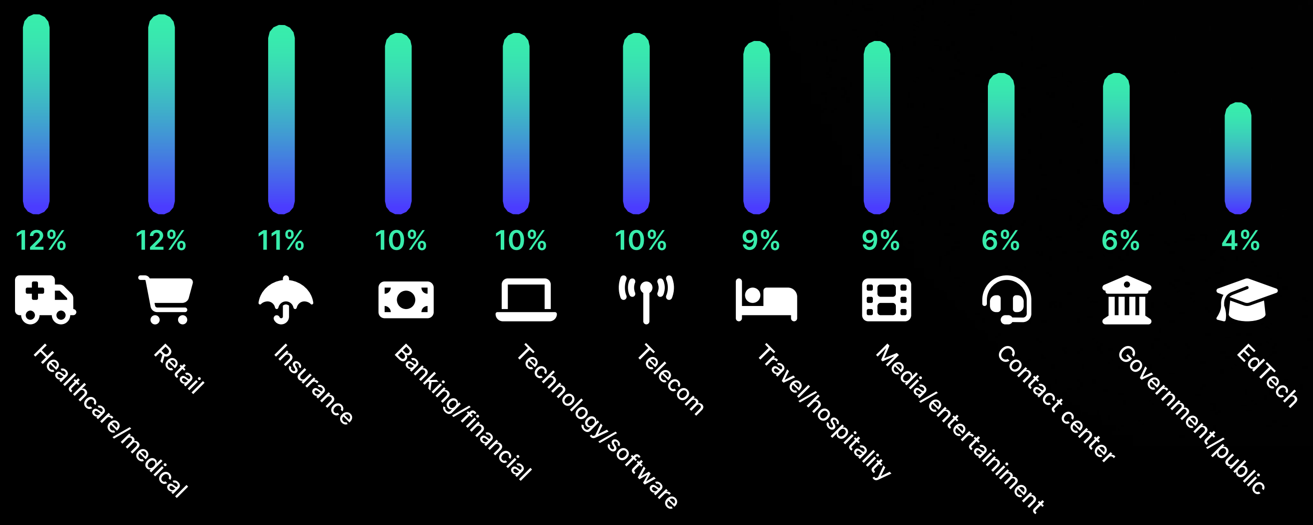
As far as the industries represented, we made sure to draw from a diverse base. The best-represented industries are healthcare and retail—each representing 12 percent of the respondents. Other industries include telecommunications, contact centers, banking, and educational technology. Here is a more complete breakdown of industries covered in this survey.

Regardless of industry, most companies have a similar set of internal departments. For example, it doesn't matter if you're running a hospital system or a large consumer bank. Customer service, business analytics, and data science are all functions you'd expect to find in nearly any organization operating at scale. Nearly half of the respondents in our data were from customer experience, marketing research, or business analytics departments.

Current job title



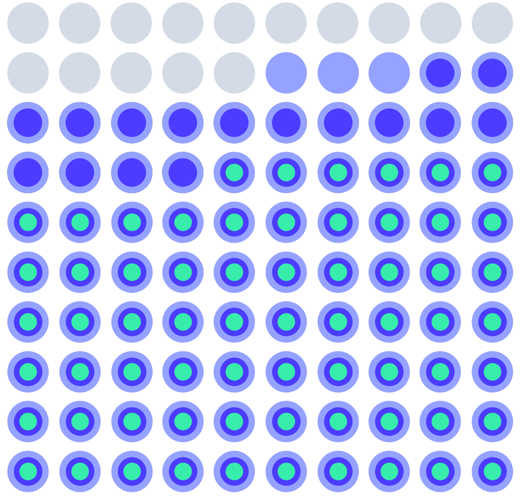
Primary industry



Executive summary

Major finding: Voice is a well-understood, yet under-utilized resource

1 Near-term Expectations are high



85%
expect widespread deployment of voice technology in the next five years

82%
use voice technology today

66%
see voice-enabled experience as "important to the future of their companies' strategies

2 That includes expectations for voicebots

56%
expect life-like voicebots in 1-3 years

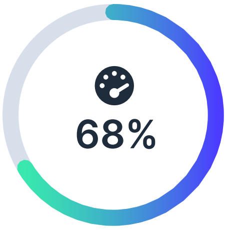
40%
see voicebots achieving humanlike levels of interaction in less than a year

3 Voice remains an underutilized resource

84% of respondents
transcribe less than 50% of their available audio data

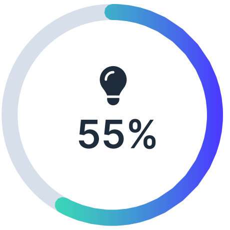
Only 1% of respondents
transcribe over 75% of their available audio data

4 Current deployments reflect the pragmatic considerations of challenging economic times



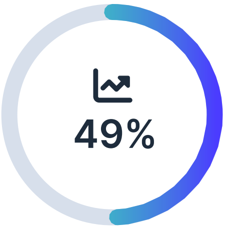
68%

improved productivity



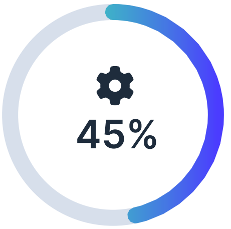
55%

find new business opportunities



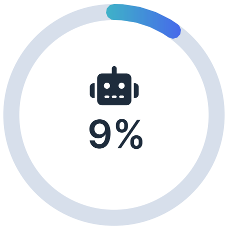
49%

increase revenues



45%

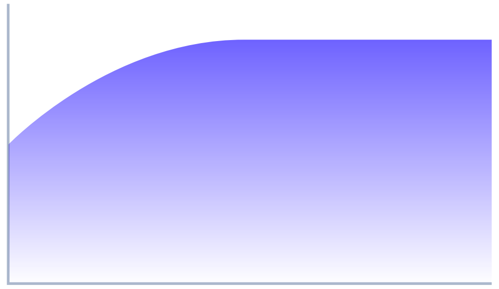
promote operational efficiencies



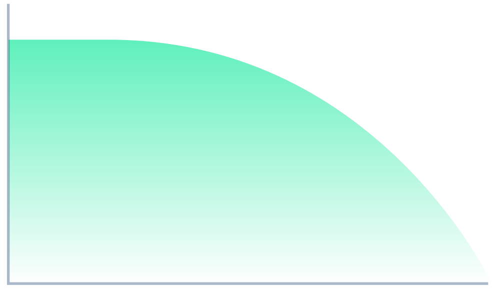
9%

only 9% see convo AI as "transformational"

5 Not implementing voice technologies has real repercussions



Revenue and customer acquisition expected to plateau



CSAT and Net Promoter Scores expected to decline

The voice landscape

The voice technology landscape is vast, spanning from simple speech-to-text (SST) and speech analytics, to advanced natural language understanding (NLU) capabilities like sentiment analysis and automatic language detection. Advances in the state of the art of deep learning and computational linguistics also enable new capabilities such as speech synthesis (also known as text-to-speech, or TTS) and real-time machine translation, among many others.

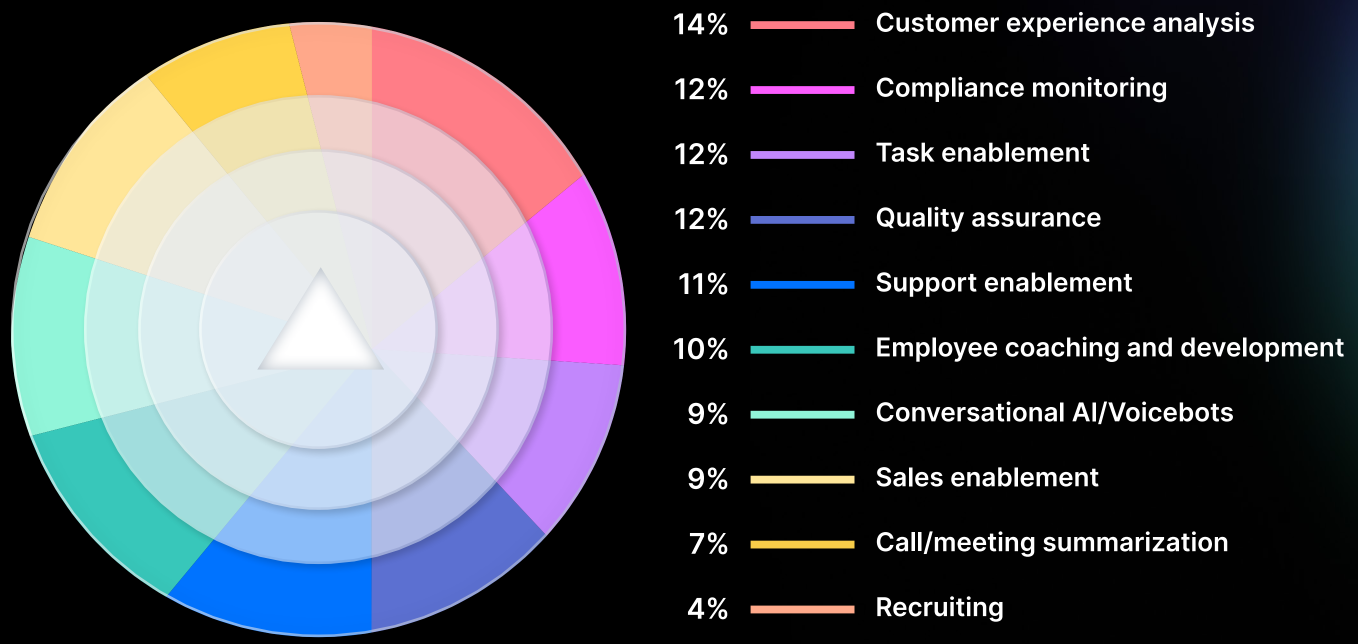
Our aim in updating this yearly report is to gain a better understanding of the present situation and how it's evolving over time. In this section, we'll explore how organizations are leveraging voice technologies in their operations.

Current uses for voice tech

Voice technologies are finding applications across industries and departments for a diverse range of use cases. However, the primary uses of voice tech remain largely consistent, regardless of the department or vertical in question. The most common voice AI technologies in use today are speech-to-text, speech analytics, and voice as a component of business intelligence.

As far as business use cases are concerned, we're seeing that Speech AI continues to give organizations unprecedented visibility into understanding interactions with their customers and thus identifying data-driven paths to improving the experience they deliver to customers.

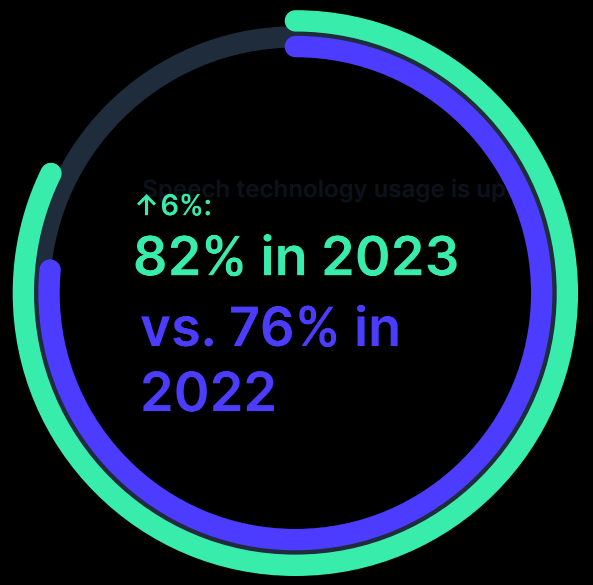
Q: In general, what do you believe to be the most transformative use cases for speech technology?



The relatively even distribution of responses highlights the diversity of needs that voice AI can meet. 50 percent of respondents said the most transformative use case for voice technology is in customer experience analysis, compliance monitoring, task enablement, and quality assurance, but the drop-off in responses for other use cases was not terribly steep.

Who is using voice technology?

Adoption of speech technology continues to rise, with 82 percent of respondents saying that they do use speech technology within their organizations. That's up from 76 percent in our 2022 survey.

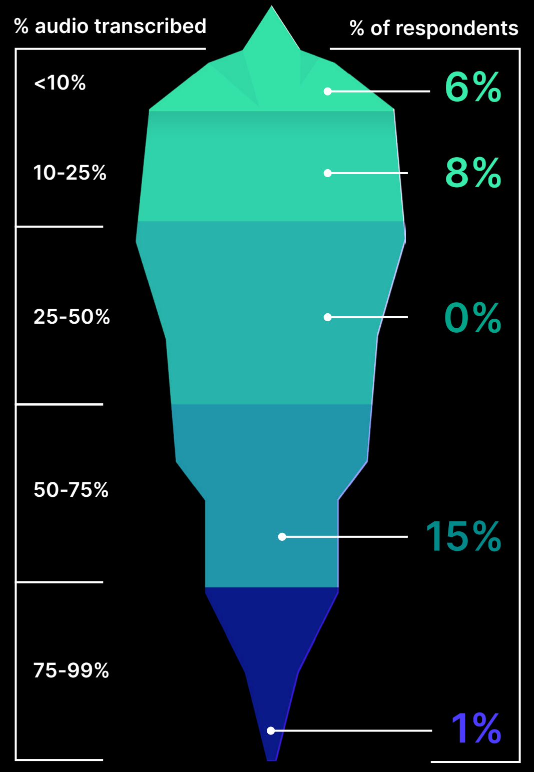


It's clear that as a company scales up, its need for some type of voice technology grows as well. This is evidenced by the fact that the largest organizations surveyed (e.g. those with 2,500 or more employees) were the biggest adopters of voice technology, with fewer than 1 in 8 organizations using no voice technology solution at all. 2023's survey finds the same pattern, with 92 percent of companies in the 2,500-plus employee range stating that they use voice technology.

Adoption of voice technology

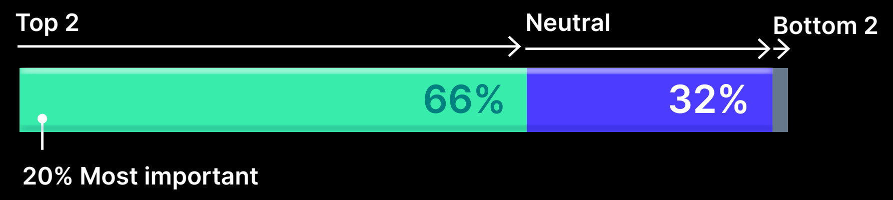
Given that over 4/5ths of respondents say they use voice technology, it's notable that, of those, 84 percent of respondents transcribe less than half of their audio data for business use cases. This leaves plenty of room to scale up for greater insights and more reliable data points to drive decisions.

Q: What percentage of your audio is being transcribed for business use cases?



2/3 of respondents see voice-enabled experiences as important to the future of their companies' strategies. Expectations are high for voice to improve both customer and employee experiences and contribute to improving bottom line results.

When asked "how important do you think voice-enabled experiences are to the future of your company's enterprise strategy?" 66% saw it as important with 20% citing it as "most important."

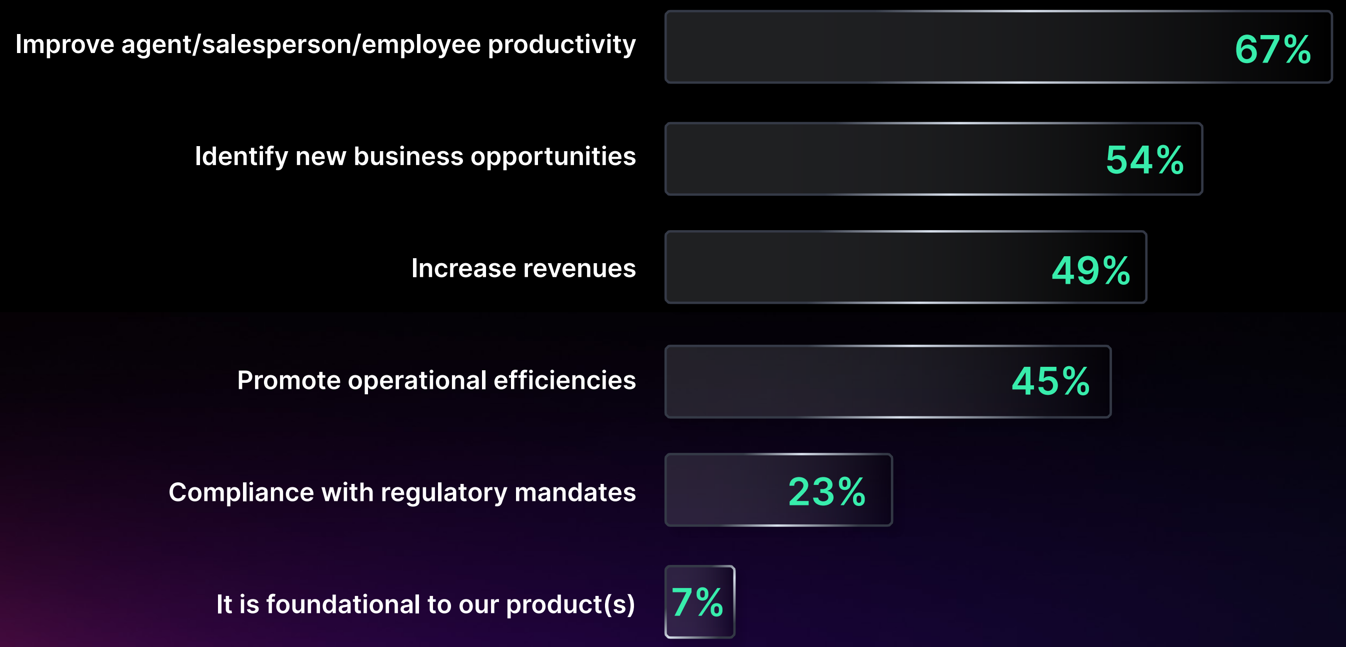


Motivations for adoption

When it comes to adopting voice technology, efficiency is usually the motivation.

- 67 percent of respondents said that improving agent and other employee productivity was a factor driving adoption.
- 45 percent said they adopted speech technology to promote operational efficiencies.

Speech technology is for growth-minded companies too: 49 percent of respondents said they adopted speech tech to increase revenues.



Which Speech Provider Is Best, and Why?

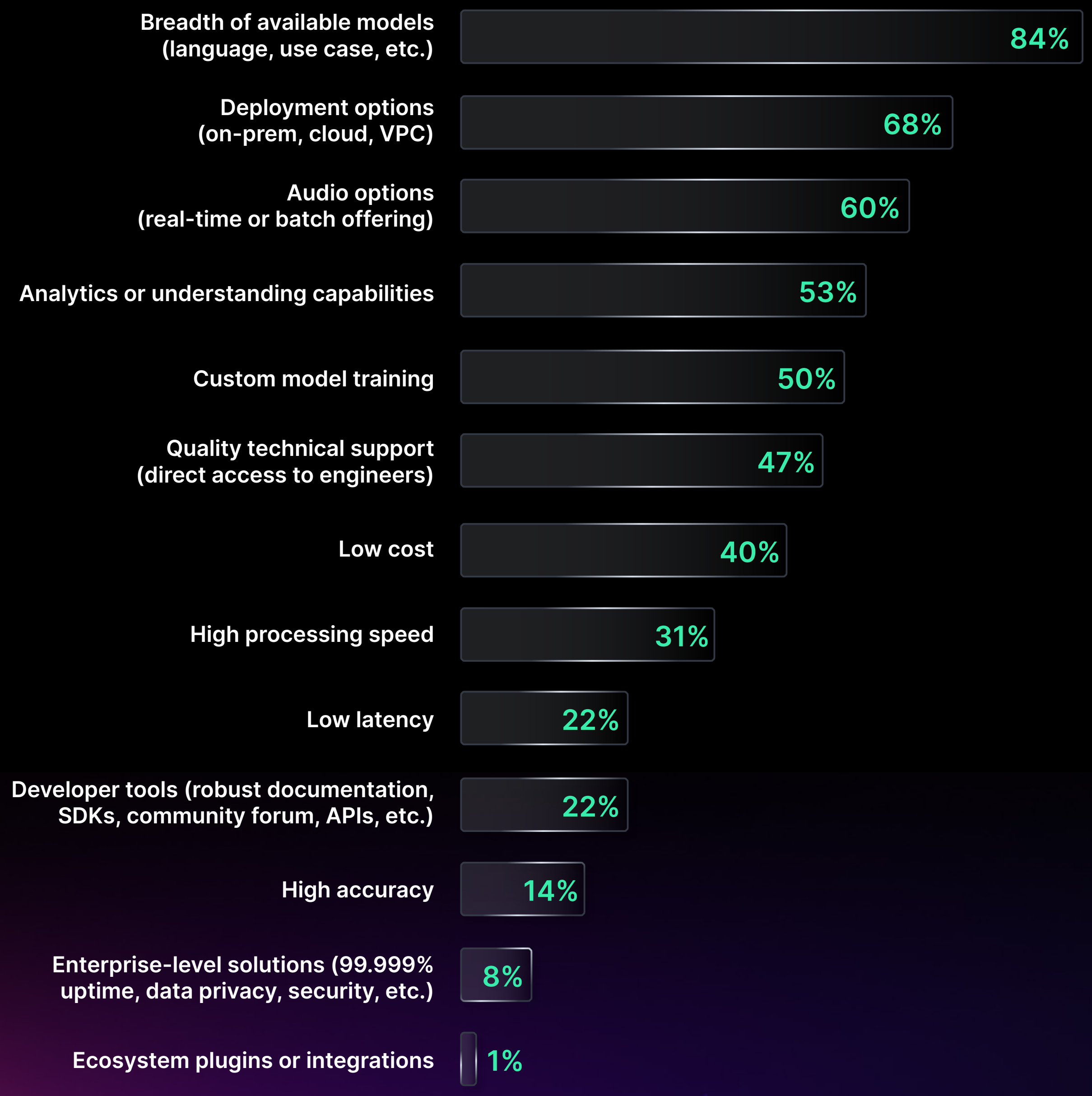
What makes a speech provider “good?” This year’s survey data points to several trends.

First and foremost: capability is paramount. 84 percent of respondents said that the breadth of available models—the number of languages they can work with, whether they are fine-tuned to a specific use case, etc.—is the most important thing to consider when assessing the quality of a speech provider’s offerings.

Next, flexibility is key; the more ways a provider can adapt to customers’ needs, the better it is viewed. 68 percent of respondents said deployment options are important, and 60 percent wanted flexibility in audio formats (e.g. streaming versus asynchronous batch processing).

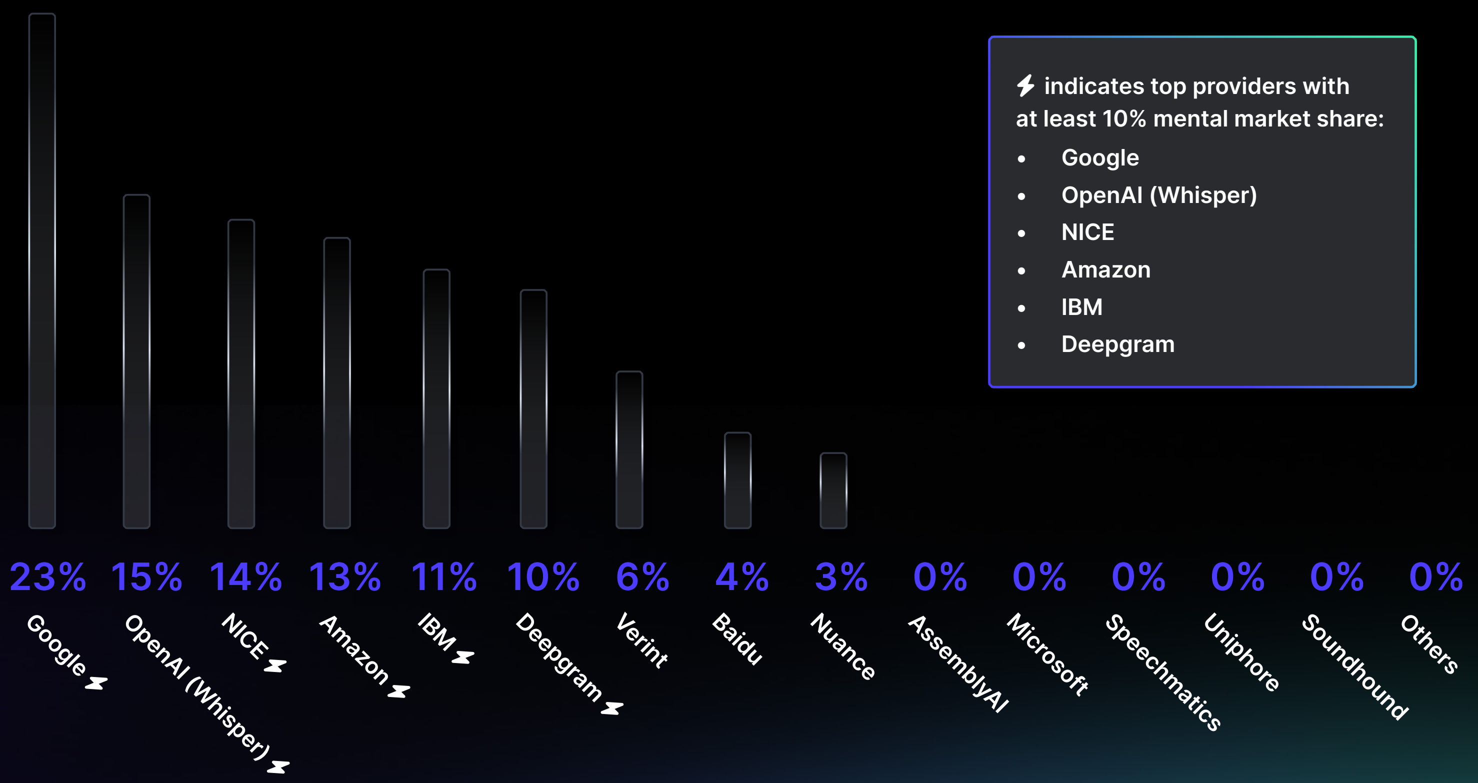
Finally, and perhaps most surprisingly, much less than half of respondents said that low cost, high processing speed, and low latency were not major factors.

Q: What does “best” mean to you when referring to speech recognition?



By these metrics, who did respondents say was the best speech recognition provider? 23 percent said Google, which has been in the ASR market for years. OpenAI’s open source speech recognition model, Whisper, was only released in September 2022, yet 15 percent of respondents already said it was the best. 10 percent of respondents said Deepgram was the best ASR vendor for businesses.

Q: Which of the following technology companies provide the best speech recognition for business?

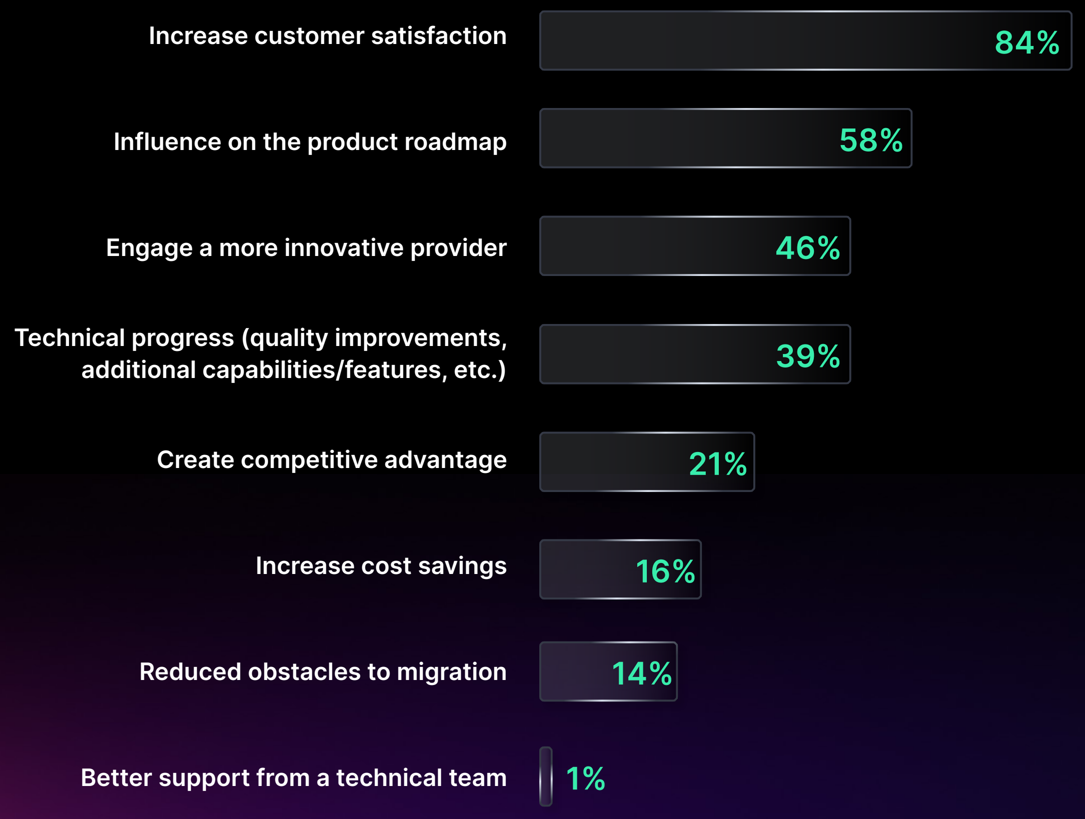


⚡ indicates top providers with at least 10% mental market share:

- Google
- OpenAI (Whisper)
- NICE
- Amazon
- IBM
- Deepgram

What would it take for a company to implement or switch speech technology providers? Good news for startups in the space: influence on product roadmap was a major factor for 58 percent of respondents, and 46 percent would switch to engage a more innovative provider. Incumbents definitely have scale on their side, but nimble teams building something better, while also providing a better partnership experience, stand a fighting chance to gain market share.

Q: Which of the following would push you most to implement or switch speech technology providers?

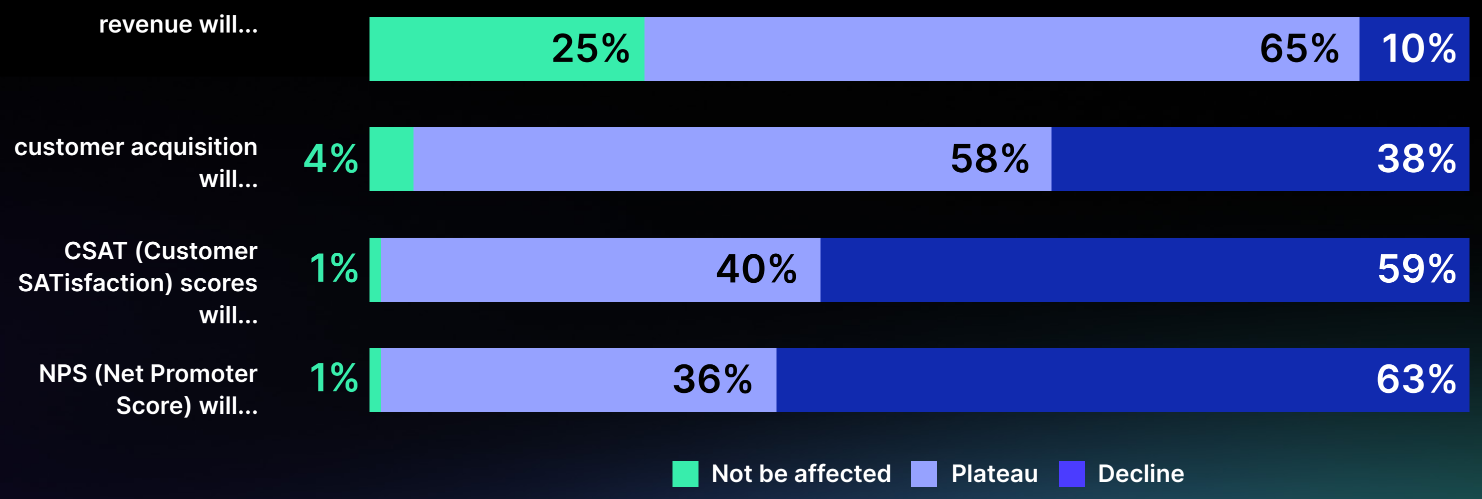


How (and Where) Voice Tech Adds Value

From the factory floor to the corner office, executives are turning up the volume on voice technology. As our survey results show, voice technology is making waves across multiple industries, with many executives seeing its potential to transform their businesses. In this section, we delve into the specific ways and contexts where voice tech can add value, from boosting productivity to improving customer experiences.

One way to think about how a technology adds value to an organization is to ask, “What happens to your industry peers if they don’t adopt it?” And that’s exactly what Opus asked in the survey. A majority (59 percent) of respondents said that businesses which do not implement voice technology can expect their customer satisfaction scores will decline, which would make customers less likely to enthusiastically recommend their products and services. To that end, 63 percent of respondents said that Net Promoter Scores (NPS) would decline without voice capabilities.

Without voice capabilities I believe that businesses can expect that:

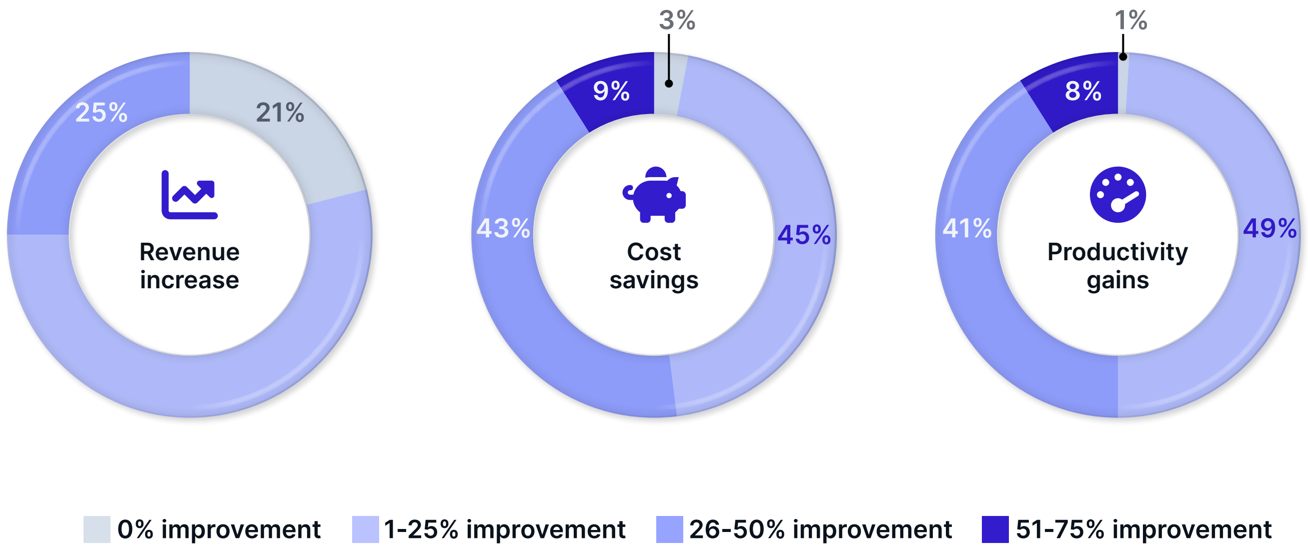


What's the ROI?

In addition to asking about impact, we also gave respondents the opportunity to share their assessments of the savings and gains they've made as a result of implementing voice technology.

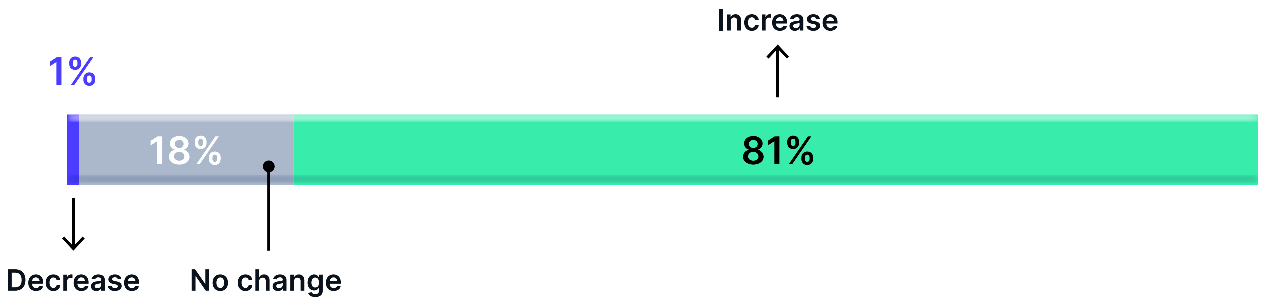
The biggest impact seems to be a new focus on cost savings, a change from last year where productivity saw the biggest boost. 51 percent of respondents said they saved between 26-75 percent on costs by using voice technology. Productivity gains aren't far behind though: 49 percent of respondents said that productivity increased between 26-75 percent as a direct result of implementing voice technology.

Savings and gains



Our survey results also show that investment in voice technology will continue apace. 81 percent of respondents said they'll either "Increase" or "Increase Significantly" their speech technology budgets. Basically nobody is planning for cuts. This points to the valuable, deeply embedded role of speech technology in companies ongoing digital transformation efforts, as well as day-to-day business operations.

Q: Over the next 12 months, do you expect to increase, decrease, or keep your speech technology budget the same?



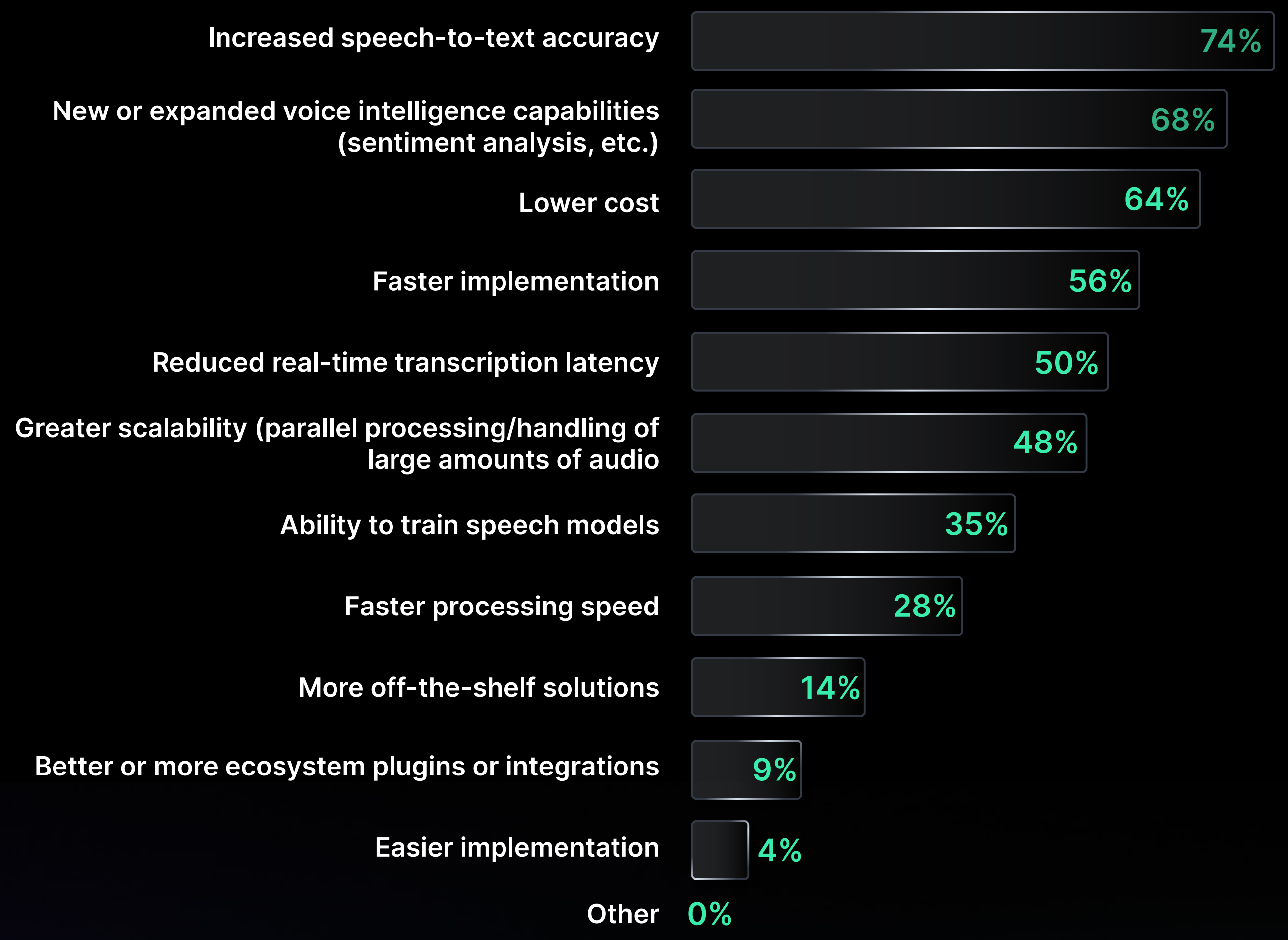
Barriers to adoption

It’s clear that voice technology adds value for companies which adopt it. But what are some of the roadblocks to even higher adoption rates? Put differently: What factors would help increase adoption?

- 74 percent said improved speech-to-text accuracy would drive higher adoption.
- 68 percent said new or expanded intelligence capabilities would increase adoption
- Lowering cost was a factor for 65 percent of respondents
- Making implementation faster was a factor for 56 percent of respondents

To generalize a bit, improving speed, accuracy, scalability, and customizability of speech models would likely increase adoption. The sophistication and capability of speech technology will only increase in coming years, and, with it, adoption rates.

Q: Which of the following factors would help increase your organization’s use of speech technology

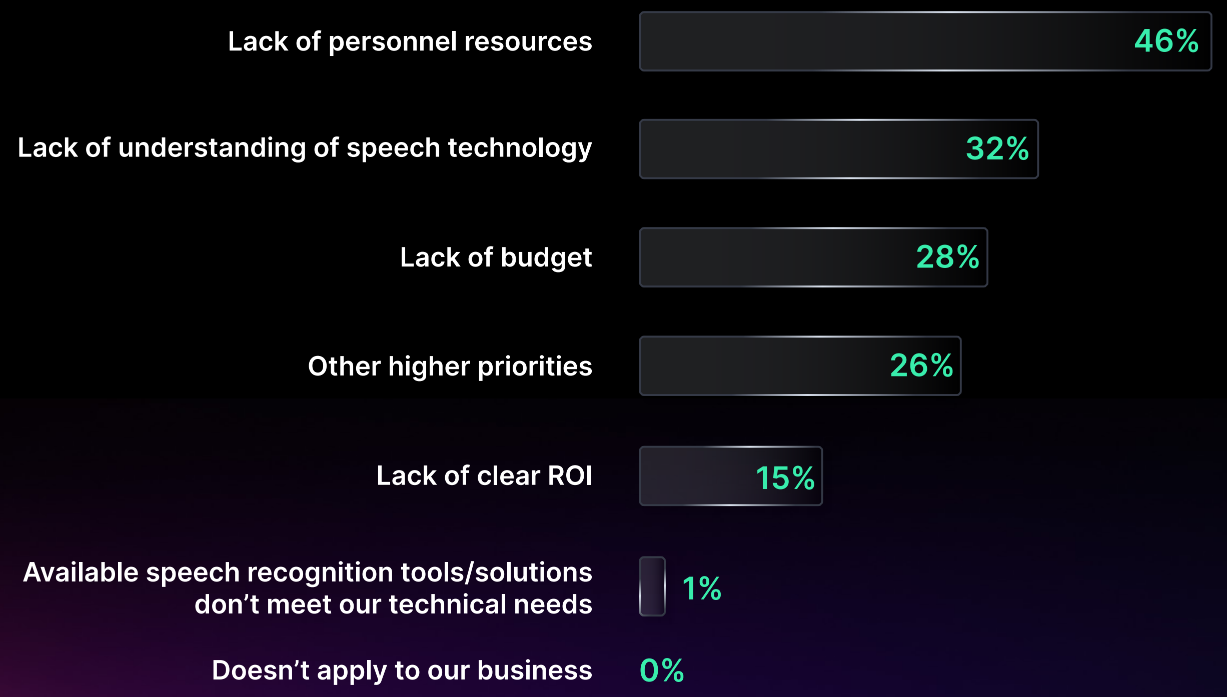


OK, so, what's actually blocking companies from unlocking more value through speech technology? We asked respondents who are not using speech tech why. Here are the top three reasons:

- 46 percent said they lacked personnel
- 32 percent said they lacked an understanding of the technology
- 28 percent said they lacked budget

Nobody said that speech technology was simply irrelevant to their business.

Q: If you're not currently using speech technology, why not?



So what does this say about the present and future of speech technology adoption? What needs to change? We find three opportunities for speech companies to consider:

- It's the nature of computing to get less expensive over time, and speech models become more efficient as the state of the art progresses. Costs will come down for end users over time, removing a significant barrier to adoption.
- Advancements in model architecture and training can eke out even higher accuracy, and new capabilities like sentiment analysis add new dimensions to data.
- It's also clear that customer education is another major opportunity. Informing customers about what is possible, clearly explaining the ROI, showing how off-the-shelf solutions mean they don't need to hire an army of in-house speech technologists, and continuing to streamline the implementation process all serve to take the uncertainty out of a speech technology buying decision.

How could LLMs help?

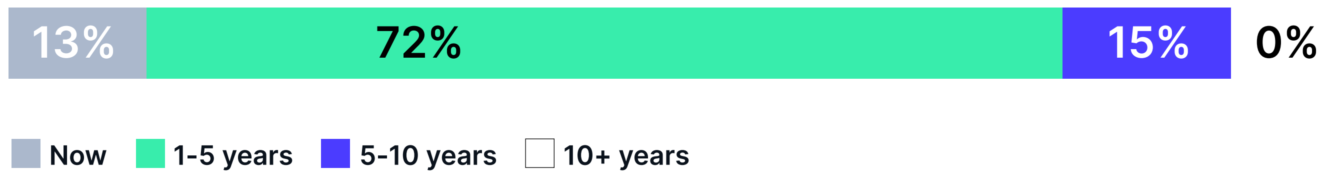
By design, they're phenomenal at natural language processing (NLP) tasks and text transformations. A large language model could produce better call summaries, improve on-the-fly translation, help generate cues for sales and support staff based on ongoing conversations, and even help automate rote tasks in a way that feels a little less robotic. As large language models gain prominence, expect more of their capabilities to be rolled into speech AI technologies.

What's next for speech tech?

In many ways, it feels like the future is already here. End-to-end deep learning powers the most capable speech AI models in the market. Intelligence and understanding features like language detection, translation, and sentiment analysis inform users about not just what was said, but who said it, how they said it, and maybe even why they said it.

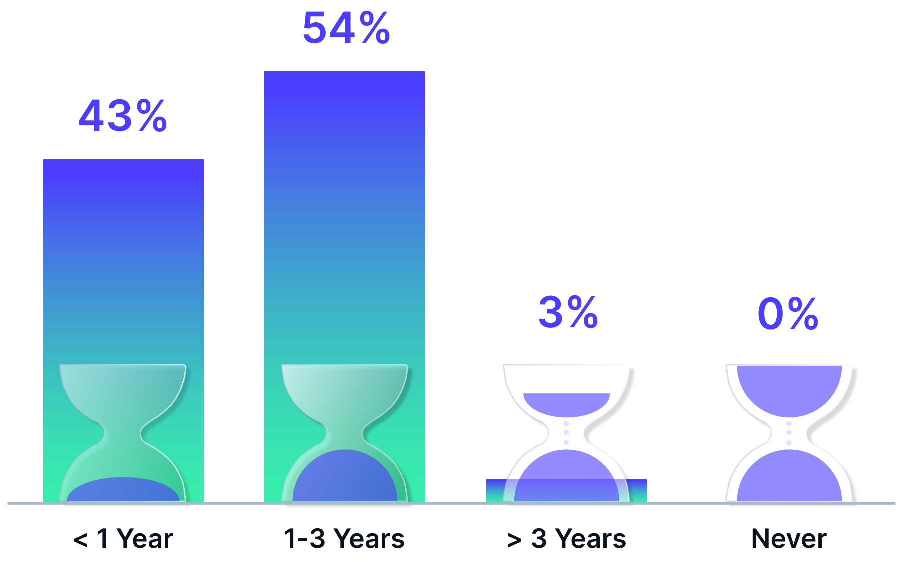
The foundation is there. Now it's time to build out rich, interactive, voice-enabled experiences. How long until we see widespread use of voice-enabled experiences? 13 percent of survey respondents said that widespread use is already here, and 72 percent said that voice-enabled experiences will gain mass adoption over the next one to five years. In other words, comfortably before the end of this decade, we can expect to see these experiences become the norm.

Q: When will we see widespread use of voice-enabled experiences in business?



Considering the advances we've seen in conversational AI, brought about by large language models like OpenAI's ChatGPT, the voice-enabled future may come even sooner. Asked when voicebots will attain human-like levels of interaction, 43 percent of respondents said it's going to happen in less than a year, and 54 percent said that human-like AI voicebots are between one and three years away.

Q: In what timeframe do you think voicebots will reach humanlike levels of interaction?



At least right now, talking to a large language model can still feel a little bit robotic, and it's easy to find hard-coded limits on the actions it can perform or the outputs it produces. However, for narrowly-scoped interactions like the ones we regularly have with human customer service, technical support, and others, it's not hard to imagine a future where it's unclear whether you're speaking with a human or a large language model.

The future:
Large Language Models (LLMs)

On November 30, 2022, consumer perception of generative AI's capabilities changed forever with the release of OpenAI's ChatGPT, a large language model (LLM) fine-tuned for conversation through a technique known as reinforcement learning with human feedback (RLHF).

ChatGPT certainly wasn't the first LLM, but its intuitive chat-based interface, and ostensibly limitless knowledge gleaned from the massive corpus of text data it was trained on, was enthralling. Investment analysts at UBS asserted in early 2023 that ChatGPT was the fastest-growing consumer app of all time, hitting 100 million monthly active users in just two months.

This points to the sustained future of LLMs up and down the entire language technology stack: from the most sophisticated models running on supercomputers in the cloud to what might, in a not-too-distant future, become a native feature of smartphone operating systems.

This section considers the trends, history, and applications of LLMs in speech technology.

Key trends

Moving beyond computational benchmarks to human capability assessment.

Following OpenAI's launch of ChatGPT, which ushered in the era of broad awareness of LLMs, academic researchers and other experts have been testing models' aptitude for academic and professional tests of human cognitive abilities. The fact that an LLM could achieve anything approaching human capabilities on such tests was, if nothing else, a curiosity, or even perhaps an indictment of the tests themselves. However, GPT-4 and related models show that LLMs have transcended legacy benchmarks of NLP performance, and have transitioned to competing against humans on similar tasks. Does this mean new LLMs, like GPT-4, have surpassed human intelligence in a generalizable way? Absolutely not. Do they light a path to the fabled threshold of artificial general intelligence (AGI)? Maybe, but it's clear that there are, as yet, years, if not decades, of work to be done to achieve it.

Multimodality is the future.

LLMs excel at complex natural language processing tasks, from writing prose in the style of basically any widely-published author or publication, to composing computer code from human language inputs. But much like how we conceptualize the world in terms of language, developers of LLMs are just beginning to instill analytical and inferential skills beyond the domain of text.

Such multi-modal models can, for example, transform a text string like "Show me a quick clip of a happy dog going down a water slide" into a still image or brief video. Partnering LLMs with other foundational generative models, in an "ensemble," enables the production of synesthetic outputs from a bespoke blend of media.

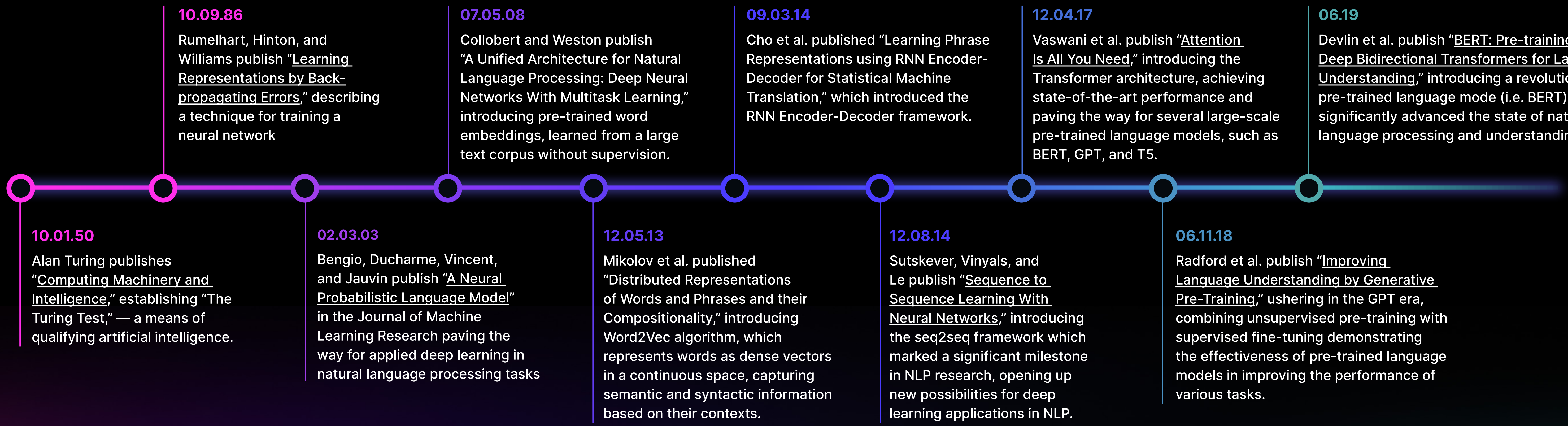
Parallel paths to the future of LLMs.

Much like other types of generative AI models, LLMs are likely to evolve as a function of how and where they're deployed: on-device vs. in the cloud. We can already see the fork in the road.

- OpenAI's GPT model family, alongside LLMs from Google, Anthropic, BigScience, Cohere, and others are, generally, huge (on the order of tens or hundreds of billions of parameters), and like all machine learning models require significant, cloud-scale computational resources to not just train, but run inference too.
- Much like how the Stable Diffusion model enables image generation on the edge, smaller, more streamlined models—such as Meta's LLaMA or Stanford's derivative model Alpaca—can be run (albeit slowly) on consumer hardware, from off-the-shelf gaming GPUs to smartphone SoCs. We expect this schism to widen over time, and it raises questions about where the value of LLMs will ultimately land.

The briefest history of LLMs

Unlike how in the latter half of 2022 when generative AI image models like DALL-E 2, Midjourney, Stable Diffusion, and others had a fleeting, if nonetheless exceedingly bright, moment in the spotlight, interest in large language models has had a bit more of a long, slow burn before bursting into the public consciousness.



10.01.50
Alan Turing publishes "Computing Machinery and Intelligence," establishing "The Turing Test," — a means of qualifying artificial intelligence.

10.09.86
Rumelhart, Hinton, and Williams publish "Learning Representations by Back-propagating Errors," describing a technique for training a neural network

02.03.03
Bengio, Ducharme, Vincent, and Jauvin publish "A Neural Probabilistic Language Model" in the Journal of Machine Learning Research paving the way for applied deep learning in natural language processing tasks

07.05.08
Collobert and Weston publish "A Unified Architecture for Natural Language Processing: Deep Neural Networks With Multitask Learning," introducing pre-trained word embeddings, learned from a large text corpus without supervision.

12.05.13
Mikolov et al. published "Distributed Representations of Words and Phrases and their Compositionality," introducing Word2Vec algorithm, which represents words as dense vectors in a continuous space, capturing semantic and syntactic information based on their contexts.

09.03.14
Cho et al. published "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," which introduced the RNN Encoder-Decoder framework.

12.08.14
Sutskever, Vinyals, and Le publish "Sequence to Sequence Learning With Neural Networks," introducing the seq2seq framework which marked a significant milestone in NLP research, opening up new possibilities for deep learning applications in NLP.

12.04.17
Vaswani et al. publish "Attention Is All You Need," introducing the Transformer architecture, achieving state-of-the-art performance and paving the way for several large-scale pre-trained language models, such as BERT, GPT, and T5.

06.11.18
Radford et al. publish "Improving Language Understanding by Generative Pre-Training," ushering in the GPT era, combining unsupervised pre-training with supervised fine-tuning demonstrating the effectiveness of pre-trained language models in improving the performance of various tasks.

06.19
Devlin et al. publish "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," introducing a revolutionary pre-trained language mode (i.e. BERT)that significantly advanced the state of natural language processing and understanding.

Applications of LLMs in voice

As we've seen throughout this report there are seemingly endless applications for voice technology in the enterprise. For the vast majority of these use cases, recent developments in large language models will only improve upon the current state of the art, delivering even richer, more interactive voice-enabled experiences to end users. With a little help from the newly released GPT-4 model from OpenAi, we find a wide array of applications of LLMs in the voice technology space.

Speech recognition: Large language models can improve the accuracy of speech-to-text systems by better understanding context and predicting likely phrases or sentences. This leads to more accurate transcriptions of spoken language, benefiting applications like voice assistants, transcription services, and voice commands.

Text-to-speech synthesis: Language models can be combined with speech synthesis techniques to generate more natural-sounding, expressive, and human-like voices for text-to-speech applications. This technology is used in voice assistants, audiobooks, accessibility tools, and more.

Natural language understanding: Large language models excel at processing and understanding spoken or written language. In voice technologies, this capability is crucial for tasks like intent recognition, sentiment analysis, and topic detection, which are used in applications such as customer support, virtual assistants, and conversation analytics.

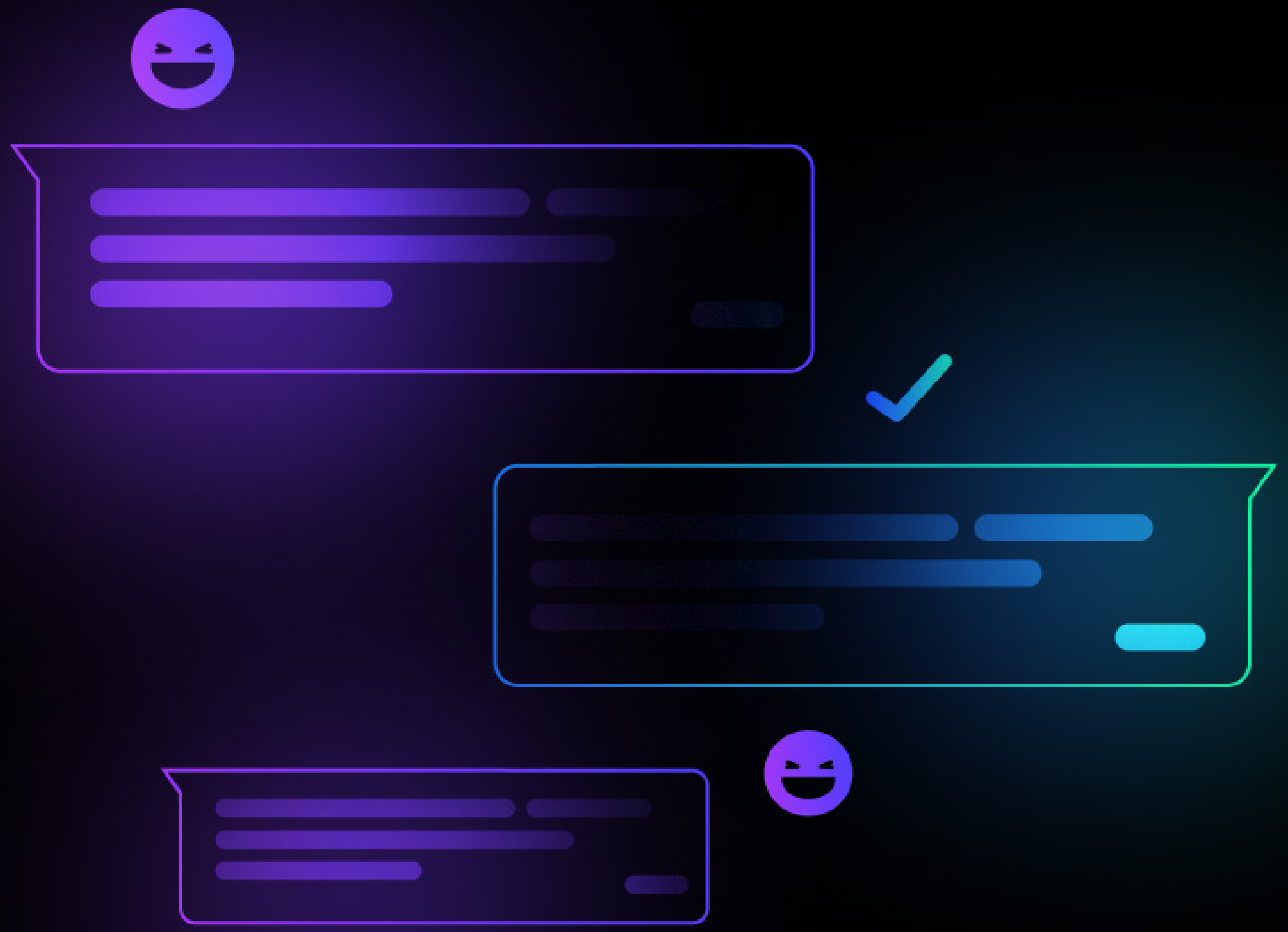
Dialogue systems and chatbots: Language models are at the core of dialogue systems, enabling chatbots and voice assistants to understand user inputs, maintain context, and generate coherent responses. This technology is used in customer support, e-commerce, and personal assistants, among other applications.

Automatic summarization: Language models can be used to generate concise summaries of long audio recordings, such as meetings, interviews, or podcasts. This application can save time and provide quick overviews of essential points for users.

Voice-based content generation: Large language models can help create voice-based content, such as generating text for voiceovers, scripting podcasts, or writing dialogue for interactive voice response systems.

Language translation: Language models can be employed in speech-to-speech translation systems, enabling real-time translation of spoken language and facilitating communication between speakers of different languages.

Speaker diarization and recognition: While not a direct application of language models, combining them with other machine learning techniques can improve speaker diarization (identifying individual speakers in a conversation) and recognition (identifying a specific speaker based on their voice).



Emotion and sentiment analysis: Language models can be fine-tuned to detect and analyze emotions or sentiment in spoken language, which is useful in applications like customer service, mental health support, and marketing research.

Voice-based data analytics: Large language models can be used to extract insights from spoken data, such as identifying trends, summarizing customer feedback, or analyzing employee engagement. This can aid businesses in making data-driven decisions.

Spoken language teaching and assessment: Large language models can be used to design voice-based language learning systems that provide personalized feedback on pronunciation, grammar, and vocabulary. Additionally, they can be employed for evaluating spoken language skills in educational or professional contexts.

Voice-triggered automation: Large language models can be combined with voice recognition to create voice-triggered automation systems for smart homes, offices, or factories, allowing users to control appliances, lighting, or machinery with spoken commands.

Voice cloning for preserving cultural heritage: Language models can be used to generate synthetic voices of individuals who represent endangered languages or dialects, helping to preserve and document these languages for future generations.

Customized voice biomarkers for health monitoring: Large language models can be employed in analyzing vocal patterns to detect potential health issues, such as neurodegenerative diseases, stress, or anxiety. These voice biomarkers can then be used for early diagnosis or continuous monitoring.

Frontier challenges in LLM development

Despite significant progress in the field of large language models, several challenges remain unsolved in their design, implementation, architecture, and usability. Some of the most prominent challenges include:

Scalability and computational resources. Training large language models requires massive amounts of computational power, memory, and energy. Developing more efficient architectures and training algorithms that can reduce resource requirements is an ongoing challenge.

Data quality and curation. Large language models are typically trained on vast amounts of text data collected from the internet, which may contain noise, inaccuracies, or biases. Ensuring data quality and developing strategies to mitigate these issues are crucial challenges.

Model interpretability and explainability. Understanding the inner workings of large language models and explaining their decision-making processes is difficult due to their complexity. Improving interpretability and explainability is essential for building trust and ensuring responsible use.

Bias and fairness. Large language models can inherit and amplify biases present in the training data, leading to unfair or discriminatory outputs. Addressing these biases and ensuring fairness in language model outputs is a significant challenge.

Fine-grained control of output. Language models can sometimes generate outputs that are undesirable or inappropriate. Developing methods to control and fine-tune the generated content without sacrificing creativity or quality is a crucial challenge.

Adaptability and transfer learning. While large language models exhibit impressive generalization capabilities, they can struggle with tasks that require reasoning, common sense, or domain-specific knowledge. Enhancing their ability to adapt to new tasks or domains with limited data remains a challenge.

Multimodal learning and reasoning. Integrating information from multiple modalities, such as text, images, audio, and video, is crucial for many real-world applications. Developing models capable of multi-modal learning and reasoning is an ongoing challenge.

Long-term context and memory. Although large language models can capture some context, they often struggle with understanding and retaining long-term dependencies or complex relationships. Enhancing their capacity to handle long-term context and memory is an important research area.

Robustness and adversarial resistance. Language models can be vulnerable to adversarial attacks or subtle input manipulations, leading to incorrect or nonsensical outputs. Improving their robustness and resistance to such attacks is an essential challenge.

Privacy and security. Large language models may inadvertently memorize and leak sensitive information present in the training data. Developing techniques to ensure privacy and security without compromising performance is an ongoing concern.

Addressing these challenges is vital for unlocking the full potential of large language models and ensuring their responsible, fair, and efficient use across a wide range of applications.

Deepgram’s Vision for Language AI

Deepgram is a foundational AI company focused on creating the essential building blocks of Language AI.

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 Speech AI performance.

Deepgram’s mission is to build the essential components that will enable humans to communicate with computers using natural language, revolutionizing human-computer interaction. We believe language is the universal interface that will unlock the full potential of AI, and we are committed to driving real-world adoption of this technology across industries.

Our vision is to create a future where humans and machines can communicate seamlessly, enabling businesses and their customers to be more productive. In the past, humans had to learn machine language, but with recent advances in AI, machines can finally understand our natural, human language. We are committed to constantly pushing the boundaries of what is possible with AI-powered human-computer interaction.

Together, let’s harness the power of natural language processing to drive real-world innovation and create a future where language is the universal interface between humans and machines.

We have transcribed over 2 trillion words for hundreds of customers, and we’re just getting started.

About the Report

Methodology

Opus Research recently fielded a survey of 400 decision-makers seeking to identify, evaluate and quantify emerging trends for speech recognition technologies and related resources. The specific areas of interest are “speech-to-text” (STT) conversion, which captures and analyzes transcriptions of spoken words (Conversational Intelligence) which employs speech analytics or customized grammars to support understanding or recognition of the meaning or intents of your company’s employees or customers. The 400 respondents represented eight vertical industries (Banking / Financial Services, Contact Center, Government / Public Sector, Healthcare / Medical Services, Insurance, Retail, Telecom, Travel & Hospitality, Media & Entertainment) with decision-making roles across varied business units.

About Opus Research

Opus Research is a diversified advisory and analysis firm providing critical insight on software and services that supports 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**.

opusresearch

About Deepgram

Deepgram is a foundational AI company on a mission to understand human language. We give any developer access to the most advanced speech AI transcription and understanding with just an API call.

Our models deliver the fastest, most accurate transcription alongside contextual features like summarization, sentiment analysis, and topic detection and others.

To learn more visit **deepgram.com**, **create a free account** or **Contact us** to get started.

Deepgram