SoundHound Inc.

Voice Assistants Evolve to Become Revenue Generators

Table of Contents

Introduction

The Challenge of Monetization

Voice commerce, voice ads, and monetization opportunities for brands with voice assistants have been a topic of discourse in the voice-first arena for the last few years. The ubiquity of voice assistants, improvements in voice AI technology, and the realization that brands need a voice AI roadmap to meet growing customer demand are fueling the drive toward more voice-enabled products and services.

Meanwhile, the trends toward custom, branded voice assistant solutions and the investments in time and resources required to implement them are highlighting a gap between the costs and return on voice AI adoption.

Non-Intrusive Voice Commerce

While most companies are viewing the implementation of voice technology as a key element for improving customer satisfaction, most have yet to find a non-intrusive way to show real revenue from the voice user interfaces in their products and services. Instead, many businesses are relying on increased Net Promoter Scores (NPS) to prove ROI. These same organizations may be avoiding pursuing monetization opportunities for fear of lowering customer satisfaction ratings due to what their users would consider unwelcome and intrusive advertising.

"When you think about monetisation, think about the things that happen throughout the customer journey that contribute to moving people along that decision-making funnel towards the point of purchase. Are there things that you could be doing further up the funnel to help solve problems and bring people down the funnel? Facilitating purchasing decisions with in-the-moment product assistance, for example."

Kane Simms, Co-Founder, VUX World, @kanesimms

In a report on mobile app monetization, Statista reported that 42% of app developers feel ensuring non-intrusive ad content was one of the biggest challenges when monetiz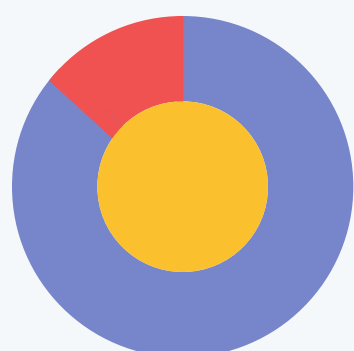ing apps. A recent Opus Research Report supported that finding by revealing a disparity between those companies ranking monetization as an important goal (84%) and those who would consider implementing voice ads (53%).
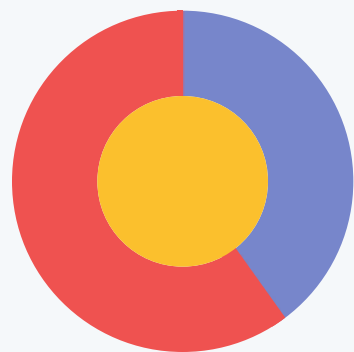
42% of app developers stated that ensuring non-intrusive ad content was one of the biggest challenges when monetizing apps

Statista

84% of business leaders ranked monetization as important compared to other goals

Opus Research Report
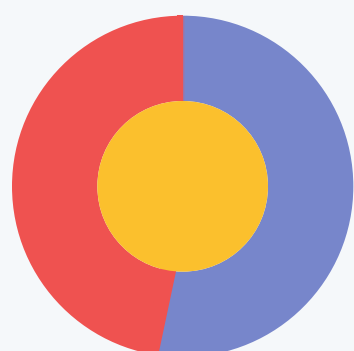
53% of those surveyed would consider implementing voice ads

Opus Research Report

Voice Commerce Pioneers and Roadblocks

A few pioneers, like Pandora, have started to innovate their voice experiences with interactive voice ads that feel more like entertainment than an advertisement. According to Pandora, the ads have been fairly well received with 47% of users stating they either liked or loved the concept and an additional 30% were neutral in their responses—neither liking nor hating the experience. The ads with the highest engagement, according to the company, are those that are entertaining, humorous, or those that used a recognizable brand. These findings represent a benchmark by which other monetization models can be measured for delivering pleasurable user experiences.
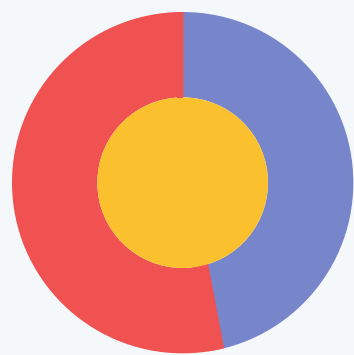
47% of users stating they either liked or loved the concept of interactive voice ads

TechCrunch
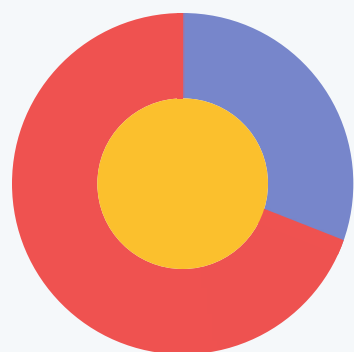
30% were neutral in their responses to questions about voice ads

TechCrunch

According to Voicebot.ai, 22% of smart speaker owners have already tried voice commerce. While this statistic is encouraging for brands hoping to monetize on their voice AI implementations on third-party platforms, there are several common challenges for companies who participate in smart speaker voice commerce.
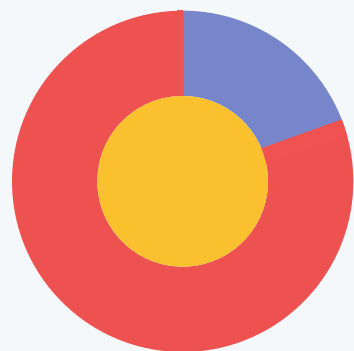
22% of smart speaker owners have already tried voice commerce

Voicebot.ai

Obstacles for voice commerce via a third-party smart speaker include:

- Lack of accurate speech recognition
- Limited "shelf space" available to merchants
- Inability to understand context
- Rising advertising costs
- Competition with voice assistant private labels

These roadblocks to monetization become more feasible to overcome when a voice assistant is wholly owned by the company. Custom voice assistants built on advanced voice AI technology with a custom wake word or wake phrase give control of the brand back to companies. As more monetization models emerge, these custom voice assistants will allow organizations to meet two goals: first, to monetize on their voice AI investment, and second, to deliver a branded omnichannel voice experience to their customers.

"Voice technology is about integration rather than replacement. Meet users where they are at. Try one use case with them and learn from how they use the technology. The ROI for voice is about looking at user data and crafting high quality user experiences based on how users are interacting with your product."

Dr. Joan Palmiter Bajorek, CEO and Founder, Women in Voice, @JoanBajorek

The Voice AI Technology Advantage

Conversational AI that's able to accept a complex set of instructions or queries, understand the context of a conversation, recognize users via two-step authentication, and personalize the experience is at the heart of what's required to make monetization a viable option for voice assistants.

Without the right technology delivering the right content at the right time, user engagement will continue to be a challenge. People will gravitate to the user interface that's easiest to use and the most helpful—a simple concept that should be the goal of any brand implementing or iterating on their voice user interfaces.

In addition, without access to an extensive language library that covers at least the majority of the world's most popular languages and the ability to understand accented language and imprecise speech, global companies will have a hard time implementing monetization opportunities for their diverse customer bases.

The Future of Voice AI is Voice Commerce

The future of voice AI is in the ability of companies to monetize their investments and to show bottom-line growth from their innovation venture. While some industries, such as retail, are already pursuing revenue opportunities with voice shopping and voice commerce capabilities through mobile apps and conversational voice bot assistants, others have yet to discover paths to create monetizable moments with their voice experiences.

"Voice commerce doesn't need to be confined to transactions on a voice-enabled platform itself, such as reorders or in-game purchases. It can, for example, manifest as exclusive deals for voice users, able to be leveraged elsewhere, or adding products to a shopping or wishlist. Opening the aperture on how voice commerce is thought about across the shopping journey can allow brands to get creative and find a fit that makes the most sense for their brand and their customers."

Elissa Dailey, Director of Strategy, RAIN, @RainAgency

We've identified 4 industries that have moved from the nascent stage of voice AI technology adoption to a growth phase where monetization strategies are likely to take hold first. While some industries are notably further along the voice AI journey than others, companies within those leading industries are innovating at different rates. We may see market leaders spring from within industries that aren't known for providing highly-developed voice user interfaces. These companies will be the winners in the war for market share and customer loyalty.

4 Industries on the cusp of monetizing their voice AI investment:

1   Auto Industry

3   Hospitality and Travel

2   Smart Homes and TVs

4   Internet of Things (IoT)

Given the ubiquity of voice assistants and the growing demand for faster, more convenient, and touchless interactions, now is the time to put monetization strategies on your voice AI roadmap. If you're not sure where to start, or you're looking for more information about how monetization will work for your voice assistant, we're here to help.

In this guide you'll discover:

- How voice assistants can be transformed to revenue generators

- The possibilities of revenue generation without an investment in advertising

- Why complex and compound query processing is key

- How multimodal interfaces enhance the voice experience

- The importance of natural language technology that's context aware

- Key industries on the cusp of monetization

- The role of personalization

- The future outlook for voice commerce and monetization

- Why multilingual capabilities are key

- The power of conversational AI

- How to create monetizable moments that don't disrupt the customer experience

Chapter 0ne

An Evolution to Revenue Generator

Chapter One

An Evolution to Revenue Generator

In 2020, 4.2 billion digital voice assistants were being used in devices around the world, and by 2024, that number is projected to reach 8.4 billion units—larger than the world's population, according to Statista. The predictions for the future of voice AI reflect the trend toward businesses offering voice user interfaces in a variety of channels to create omnichannel voice experiences that extend the brand's image and values. Most importantly, these voice assistants will continue to drive increased user satisfaction and brand affinity and unlock revenue opportunities for the brands that own them.

$4.2 Billion
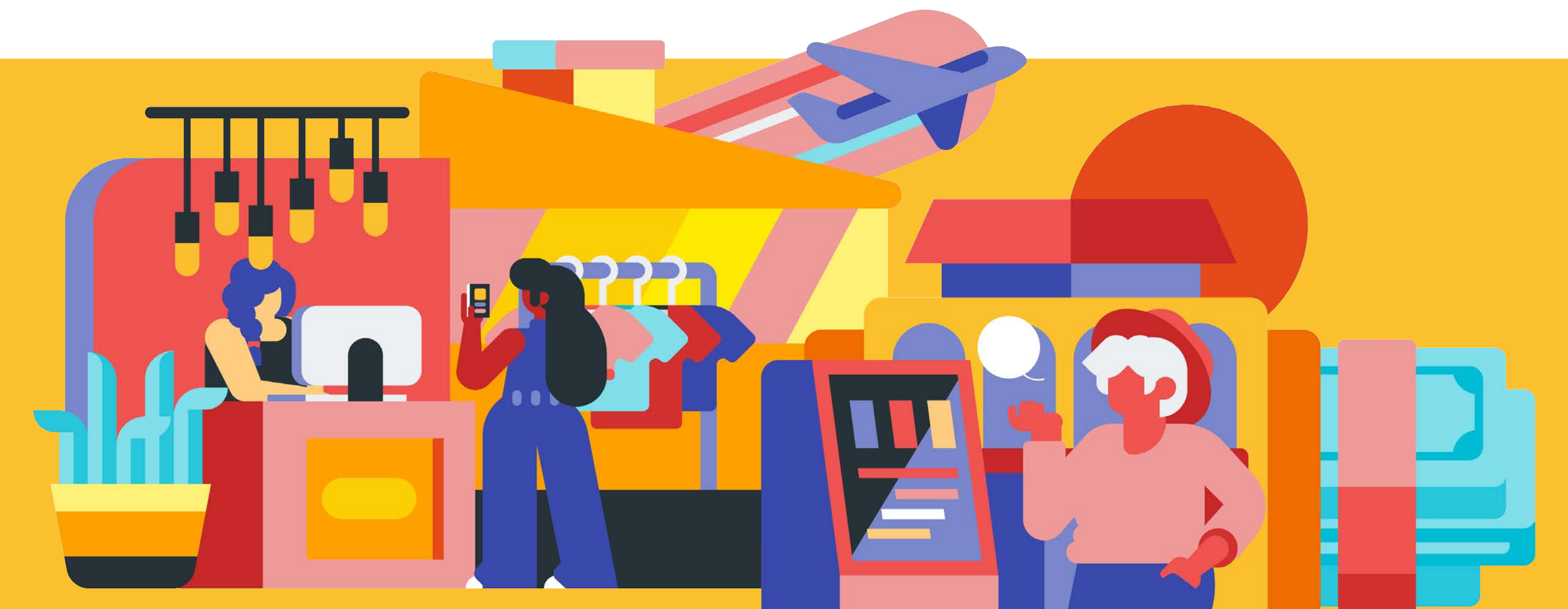
digital voice assistants in devices globally in 2020

Statista

$8.4 Billion

projected units by 2024

Statista

Business leaders in large organizations across industries and around the globe already understand the business value of voice assistants and have implemented voice AI in some capacity, according to a recent study by Opus Research.

The proliferation of voice assistants and the increasingly affordable access to the technology has had the effect of trickling down to small business entities. Voice AI technology, previously out of reach for small businesses, is making its way into voice ordering solutions for restaurants, voice shopping for small retailers, voice-enabled mobile apps, and customer service solutions for small companies of all types. These businesses are mirroring their larger counterparts by looking for more omnichannel voice experiences for their customers. That trend is supported by a study by Small Biz Genius that showed an up to 23% increase in revenue for brands presenting consistently across all platforms.
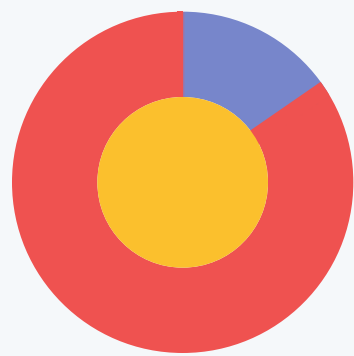
23% increase in revenue as a result of brand consistently across all platforms

Small Biz Genius

Voice Assistants Evolve to Become Revenue Generators

Market and business readiness for voice commerce is best explained from an historical perspective. As with any new technology, voice assistants were met with a mix of enthusiasm and skepticism. In some ways, the smart speaker was a natural evolution from using the smartphone in the home for music listening, entertainment, reminders, and information search. Since people were already engaging in those activities, the smart speaker simply made the process faster, more convenient, and hands-free. As an extra bonus, people no longer had to hunt down their phones to ask for what they wanted.

Around the same time, Honda was the first automaker to offer in-car voice experiences with a voice-enabled navigation system and the ability to make phone calls. The embedded voice assistant gave drivers a way to control the audio, DVD, and climate without taking their eyes off the road. The introduction of voice technology into cars gave auto manufacturers a way to improve the driving experience while making it safer and hands-free.

From there, the voice-first era began its steady climb from research and development to rapid adoption as more use cases were identified and more companies discovered the dual value of offering convenience to customers while creating operational efficiencies. Many companies looking to create a unique voice experience for their users first turned to mobile apps as the low-hanging fruit of voice AI implementation.

Since people were already using their voices to invoke Siri, and the hardware capabilities already existed on their devices, adding a custom voice AI to a mobile app was an easy way for brands to quickly implement their own voice assistants.

The trends that pushed companies to develop custom voice assistants for their mobile apps are now accelerating the need for omnichannel voice experiences.

"The name of the game is omnichannel commerce and finding where voice plays a role in the journey — whether it's using it as a reminder to buy something later, checking on the status of an order, or completing a purchase that began elsewhere. Voice commerce will not work in a vacuum, it must be additive to the shopping experience."

Nick Rovisa, Director of Business Development, RAIN, @nr

As the investment in voice AI grows, so does the desire to show more of a direct impact on the bottom line rather than relying solely on increased customer satisfaction and growing NPS scores. While members of the C-Suite are still considering the implications of voice commerce on customer satisfaction scores, many voice app developers have already made the mindset shift from voice AI as a way to

improve user experiences to voice assistants as revenue generators. According to a report by Voicebot.ai, 30% of voice app developers in 2020 wanted to construct voice apps with monetization built-in. While developers understand the potential for voice AI to deliver monetizable moments, the creation of unintrusive voice commerce opportunities is elusive.

30%
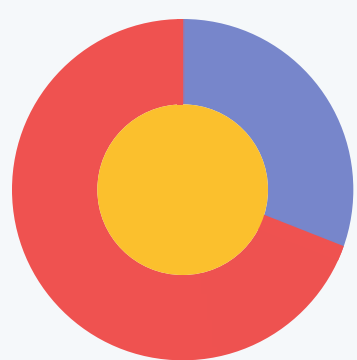
of voice app developers in 2020 wanted to build-in monetization to voice apps

Voicebot.ai

A Monetization Mindset Shift

Business leaders across industries agree that the first priority for any voice user interface must be the user experience. As voice ads and other voice commerce opportunities become more widely available, there is likely to be hesitancy for early adoption due to fears of intruding into the user's space and interrupting the flow of conversation. Adopting new forms of voice commerce will likely require a shift in how we think and talk about voice commerce. The migration from voice ads to monetizable moments is not unlike recent changes to TV advertising.

> *"Your monetization model needs to fit the experience. Some experiences lend themselves well to subscriptions, while others are better for one-time purchases. And don't forget about other avenues, where you can use the voice interaction to drive your customers to another platform where you can monetize, such as a mobile app."*

Sarah Andrew Wilson, Chief Content Officer, Matchbox.io, @SarahAndrewWils

When viewers started recording shows and fast-forwarding through commercials or choosing streaming services to avoid commercial breaks, TV advertisers quickly pivoted from 1-2 minute commercials that interrupt the storyline of a show to paying for product placements that show up in the scenes of those shows. In the world of voice AI, the same level of innovation will need to happen to move advertising from an unwanted interruption to a natural and pleasant part of the voice experience.

Instead of voice ads that barge in on conversational experiences, interactive voice ads are introduced as part of the conversation. Similarly, other voice commerce opportunities, like those recently announced by Snapchat, enhance the user experience by allowing them to identify items they are already selecting to photograph and helping users to discover where to purchase them, if they so desire.

These monetizable moments are not only unintrusive, they are actually helpful and are viewed by the voice user as beneficial. In the future, having these kinds of valuable interactions may lead consumers to engage with

the companies that provide relevant and timely purchasing opportunities over those that do not.

Unintrusive Voice Commerce Models

Reducing friction in commerce interactions began as early as 1906 when Charles Kettering designed the world's first electric cash register. Since that time, internet access and the rise of more ways to conduct business, including mobile devices, POS systems, and touchless payments, have opened up monetization opportunities beyond the physical transfer of money. The benefits of speed, convenience, and ordering ease continue to drive innovations for commerce and purchasing opportunities of all types. These key benefits are also propelling the trends in voice commerce, along with an increased focus on providing purchasing opportunities that are

both unobtrusive and effective.

When considering a voice commerce model, voice AI developers and business leaders agree that continuing best practices for voice user interface design and adhering to the guidelines already established for their branded voice experiences are paramount to success. For most companies, the goals associated with their voice assistant strategies adhere to a common set of objectives that begin with engagement and end with brand loyalty. These goals are built on five pillars of the user experience that support a successful voice assistant strategy.

The core goals of any voice experience include five main pillars:

1 Entice users to engage with the voice assistant

2 Make the experience pleasurable so they return

3 Expand usage across channels (omnichannel voice experiences)

4 Encourage increasing use cases amongst users

5 Create brand evangelists who consider your company their go-to for all their relevant needs

Voice commerce strategies must follow these same tenets of good UX design to encourage continued and expanded user engagement while meeting the revenue goals of companies. Responsive and proactive suggestions that become monetizable moments with purchasing opportunities for users based on their location, context, and perceived needs accomplish both business goals—bottom-line contributions and positive customer satisfaction ratings.

"Monetization around an AI-based voice assistant, will surely need reflection around defensibility of the AI brick, mastering the AI inconstancy, and clearly understanding the role of the AI agent within the process."

Joris Guerry, Co-founder & CTO, Rumble Studio, @jor_gu

Responsive Voice Commerce

Voice interface users are constantly communicating their needs by asking questions, and voice assistants are already providing results based on pre-set criteria. For most interactions, the results returned by the voice assistant are those considered "most relevant'' to the question or request. However, relevancy can still be maintained and a preference given to a business that has signed on as a sponsor. When the voice assistant responds with suggestions based on the user's needs and prioritized based on a subscription model for service providers, monetizable moments are created.

Imagine this scenario: while driving, someone in the car might wake up the voice assistant and say, "I'm hungry". When the natural language voice assistant has cloud connectivity, it responds with a shortlist of nearby restaurants. Savvy restaurateurs know that the first few choices presented through the voice user interface are those most likely to be selected.

Now, suppose those recommendations were based on a combination of meeting the user's needs and surfacing results based on a subscription program. When the user selects the prioritized business listing and either orders food to-go or makes a reservation, the car manufacturer can share in a percentage of the food order revenue. Meanwhile, the restaurant is enjoying more food orders and a growing customer base, making the revenue share a low-impact marketing cost.

"This is what most brands fail to see: the abundance of conversational channels and the sheer scale of positive conversations that they could be having that either directly generate revenue or that move customers along the funnel seamlessly."

Kane Simms, Co-Founder, VUX World, @kanesimms

In this monetizable moment, the voice assistant has unlocked the revenue potential for both the carmaker and the restaurant, while the user experiences a helpful suggestion that feels more like a personalized moment than an advertisement.

In this model, any number of service or product providers could opt-in without the up-front costs associated with traditional advertising and pay the hardware manufacturer only after a transaction has been completed. Managed through the voice AI platform provider, voice AI-generated payments would be distributed to both entities automatically as transactions happen. Meanwhile, the users can get their needs met quickly, conveniently, and hands-free.
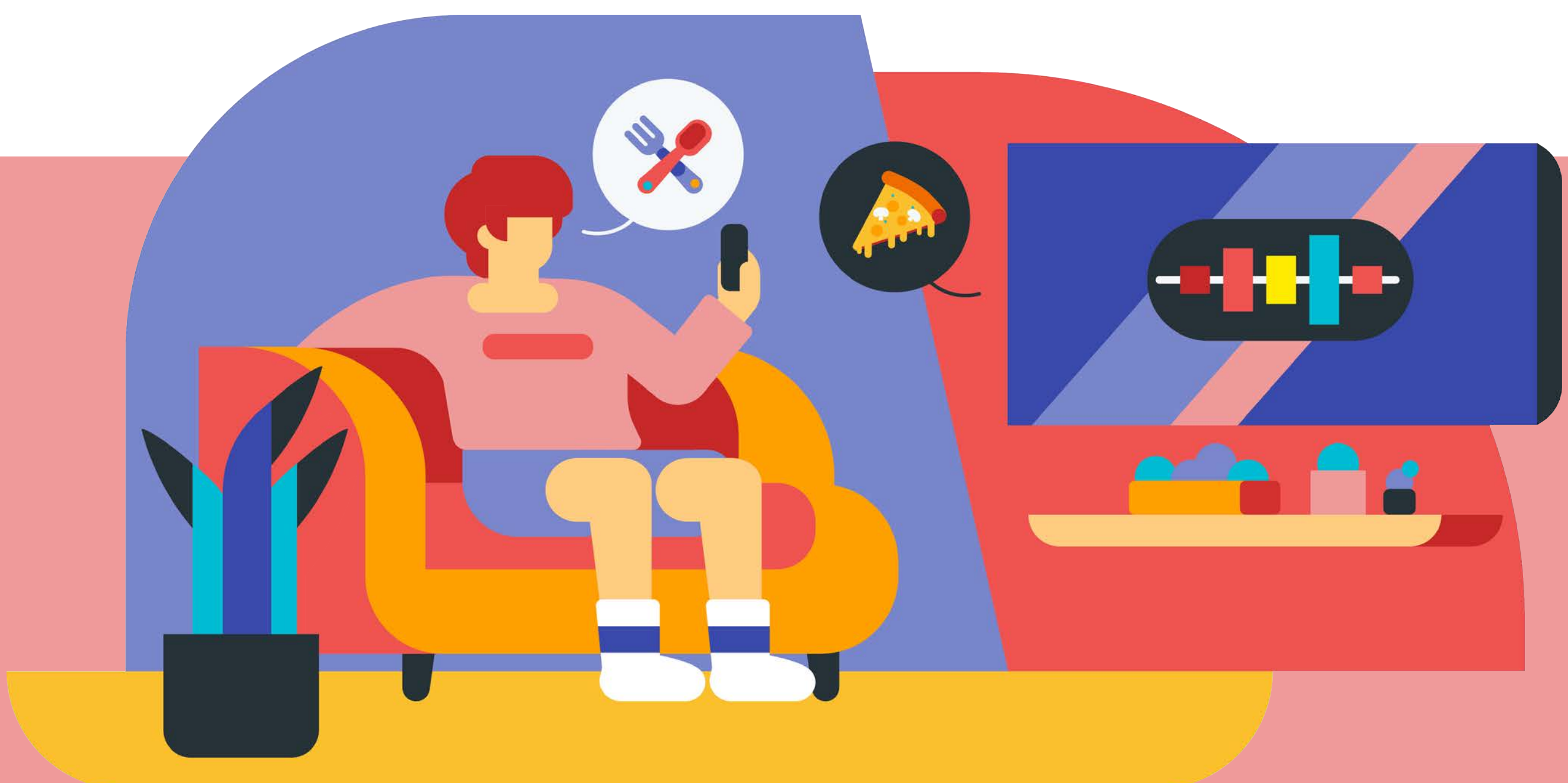
Proactive Voice Commerce

Sometimes, your voice interface user isn't yet aware of a need. In that case, the voice assistant can step in with a suggestion in anticipation. These monetizable moments are based on the AI abilities of the voice assistant to understand the context of a question and make relevant suggestions.

Proactive suggestions can take place at any time and on any device. While traveling, a hotel guest may ask the in-room voice assistant for the weather report in a distant city. Given the context of traveling and the distant location, the voice assistant may deduce the guest is continuing to travel. After responding with the weather information, the voice assistant can follow up with a helpful suggestion, "Do you need travel or hotel accommodations?"

Airline partners in the monetization model benefit from having their flight times mentioned in the first results from the voice assistant. When the hotel guest chooses the partner airline's prioritized business listing, the hotel shares in the booking revenue, the airline has the benefit of avoiding the daily online price wars with competitors, and the hotel guest has secured the next leg of their trip. If lodging is required, the hotel has an opportunity to suggest their properties at the new destination, even when they don't share the same name.

> *"Starting with a posture toward providing the best voice-led customer service possible can help brands turn a reactive, potentially negative situation into a proactive selling opportunity."*
>
> Nick Rovisa, Director of Business Development, RAIN, @nr

Other examples of proactive suggestions include destination parking availability to drivers through the in-car voice assistant or triggering a grocery order when the smart refrigerator senses the owner is low on a key ingredient.

By understanding the context of the user, smart, conversational voice assistants create new revenue streams for manufacturers and service and product providers across a wide range of industries. Providing suggestions at just the right time increases the chances that voice AI users will initiate purchases and perceive them as helpful recommendations.

A New Revenue Stream

Unlike traditional advertising models where the money flows in one direction—from the advertiser to the entity providing the advertising space—voice commerce monetizable moments create revenue sharing between companies and opens up opportunities that wouldn't be available through traditional advertising models.
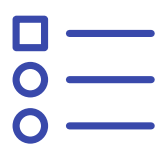
Revenue and marketing opportunities through monetizable moments include:

Revenue opportunities for hardware manufacturers

New marketing channels for service providers

Preferred business listings for service and product providers

Additional services and use cases for voice assistant users

Revenue sharing models allow content domain providers, retailers, and service providers to sponsor their offerings and receive prioritized listings when users are given purchasing choices—increasing sales and expanding the customer base. Even when the sponsored brand is not chosen, prioritized listings have the benefit of improving brand awareness for future purchasing decisions.

"Many large scale organizations are still trying to sort out their traditional monetization strategies, so it's our responsibility as a partner to educate them on how VOICE tech integrates into that roadmap, and empower them to have meaningful discussions with other stakeholders."

Brandon Kaplan, CEO, Skilled Creative, @skilledcreative

When the owners of voice user interfaces can participate in the revenue from the products and services suggested by their voice assistants, they're incentivized to integrate products and services that benefit their users, providing ongoing opportunities for service providers to participate in more voice commerce opportunities across industries, hardware, and users.

Unlocking the potential of reactive and proactive voice commerce opportunities hinges on one key element: voice AI technology. Without a conversational interface that understands context, has access to the cloud and a growing library of content domains, understands multiple languages and accents, and performs with the highest accuracy and speed, monetizable moments based on the subtleties of user needs are not possible.

Chapter Two

The Real Power of Voice AI for Unlocking Revenue

Chapter Two

The Real Power of Voice AI for Unlocking Revenue

According to a report by **Voicebot.ai**, 62% of voice app developers agree that building custom voice apps for businesses provides a clear path to monetization. These technology experts understand the importance of customization as an essential element of monetization.
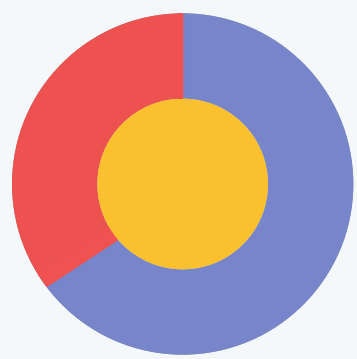
62%

of voice app developers agree that building custom voice apps for businesses provides a clear path to monetization

Voicebot.ai

The need for greater control over the customer experience and custom features and the realization that data ownership is key to personalization and monetization are essential factors propelling the conversion from voice experiences intermediated by a third party to custom, branded voice assistants. While a majority of brands with voice experiences began with skills and actions for Alexa or Google, more than half have deployed a voice assistant (64%) and 84% of those companies have a voice-enabled mobile app, according to
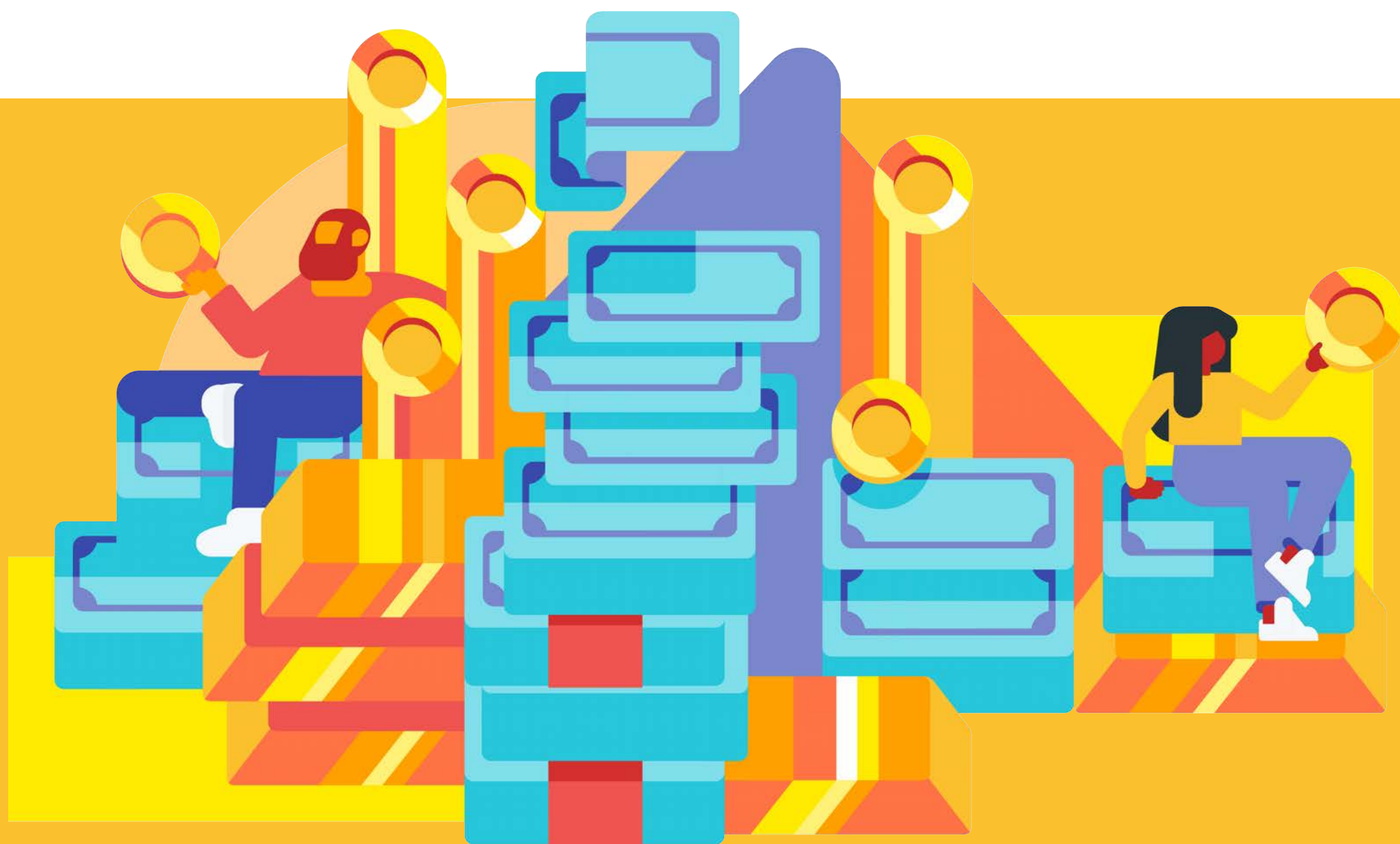
[Opus Research](). An additional 43.4% of companies have embedded custom voice assistants into their products and 33% have added customized voice AI to their websites, according to a study by [Voicebot.ai]().

67% of companies have deployed voice assistant technology

Opus Research
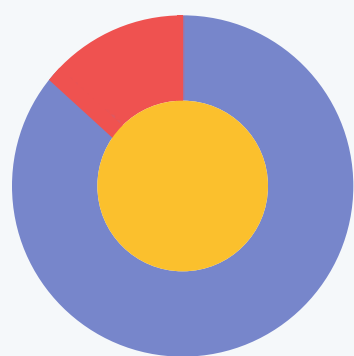
84% of companies with a voice assistant have a voice-enabled mobile app

Opus Research
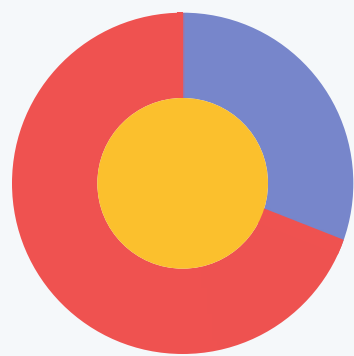
33% of brands have added customized voice AI to their websites

Voicebot.ai
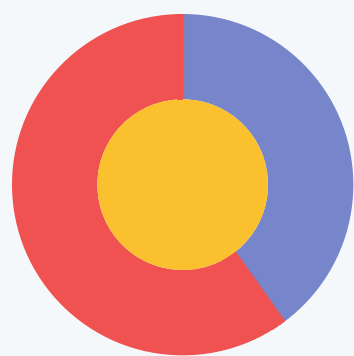
43% of companies have embedded custom voice assistants into their products

Voicebot.ai

As voice AI continues its rapid ascent from nascency to ubiquity, the technology continues to improve, monetization opportunities provide a return on investment in the form of revenue, and adoption of custom voice assistants will continue to rise.

The major driving force behind the shift to custom omnichannel voice experiences will be the quality of the voice AI technology powering the voice assistant. Without a robust voice platform built on natural language technology designed to grow and improve over time, the dream of voice-enabled customer interactions and voice commerce will remain unrealized for many brands.

Conversational AI: The Foundation of a Great Voice Experience

At the core of every good voice experience is a conversational voice AI platform built on advanced technology that delivers highly accurate and fast results. Across industries, business leaders overwhelmingly agree (97%) that accuracy and speed are the top success metrics for voice assistants, according to Opus Research. That's not surprising, given that delivering results is the function of a voice assistant and the very foundation upon which voice experiences are built.

97% of business leaders agree that accuracy and speed are top success metrics for voice assistants

Opus Research

Although conversational intelligence is the goal, many voice assistants fall short due to their lack of real understanding. In the typical voice AI model, th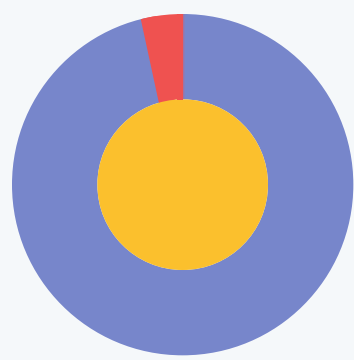e sound of the user's voice is turned into data to communicate with the digital assistant, once the message has been received and the correct data has been retrieved, the computer language is turned back into sound—a two-step process.

"Successful voice assistants create a platform from which companies can upsell. As conversational AI decreases transactional friction, more space opens up to increase per-user revenue."

Bradley Metrock, CEO, Project Voice, @BMetrock

In an advanced model, like Houndify's Speech-to-Meaning® technology, the Automatic Speech Recognition (ASR) and Natural Language Understanding (NLU)

technologies are combined to understand speech at the speed of thought. Just like the human brain, advanced voice AI technology begins listening and processing even as the person is speaking and delivers an accurate result immediately.

The ability to understand quickly and precisely opens up the possibility of establishing a conversational tone with the user. When a conversation can flow naturally between the voice assistant and the customer, the friction of waiting for inaccurate replies and the need to repeat the same request multiple times is removed. Users will gravitate toward the frictionless interface as a truly convenient way to complete a task or get information.

"Voice is not only a channel, but a UI. From a UI perspective VOICE tech can simply represent an opportunity to make the monetization process more efficient or more engaging."

Brandon Kaplan, CEO, Skilled Creative, @skilledcreative

Add to that, the ability to listen and respond to multiple questions or multi-layered requests, and the human-to-machine interaction becomes as natural as speaking to another human. When actual conversations are possible, the voice assistant can become a trusted companion and personal helper for your customers—opening the door to more types of interactions.

When the voice assistant can accept compound and complex queries, conversations take on a new dimension.

For example, users can find just the right place to eat the first time they ask:

Beginning with:

"Find me Asian restaurants, except Chinese and Japanese, within 2 miles of Union Square, with at least 3 stars, that are currently open."

And follow-up with:

"Which ones are kid-friendly and have free parking?"

Or ask questions about two related topics:

"Who's leading the National League West and how many home runs have the San Francisco Giants hit this year?"

Regardless of the use case or context, customers who are able to interact with a voice assistant that is at least as smart as a human are increasingly likely to find more things to ask for and more ways to engage.

During normal discourse, people will often interrupt each other to offer clarification. Communication doesn't always happen turn by turn with each person waiting for the other to state all they have to say on a topic. Creating natural conversations that allow for interruptions, or "barge-ins," allow users to forgo listening to more information or to change the topic once they have everything they need.

The ability to listen while talking isn't always easy for humans, but it should be second nature to a mature voice assistant solution.

Once the foundation of the voice assistant is established on an advanced voice AI platform that offers conversational interfaces and fast, accurate answers, brands are ready to consider moving to the next phases of their voice AI roadmaps. For most companies, the next steps include omnichannel experiences, increasing user engagement with more use cases and more channels, developing multimodal experiences, and finding ways to monetize the investment.

The statement that best describes the future outlook for voice assistants among business leaders:

Increase breadth of use cases supported by voice

Expand voice into a multimodal experience

Find ways to monetize voice solutions

Gather more user data to inform the roadmap

Have a consistent customer experience across the company

Increase number of channels that are voice-enabled

Opus Research

Personalization is the Key to Monetization

Personalized voice experiences can be helpful and satisfying to users who want a restaurant voice assistant to remember their name, payment method, and past orders based on a phone number or other identifying feature.

In other instances, companies employing voice assistants may want to implement an "opt-in" system where users can agree to certain information being collected and stored to make future interactions more personal.

"One of the roadblocks to successful monetization of voice AI is the ability to personalize the experience. Businesses with contact centers possess the data to help overcome this, which is why we've seen high adoption rates in these types of organizations."

Bradley Metrock, CEO, Project Voice, @BMetrock

Personalization features that feel too intrusive or are activated without user consent risk negative reactions and fears about invasions of privacy. Companies that can establish trust with customers through transparent data collection, usage policies, and the implementation of custom, branded voice assistants that are not intermediated by a third-party platform provider, have an opportunity to improve voice assistant experiences, expand use cases, and create revenue opportunities.

Familiarity with the user allows the voice assistant to address a specific speaker by name, differentiate that person from others in the household or car, and provide results and suggestions based on who they are, where they are, and what they are doing at the time of the voice interaction.

Voice profiles can also be created when a customer calls or interacts with a company for the first time. Basic information such as name, personal preferences, and contact information can be stored and activated when that person re-engages with the same business.

In a practical application, a voice profile could sound something like this:

"Thanks for calling Dr. Burgers. I'm your automated ordering assistant. Your phone number is in our system. Am I speaking with Kevin?"

"Yup, that's me!"

"Hi Kevin, are you calling to place an order for delivery or pick-up?"

"Delivery please."

"Sure, are you still located at 555 First St. in San Jose?"

"Yup."

"Great, would you like the usual?"

"Yes, that'd be great."

"Ok, I have a BBQ Bacon Burger with no Tomato and a chocolate milkshake. Anything else?"

"No thanks!"

"Ok, I've submitted your order and it will be delivered within the next 45 minutes."

Voice profiles have other use cases that help customers speed their interactions with contact centers and other customer support and account-based applications. Voice biometrics offer many of the same benefits of voice profiles when implemented into products and devices.

"Biometrics will open new doors for what voice can do in certain verticals, such as banking and healthcare, where authentication is a requirement."

Bradley Metrock, CEO, Project Voice, @BMetrock

The process of implementing voice AI within business operations or marketing endeavors forces an organization to define key attributes about itself, sometimes for the very first time.

Voice assistant users can opt-in to setting up a voice profile for their car or any personal or smart home device. In some cases, voice biometrics can be coupled with other forms of identity verification to create multimodal authentication and provide users with an additional layer of security.

Using a unique voiceprint, voice assistant users can shortcut their interactions with a voice user interface, further increasing convenience and
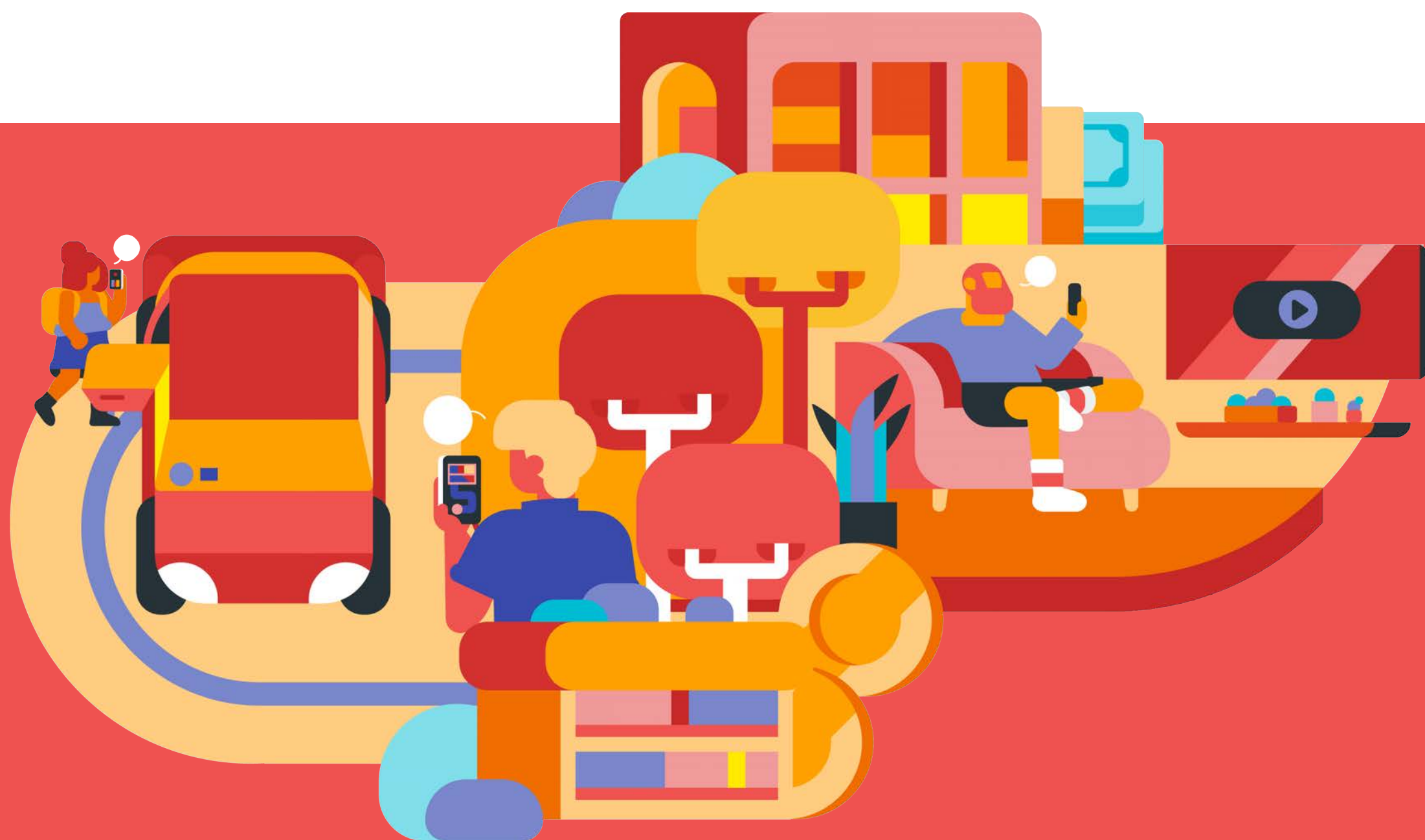
speed of service. Imagine the delight when a driver gets in the car and after speaking a passphrase, the seat, steering wheel, and mirrors automatically adjust to preset positions while the voice assistant offers to recant the day's calendar events, an alternative route to work based on traffic flow, and a reminder. Depending on the level of opt-in by the driver, more conveniences can be presented, including

finding the nearest gas station and paying for gas through a digital payment system.

Companies that build trust with their customers and offer personalization that is secure and safe can provide more services to those customers through helpful suggestions and context-based opportunities.

"Trust is a key factor. When relying mainly on voice interaction, customers need to know exactly what they're getting and that their selections and responses have been heard correctly."

Sarah Andrew Wilson, Chief Content Officer, Matchbox.io, @SarahAndrewWils

The buying journey can start at various touchpoints. Knowing who the person is, where they are, and understanding the context of their conversations with the voice assistant are critical to providing buying opportunities relevant to their needs. Without those elements, suggestions might be perceived as annoying or intrusive.

A key to this level of customer awareness is a voice AI platform that's context-aware and can remember the beginning of a conversation in order to continue to offer relevant responses as the conversation continues—without requiring the user to repeat key information.

Multilingual Voice Assistants for Global Enterprise

Globalization of marketing and sales efforts has always been about a lot more than simply translating a website or having bilingual employees. Successful global organizations have

localized their messaging, products, and services to meet the needs and answer the questions their customers are asking in their own languages and geographic locations.

The trajectory of global voice assistants needs to follow the same path. The first step is to ensure your voice assistant can speak your customers' native language. If your voice user interface doesn't support the languages spoken by your customers in various geographic locations, you risk lack of adoption and engagement.

Beyond simple translation, implementing a multilingual voice assistant requires that it be trained using the conversational language of users in the target market. Cultural norms and values should also be taken into consideration to determine the persona, gender, voice, and tone of the voice user interface.

The sheer volume of data and time requirements make developing multilingual voice assistants prohibitive for internal development teams. That's why it's important to partner with a voice AI technology platform provider that has already tackled the challenges of multiple languages and has the foundational models upon which to build more.
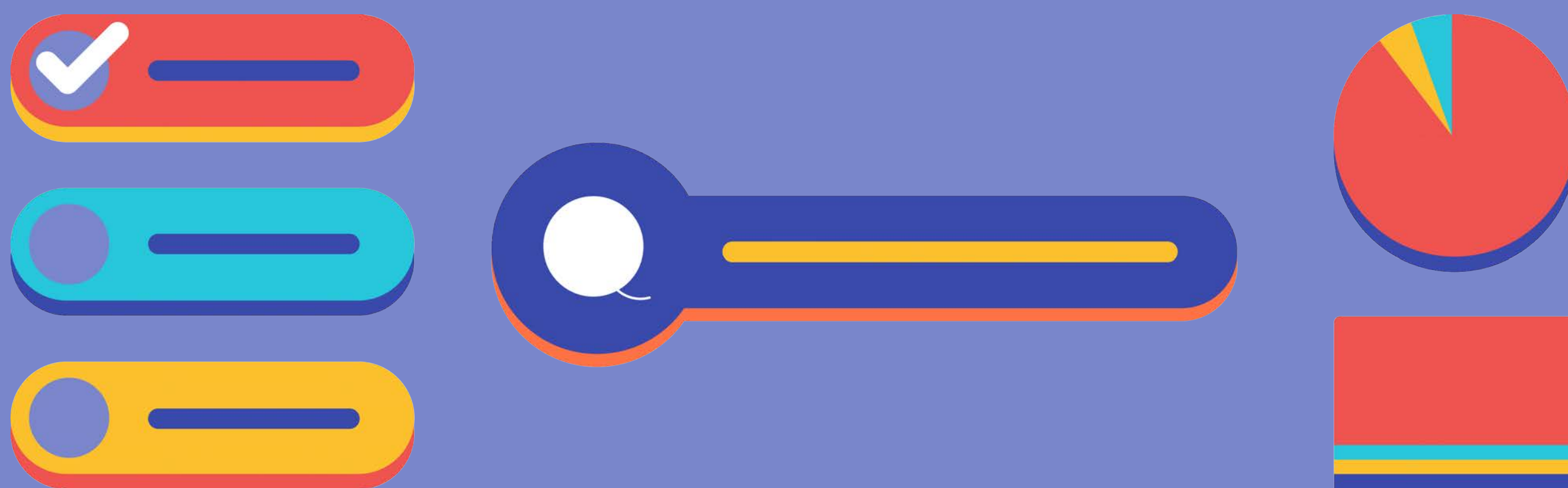
Solving for languages in a voice assistant is more complicated when Natural Language Understanding (NLU) and Automatic Speech Recognition (ASR) live as two separate components of the same system. In this traditional voice AI technology configuration, it takes a long time to build the query-understanding module for a new language and there are many opportunities for errors.

Voice assistant accuracy can be impacted by accented and imprecise speech. Users, unable to communicate with a voice assistant because of speech differences or accents, become frustrated quickly and tend to abandon voice interfaces. When we interact with each other, we don't expect to all sound alike, yet voice assistants often fail to understand the variety of accents and diverse lexicon of their users.

Companies seeking monetization opportunities must employ voice assistants that will help them break down the language barriers between the voice AI and its users. A voice AI platform with a

broad range of available languages and training models that include a variety of genders, accents, and speech patterns is critical.

In addition, a large library of content domains will help the voice assistant to understand context, thereby improving accuracy.

"Whether you are in the process of implementing your strategy or just getting started and trying to define a business case, if you think about the pain points that customers have throughout the customer journey, the kind of conversations you can have that'll help, and the number of touch-points that could be voice-enabled, for many brands, even a minor percentage increase in sales will be more than enough to justify the investment. The business case will write itself."

Kane Simms, Co-Founder, VUX World, @kanesimms

Accessibility to products and services for a wide range of customers across the globe includes offering the voice user interface in their primary languages and delivering a voice assistant that can understand and respond to requests, no

matter how they're spoken. Depending on the use case, companies might also want to consider offering multimodal interfaces to enhance convenience and accessibility.

Multimodal Interfaces Enhance Convenience and Accessibility

Increasingly, voice user interfaces are part of a multimodal strategy for companies looking for more ways to improve customer experiences and increase accessibility to their products and services. Voice-first is a call to action for brands across industries to add a voice assistant to their products and services roadmaps—but, voice-first

has never meant voice-only.

Among the top three end-user benefits of voice-enabled services, products, and apps are: hands-free access, greater convenience, and the ability to multitask, according to the Opus Research Report. The ability to

communicate is at the core of what makes us human, and voice represents one of the most natural and intuitive methods to interact with the digital world. The ability to simply ask for the things you want has increased accessibility for those who have been disenfranchised by technology due to differences in ability by removing the physical and cognitive barriers other forms of interfaces create.

"Like any new interface or technology, monetizing Voice AI can be a challenge, because the general public isn't yet used to purchasing without touching, clicking, or swiping. But multimodal voice interactions are the future, and soon people won't think twice about making a purchase using their voice."

Sarah Andrew Wilson, Chief Content Officer, Matchbox.io, @SarahAndrewWils

The rapid advancement of technology has created an appetite among consumers for greater convenience and faster resolution of problems and tasks—further propelling the voice AI market into exponential growth. According to [Meticulous Research](#), the voice and speech recognition market is expected to grow at a 17.2% compound annualized rate to reach $26.8 billion by 2025. As the market grows and voice assistants become a common feature, consumer expectations and voice AI technology are reaching equilibrium—setting the stage for increased use cases and omnichannel experiences.
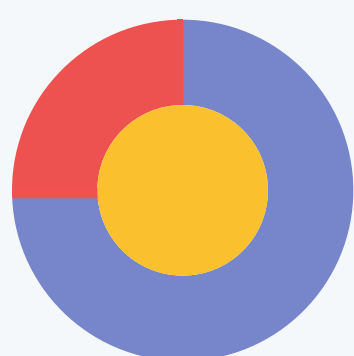
75%

of business leaders rate hands-free access, greater convenience, and the ability to multitask as the top end user benefits from voice assistants.

Opus Research

17.2% Rate

$26.8 Billion

The voice and speech recognition market is expected to grow at a 17.2% compound annualized rate to reach $26.8 billion by 2025.

Meticulous Research

Beyond the voice user interface, multimodal experiences ensure that customers have access to information, services, and product functionality when speech is not possible or not the most convenient form of interaction. Visual wake, gaze, gesture, and of course, touch are all within the realm of possibilities.

Offering a variety of methods to interact with your products, services, and apps not only increases convenience for your users, but also creates inroads for new customers who might not otherwise interact with your brand.

Knowing your audience and understanding the underserved population that would be interested in your offerings is key to determining the role of multimodal interfaces for your brand.

When considering monetization opportunities, it's important to remember that the more points of access you provide your customers, the more opportunities you'll have to encourage engagement.

The real power of voice AI to unleash revenue opportunities lies in the technology itself. If your voice experience doesn't deliver on the promise of accuracy, speed, convenience, and accessibility, efforts to monetize on the experience will fail to perform. On the other hand, building your voice experiences on a foundation of advanced AI that's context-aware, multilingual, customized, and conversational will allow you to expand your use cases and provide voice experiences that delight and deliver revenue.

"Voice AI can be slow and costly to develop yourself, and can easily become an obstacle to monetisation of your product. Don't assume voice AI is required from day one, or that you have to build everything in-house. Carefully consider at what stage of your product development you really need voice AI-powered functionality, and explore the fast-changing voice AI market for existing solutions."

Carl Robinson, Co-founder & CEO, Rumble Studio, @VoiceTechCarl

Chapter Three

Key Industries on the Cusp of Monetization

Chapter Three

Key Industries on the Cusp of Monetization

The Gartner Hype Cycle has become a popular model for explaining the current state of the voice AI industry. If we follow the trajectory of voice technology in the context of the model, we believe product-based custom voice assistants that add value to the core user experience are passing through the "Trough of Disillusionment" and are entering the "Slope of Enlightenment".

The "Innovation Trigger" occurred about 7 years ago with the introduction of the first voice-enabled home speakers. The promise of a new technology that would instantly catapult the world into a Star Trek-like reality—where people could interact with everything in the digital world using only the sound of their voice—created a peak of inflated expectations.
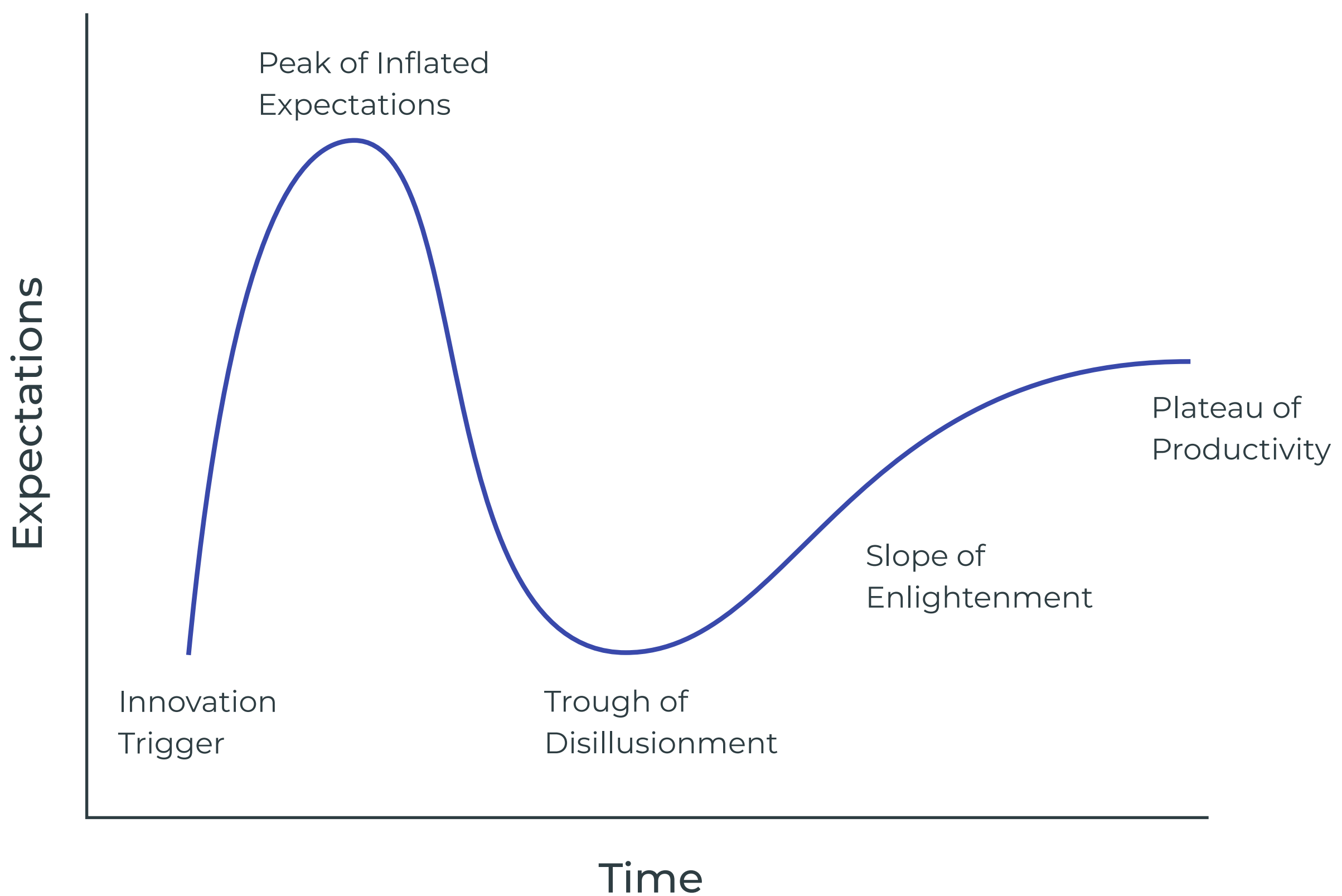
Gartner

Soon after the first wave of voice-enabled smart speakers and mobile phones, the challenges of developing and implementing a viable voice

user interface for products and services created a "Trough of Disillusionment" amongst companies across industries. In many cases, it took

companies 3 years to discover that it would take at least 10 years to develop the technology on their own.

Simultaneously, the reliance on existing smart speaker product offerings as the only vehicle to provide voice user interfaces caused problems for companies who risked losing customer relationships and trust from a lack of control over the experience and data. Privacy concerns slowed many voice initiatives and costs remained too high for medium and small businesses to participate. Consumer disillusionment over voice experiences that failed to deliver fast and accurate responses held back rapid deployment of "just ok" voice user interfaces.

"Even if a voice experience isn't on a brand's immediate roadmap, there are steps brands can take to better prepare themselves for monetization when that time comes. Assessing readiness of their payment systems, shipping services, and any other platforms or APIs that may influence (or delay) monetization features and having a pulse on the current state of their technical architecture can help brands identify potential roadblocks and save them time down the road."

Elissa Dailey, Director of Strategy, RAIN, @RAINAgency

Over the next couple of years, education and the realization that strategic partnerships with voice AI platform providers could help brands put voice assistants back on their company roadmap gave new hope for the future of voice AI, and the voice-first era was launched in earnest. The "Slope of Enlightenment" phase was also a time of reflection on the purpose and promise of voice AI. Difficult discussions brought to light the need for data privacy and custom voice assistants that could reflect the values and personalities of the brands they represented.

Companies began to understand the importance of voice AI platforms that could provide conversational AI, remember the context of conversations, and eliminate the need for users to memorize specific scripts to ask for what they want.

As more and more companies enter the voice-first era with voice-enabled mobile apps, services, and hardware devices, expectations have begun to meet technological advancements. As we enter the "Plateau of

Productivity", many companies have addressed the challenges of the earlier stages and are looking for omnichannel voice AI solutions that generate revenue. For those industries and companies that were early adopters, the time has come to expand voice experiences to include more use cases on more channels with greater opportunities to monetize their voice AI investment.

"My advice for brands looking to build a business case around monetising voice AI is to do the same thing as you did when you first started implementing voice AI: start small, take baby steps, little and often."

Kane Simms, Co-Founder, VUX World, @kanesimms

Industry norms and market competition are the drivers that have determined the general rate of adoption of voice AI solutions for companies in specific sectors. Some industries, such as automotive, saw early adoption and are moving quickly to the next phase. Within industries, the rate of adoption varies—with early adopters and laggards existing side by side. Despite the outliers, there are some industries that, in general, can be characterized as in the "Plateau of Productivity" phase and are ready for the next phase of voice AI—omnichannel experiences, expanded use cases, and monetization opportunities.

"Of the many industries that are ready to monetize voice AI, the podcasting and content marketing industries rank highly. Interactive bite-sized audio content, accessible on-demand, is becoming ever more popular with consumers. Brands that are quick to produce specialized audio content that answers long-tail questions posed by consumers via voice will enjoy a greater mindshare, and ultimately boost sales."

Carl Robinson, Co-founder & CEO, Rumble Studio, @VoiceTechCarl

Acknowledging the outliers within industries, we've identified four market sectors where companies are entering the next stage of voice assistant adoption and implementation.

1 Auto Industry

2 Smart Homes and TVs

3 Hospitality and Travel

4 Internet of Things (IoT)

Voice Commerce in the Auto Industry

As one of the earliest adopters of voice AI, the auto industry continues to push boundaries and find more ways to improve in-car experiences through voice technology. Their focus on safety through hands-free operation of cockpit controls and access to phone calls and texts began a journey into voice user interfaces that has grown to include driver and passenger entertainment systems and a vision of the car as an integral part of the Internet of Things (IoT).

As early as 2018, connected commuters spent $230 billion, according to PYMNTS. The same study revealed that 35.3% ordered food, 33.4% ordered coffee, 14.9% ordered an item to pick up at a store, and 14.5% ordered groceries during their commutes.

$230 Billion

In-car shoppers spent $230 billion in 2018

35.3% ordered food and picked up at drive-through

33.4% ordered coffee and picked up at drive-through

14.9% ordered an item to pick up at a store

14.5% ordered groceries to pick up

PYMNTS

Consumer demands for connectivity and the ability to be more productive during their commutes are driving car manufacturers to make ecommerce easier, more convenient, and safer through the implementation of in-car voice assistants. Connected cars equipped with advanced natural language capabilities and the ability to understand the context of conversations will soon provide hands-free purchasing opportunities for drivers through cloud-based voice interfaces.

Initially, in-car voice assistants were reserved for luxury vehicles and top-of-the-line offerings from car manufacturers. The introduction of Hyundai's Dynamic Voice Recognition System in the 2021 Elantra and Elantra Hybrid models in North America marked the first time a voice user

interface was made available in everyday vehicles, a trend which is expected to continue as consumer adoption persists.

In its In-Car Voice Assistant Report, Voicebot.ai reported that nearly twice as many U.S. adults have used voice assistants in the car (114 million) as through a smart speaker (57.8 million). Considering that the average commuter spends nearly an hour driving each day, it's not surprising that people are looking for ways to make this time more productive.

2X More

Twice as many U.S. adults have used voice assistants in the car as through a smart speaker

Voicebot.ai

Conversational voice assistants allow car manufacturers to respond to the growing consumer demand to make purchases in-car while maintaining their goal to make driving safer by reducing distractions and allowing drivers to keep their hands on the wheel and eyes on the road.

A conversation between the in-car voice assistant and a driver may sound like this:

"What's the typical weather in Seattle for July?"

"The usual weather is a high of 73 degrees and a low of 56 degrees in Seattle, Washington in July. Would you like to book travel there?"

"Yes. Find me a hotel checking in July second and checking out July fifth for less than 300 dollars a night, with at least 4 stars and free breakfast and WiFi."

"Downtown Hotel has 4 and half stars and is offering a Fourth of July special for 3 nights at 180 dollars per night, has free continental breakfast and complimentary WiFi. Would you like to book a room?"

"Do they have a room with a king bed available?"

"Yes, there are two rooms with king beds currently available. Would you like to book one now?"

"Yes, please."

"OK. Anything else?"

"No thanks!"

"Great. I've booked your room and sent a confirmation to your email."

The infrastructure is in place and consumers are ready for the next phase of in-car voice commerce opportunities. Soon, drivers and passengers will enjoy proactive suggestions and helpful responses when a query is made into ordering availability. When conversational voice AI is combined with strategic partnerships, the possibilities are endless.

Smart Homes and Smart TVs

Smart TVs are emerging as one of the most popular solutions to meet the demand for multimodal interfaces. While voice allows users to quickly navigate complicated menu options, the display screen provides a visual representation to help consumers make informed buying decisions. In homes and hotel rooms, televisions can be multipurpose devices—delivering entertainment, controlling other devices in the room, and offering easy purchasing opportunities.

Juniper Research predicts that more than $164 billion in transactions will be conducted via smart home devices by 2025, representing more than 630% growth in five years from the $22 billion in voice commerce in 2020.

$164 Billion

in transactions via smart home devices predicted by 2025

Juniper Research

630% Growth

in voice commerce via smart home devices from 2020 to 2025

Juniper Research

Not surprisingly, smart televisions are taking the lead as revenue generating smart devices. Juniper predicts that more than a fifth of the 2025 value, almost $33 billion, will flow through smart televisions.

Since their inception, TVs have had multiple purposes. As a key source of news and information, TVs allowed families to keep up with current events and understand the world around them. Television was initially a source of

viewing-only entertainment and a place where families might gather to watch a movie together. VCR and DVD players further enhanced movie-watching experiences as people could more easily choose what they wanted to see and when they wanted to see it. As it began to evolve, television also served as a game console, an internet streaming device, and an extension of the computer screen.

$33 Billion out of $164 Billion

$33 billion of the $164 billion in smart home purchases will occur on smart TVs
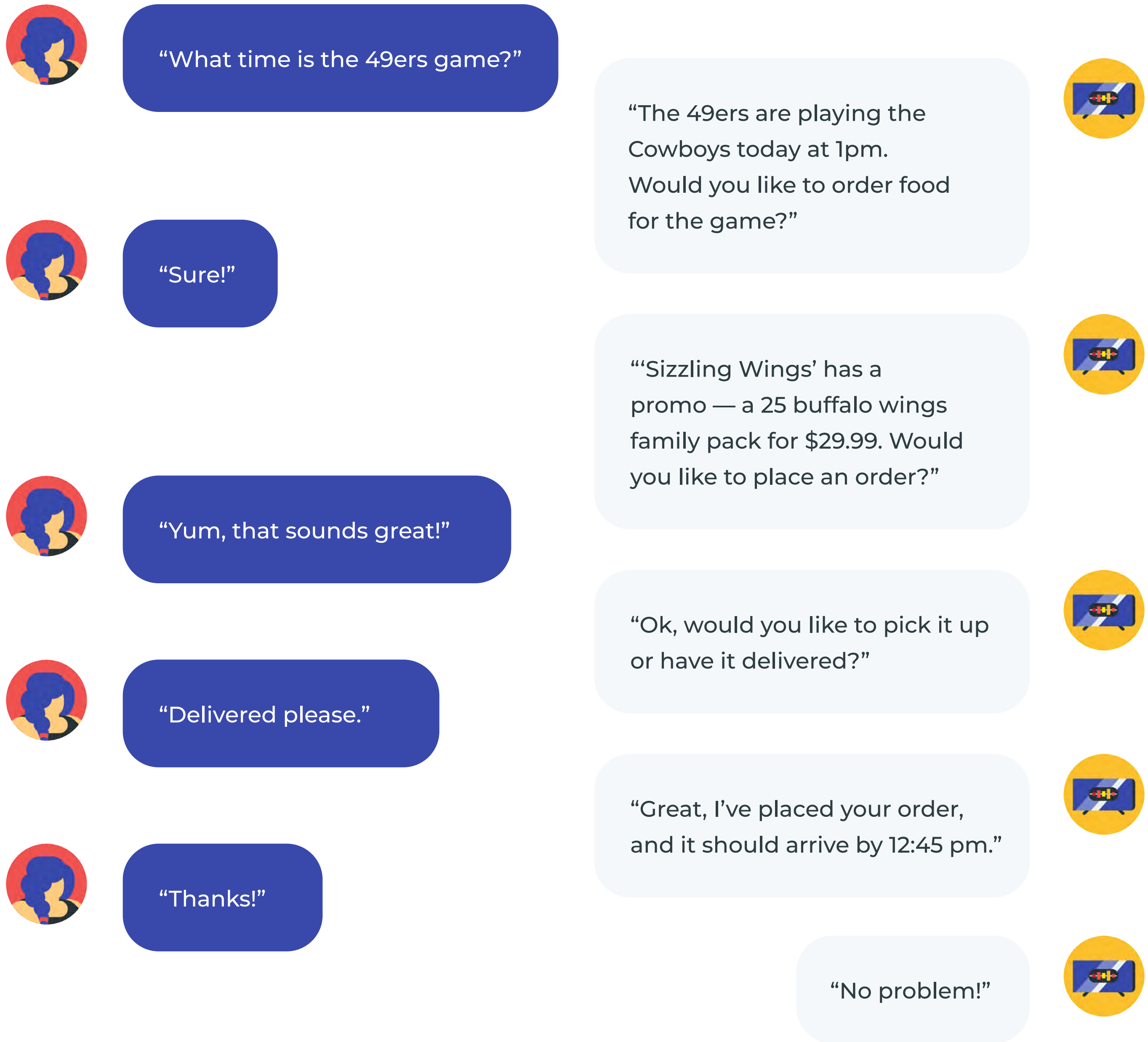
Juniper Research

"The CPG and Media space are two areas that we have hyper focused on, and where we see some of the largest growth. CPG organizations are sophisticated in terms of monetization strategy, and are very inclined towards looking forward to the next user journey touch point. Media organizations, like streaming services or entertainment publications, are looking for unique ways of generating a path to purchase."

Brandon Kaplan, CEO, Skilled Creative, @skilledcreative

Now, as an integral part of the IoT, voice-enabled television sets are poised to be major players in the voice commerce economy. Smart TVs combine the convenience of making purchases from the comfort and privacy of home with a connection to the cloud—opening the door to monetization partnerships across a wide range of product and service providers.

Here's an example of a possible voice commerce exchange via a connected TV:

"What time is the 49ers game?"

"The 49ers are playing the Cowboys today at 1pm. Would you like to order food for the game?"

"Sure!"

"'Sizzling Wings' has a promo — a 25 buffalo wings family pack for $29.99. Would you like to place an order?"

"Yum, that sounds great!"

"Ok, would you like to pick it up or have it delivered?"

"Delivered please."

"Great, I've placed your order, and it should arrive by 12:45 pm."

"Thanks!"

"No problem!"

Besides their role in the home, voice-enabled TVs can act as hotel concierges in the hospitality industry and smart room hubs to control lights, heating and A/C, curtains, and other devices.

Hospitality and Travel

Guests in hotels are already making purchases from their rooms. Whether it's ordering food from room service or making hotel arrangements for the next leg of their travel, hotel guests have become accustomed to making purchases while away from home.

Voice-enabled devices, such as TVs, add convenience and speed to the purchasing process. Instead of stopping to search for the in-room menu and then looking for the desired room service items, guests can simply ask for what they want. Checking out and making more travel arrangements can also be completed hands-free—allowing guests to continue getting ready or packing while making the transaction.

In addition to the opportunities for guests to order room service or book a reservation with the hotel for the next leg of their travel, monetization opportunities allow hoteliers to partner with service providers outside their property and share revenue for services and experiences booked from the hotel.

Here's an example of an in-room voice assistant acting as a personal concierge:

"How far is it to the City Art Museum from here?"

"The City Art Museum is 3.5 miles away. Would you like to book an Uber?"

"How much would it cost?"

"An Uber-X can take you from here to the City Art Museum for 12 to 19 dollars. Should I book one for you now?"

"What about an Uber SUV?"

"An Uber SUV can take you for 21 to 30 dollars. Should I book one for you?"

"Yes please. How soon will it be here?"

"The nearest driver is 12 minutes away and will meet you at the front door."

"Great! Thank you."

"You're welcome!"

The COVID-19 pandemic has increased the desire for touchless interfaces and experiences that don't require people to handle menus, remote controls, light switches, and other devices. Voice assistants have the ability to provide touchless experiences for guests while delivering more opportunities for monetization to hoteliers and outside service providers.

"It's taken some time for consumers to trust their interactions with voice assistants. Sustained adoption over the pandemic has shown we've crossed that threshold. Now, companies have the latitude to build on top of that trust and try new things with conversational AI that push boundaries."

Bradley Metrock, CEO, Project Voice, @BMetrock

Internet of Things (IoT)

Traditionally, discrete IoT devices were controlled through a central hub, often a smart speaker or mobile device. While the convenience of having one device to control them all seems logical, the model is laden with challenges for users and device manufacturers. While it may seem convenient to have a centralized command and control center, IoT devices are rarely in the same room as the hub speaker. For instance, it would be extremely inconvenient to operate your coffee maker or washing machine from the living room.

For device manufacturers, giving up control of the voice experience to a third party can lead to a loss of customer loyalty and a gap in consumer behavioral data. In addition, exponential growth in voice shopping on devices in the home and the availability of hybrid voice AI technology are helping nudge IoT manufacturers away from intermediated voice experiences and toward adding custom voice assistants to their products.

"Don't underestimate the value of anonymised voice data captured by your voice-enabled product. Not only can it be used to improve the performance of the voice AI itself, but its analysis can provide insights into your customers and inform your product and marketing strategies."

Carl Robinson, Co-founder & CEO, Rumble Studio, @VoiceTechCarl

For consumers, growing comfort and familiarity with voice assistants coupled with the hands-free convenience and ability to multitask have increased the popularity of voice shopping in the home.

According to Statista, approximately 14% of U.S. consumers made purchases through voice-activated devices while doing routine activities like cooking or cleaning in 2020—representing a significant increase from 9.7% in 2018.

14%

of U.S. consumers made purchases through voice-activated devices while doing other routine activities
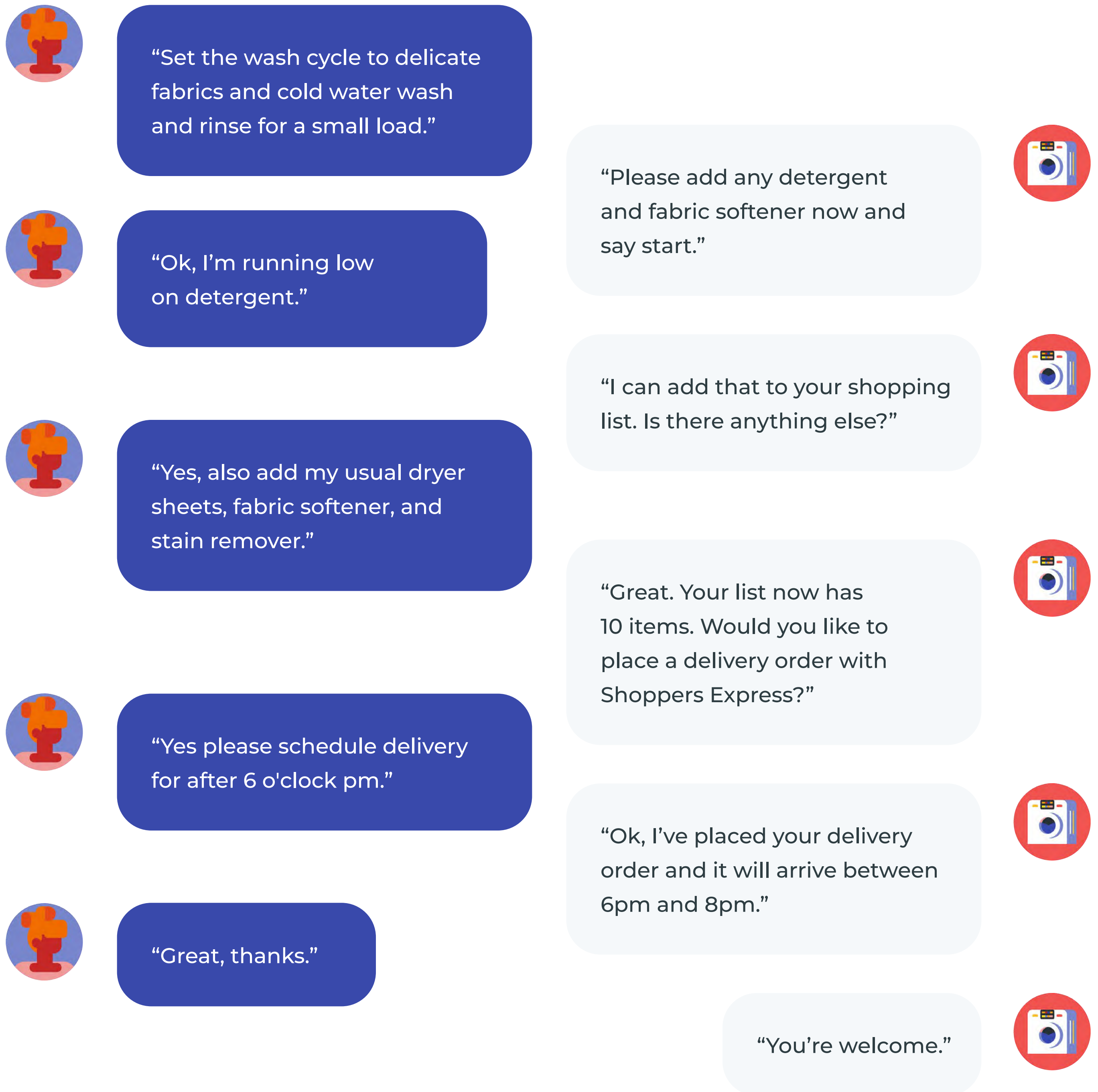
Statista

9.7%

more people in 2020 are voice shopping from IoT devices while multitasking than in 2018

Statista

Here's an example of a monetizable moment on a voice-enabled IoT device:

"Set the wash cycle to delicate fabrics and cold water wash and rinse for a small load."

"Please add any detergent and fabric softener now and say start."

"Ok, I'm running low on detergent."

"I can add that to your shopping list. Is there anything else?"

"Yes, also add my usual dryer sheets, fabric softener, and stain remover."

"Great. Your list now has 10 items. Would you like to place a delivery order with Shoppers Express?"

"Yes please schedule delivery for after 6 o'clock pm."

"Ok, I've placed your delivery order and it will arrive between 6pm and 8pm."

"Great, thanks."

"You're welcome."

Operational efficiencies, speed, and convenience for users will continue as IoT devices evolve from voice-enabled command and control experiences to channels for monetization and real revenue generation. Monetization opportunities for individual IoT device manufacturers will soon replace hesitancy over voice AI adoption due to uncertain ROI with enthusiasm for adding custom voice assistants to the product roadmap.

"Creating a voice experience that seamlessly integrates with existing order and payment platforms will lower the barrier for both company and customer. Customers should not need to link a new payment method to place an order through voice, and conversely companies should not need to reconcile voice orders outside of their core web and mobile payment platforms."

Nick Rovisa, Director of Business Development, RAIN, @nr

Conclusion

The Future Outlook for Voice Commerce and Monetization

Conclusion

The Future Outlook for Voice Commerce and Monetization

According to PYMNTS, "The future of voice commerce will depend on artificial intelligence (AI)." The article goes on to cite the necessity for voice AI technology that can handle complex queries and understand intent to meet consumers' rising expectations.
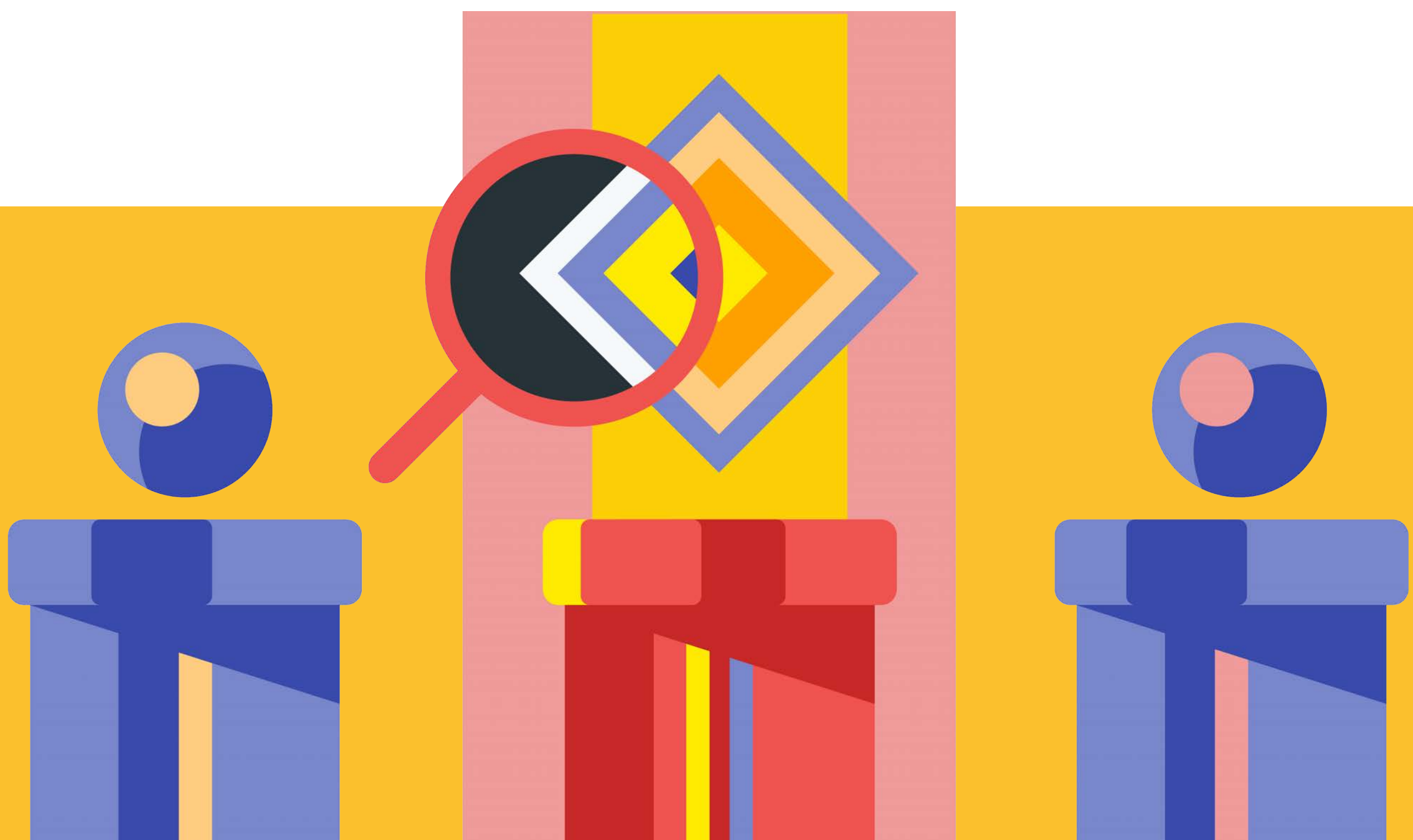
☀️👁 Artificial Intelligence

The future of voice commerce will depend on artificial intelligence (AI).

PYMNTS

Getting a piece of the $40 billion voice commerce market will require businesses across industries to overcome internal obstacles and move forward with aggressive voice AI roadmaps designed to meet the needs of customers now.

Voice strategies will also need to allow for expansion to address the rising demand to talk to the world the way we talk to each other, by simply speaking.

"Voice AI isn't a fad or something that is going to go away any time soon. Brands should be planning for the long term, and therefore in the short term, the real job is to build confidence and trust in users. You don't need to hit the jackpot on day one. Incremental revenue that demonstrates and builds user confidence is good enough for now."

Kane Simms, Co-Founder, VUX World, @kanesimms

Conversational voice AI interfaces are the key to creating buying scenarios that are both responsive and proactive. Meeting the consumer where they are, understanding the context of their queries, and responding with intent-based suggestions is the future of voice AI. That future

is already here and the businesses that can participate in the revenue share that will be available as a result of their investment in custom voice solutions will be those that continue to dominate their markets.

Hear what voice AI experts have to say about "The New Age of Voice Commerce"

Watch the Webinar

Get Started with Houndify

01   **Houndify**

02   **Blog**

03   **Resources**

04   **Sales**

HOUNDIFY

SoundHound Inc.

At SoundHound Inc., we believe every brand should have a voice. As a leading innovator of conversational technologies, we're trusted by companies around the globe, including Mercedes-Benz, Honda, Snapchat, Mastercard, and Deutsche Telekom. Houndify, our independent voice AI platform, allows developers to build custom voice assistants with a branded wake word for any product or service. Our proprietary Speech-to-Meaning® engine delivers unprecedented speed and accuracy. Our mission: To enable humans to interact with the things around them like we interact with each other— by speaking naturally.