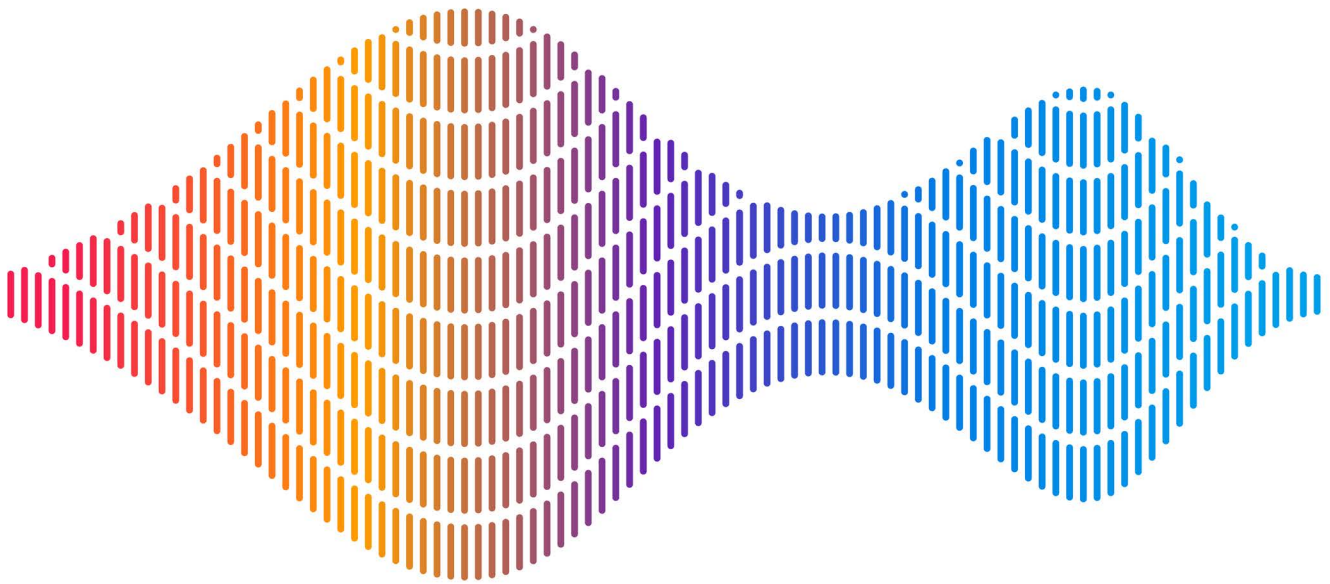


CONTACTBABEL

REAL DATA. REAL INSIGHTS.



The Inner Circle Guide to Voice AI

Research supported by





Purpose-Built **AI Agents**. One CX Platform.

AI Agents for Customers

Resolve more, end to end.

Deliver personalized, always-on support across voice and chat. Observe.AI AI Agents can understand customer needs, authenticate users, take action across your systems, and drive conversations toward resolution.

AI Agents for Frontline Teams

Give every employee an AI Companion.

Companion Agent listens, thinks, and acts alongside frontline teams—providing personalized context, real-time guidance, next-best actions, and automated assistance to help resolve complex interactions faster and more consistently.

AI Agents for Operations

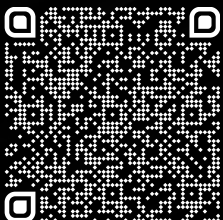
Turn every interaction into intelligence.

Evaluate interactions at scale, uncover insights, generate targeted coaching, and surface opportunities to improve performance across your entire operation.

One Connected System. Better Outcomes across CX.

Faster resolutions. More effective frontline teams. Full operational visibility.

Observe.AI connects AI Agents, human agents, and interaction intelligence so every conversation makes the entire system smarter.



Ready to Transform Your Customer Experience?

See Observe.AI in action.

The Inner Circle Guide to Voice AI (US edition)

Published September 2026

© ContactBabel 2026

Please note that all information is believed correct at the time of publication, but ContactBabel does not accept responsibility for any action arising from errors or omissions within the report, links to external websites or other third-party content.

CONTENTS

Contents.....	4
Table of Figures	7
Introduction and Methodology.....	11
What this report covers.....	11
Sources.....	11
How to read the segmentations	11
The Voice Channel and What "Voice AI" Means.....	12
Voice is not going away	12
The cost of a voice minute	13
What "voice AI" actually covers	15
Voice AI is not only self-service.....	17
The four layers of a voice AI deployment.....	19
Getting the Audio Right.....	21
Two trends working against clear audio.....	21
What poor audio costs.....	23
What can be done about it.....	24
Accent, language and comprehension	25
Voice Self-Service: From Containment to Resolution	28
How much self-service actually handles.....	28
Why sessions are abandoned	29
What voicebots can currently do.....	30
How much gets resolved without a human.....	32
When customers prefer to deal with AI	33
How customers rate AI against human agents	34
The concern that shapes everything else	35

When self-service does not succeed	37
What this means for containment targets.....	38
Voice AI Alongside the Agent.....	39
What agents actually spend their time on.....	39
The onboarding problem.....	41
What real-time assistance is measured to deliver	42
What agents make of it.....	43
What AI has done to headcount	44
Analytics, Quality and Assurance	45
How widely analytics is deployed	45
What analytics is used for	47
Why the quality process needs the help	49
Assuring the AI itself.....	51
Voice, Identity and Fraud	53
The scale of the authentication problem	53
Fraud, and the problem with using a voice as a credential	54
Proactive Outbound and the US Rules.....	56
The commercial case.....	56
A key distinction to understand	57
What the rules require.....	57
Where the exposure comes from.....	59
Enforcement and the direction of travel	60
Disclosure.....	60
Points a buyer may wish to consider	61
Implementation: What Breaks.....	62
Why voice is harder than text	62
What is holding contact centers back generally.....	63

Sourcing, and what got rolled back	64
What got stopped	66
What is blocking further adoption	67
Measuring Return, and What to Do Next	68
The building blocks of a business case	68
Performance against expectations	70
The right target, measured the wrong way	73
A sequence that the evidence supports	73
About ContactBabel	75

TABLE OF FIGURES

Figure 1: Inbound interactions by channel	12
Figure 2: Cost per inbound interaction	13
Figure 3: Historical mean call duration (service & sales), 2012 - 2025	14
Figure 4: The five levels of voice self-service technology	15
Figure 5: Current US adoption of voice self-service technologies	16
Figure 6: Current and planned use of voice-related AI applications	17
Figure 7: The four layers of a voice AI deployment	19
Figure 8: What % of your customers use a cell phone to call a contact center? 2017-25	21
Figure 9: The rise of remote & hybrid contact center working (2007-2025)	22
Figure 10: Customer issue: "I can't hear clearly and/or I have to repeat myself as they can't hear me" (by age range)	23
Figure 11: The unnecessary cost of mishearing	24
Figure 12: Current and planned use of real-time translation or multilingual support	26
Figure 13: Current and planned use of accent localization or accent softening	27
Figure 14: Overall proportion of calls handled entirely through self-service (only in respondents which offer telephony self-service)	28
Figure 15: Reasons for abandoning telephony self-service sessions	29
Figure 16: Capability of the most capable chatbot or voicebot in the organization	30
Figure 17: Controls applied to the actions bots can take in business systems	31
Figure 18: When contacting an organization, in which situations would you prefer to deal with AI rather than a human?	33
Figure 19: How does AI compare to human agents in your experience?	34
Figure 20: To what extent are you concerned about each of the following when dealing with AI in customer service?	35
Figure 21: Concerns about difficulty reaching a human agent when dealing with AI in customer service (by age range)	36
Figure 22: Quality of escalation from bot to human	37
Figure 23: Agent activity	39

Figure 24: Proportion of inbound call spent navigating between screens / applications.....	40
Figure 25: Requirements for in-call and post-call AI automation.....	40
Figure 26: "New agents find it difficult to familiarize themselves with our systems", by contact center size	41
Figure 27: Impact of real-time AI agent assistance on selected metrics.....	42
Figure 28: Agreement with statements about how agents feel about AI	43
Figure 29: Effect of AI on the number of people employed in the contact center	44
Figure 30: Use of interaction analytics, by contact center size	45
Figure 31: Use of various interaction analytics functionality (from only those respondents who use analytics)	46
Figure 32: Usefulness of post-call analytics	47
Figure 33: Usefulness of real-time analytics.....	48
Figure 34: Challenges to managing performance & quality	49
Figure 35: Effectiveness of quality assurance	50
Figure 36: How organizations quality assure interactions handled by AI	51
Figure 37: Extent to which AI-driven analytics is generating actionable operational insight ...	52
Figure 38: Caller identity authentication methods (only those contact centers which authenticate some or all calls)	53
Figure 39: Concerns about security	54
Figure 40: Current and planned use of voice biometrics or AI-based authentication and fraud detection.....	55
Figure 41: Current and planned use of proactive and predictive outbound AI	56
Figure 42: What is preventing the contact center from achieving its aims?.....	63
Figure 43: Where the majority of AI capability in the contact center comes from	64
Figure 44: What matters most when deciding where to source AI (top 3 answers)	65
Figure 45: Main reasons for stopping, pausing or rolling back an AI solution	66
Figure 46: Most significant barriers to further AI adoption over the next two to three years ...	67
Figure 47: Modelled saving per 1,000 inbound calls, by intervention (as examples only)	69

Figure 48: Differences in performance between AI-handled and human-handled interactions	70
Figure 49: Performance against the expectations set in the original business case	71
Figure 50: Which statement is closest to the expected position in five years	72



Observe.AI is the AI Agents company for customer experience, helping enterprises transform how they serve customers, support frontline teams, and run contact center operations.

The Observe.AI Agentic CX solution brings together specialized AI Agents that work across the customer experience lifecycle:

AI Agents for Customers autonomously handle customer conversations across voice and chat, understanding intent, taking action, completing workflows, and resolving requests end to end. When human support is needed, interactions can be seamlessly handed off with full context.

AI Agents for Frontline Teams, including Companion Agent, work alongside customer service representatives before, during, and after every interaction. They provide real-time guidance, surface relevant knowledge, automate workflows, support compliance, and reduce after-call work so teams can deliver faster, more consistent service.

AI Agents for Operations help contact center leaders evaluate interactions, uncover insights, improve quality, strengthen compliance, and coach teams at scale. By analyzing customer interactions across human and AI Agents, organizations gain a comprehensive view of performance and opportunities for improvement.

Underpinning these AI Agents are enterprise capabilities for building, testing, evaluating, and continuously improving AI performance, with the governance and integrations required to operate at scale.

Observe.AI works with **300+ global brands** across industries including financial services, insurance, healthcare, retail, consumer services, and technology. Customers include organizations such as **DoorDash, Affordable Care, and Simple Online Healthcare**, using Observe.AI to increase automation, improve customer experiences, strengthen frontline performance, and reduce the cost and complexity of contact center operations.

Observe.AI helps organizations move beyond AI that simply assists or answers questions toward AI Agents that can understand, reason, act, and continuously improve across the customer experience.

Contact:

w: www.observe.ai

INTRODUCTION AND METHODOLOGY

This report examines the use of artificial intelligence in the voice channel of US contact centers: what is deployed today, what it achieves, what customers make of it, and where the evidence supports the claims being made for it.

WHAT THIS REPORT COVERS

Voice AI is treated here as four layers rather than a single product, with the first section setting out that framework and defining the terms.

Other sections consider:

- audio quality, which determines what every other layer has to work with
- voice self-service and autonomous resolution
- AI running alongside the agent rather than instead of them
- the analytics and quality assurance built on the resulting transcripts
- identity, authentication and fraud
- proactive outbound and the US rules that govern it
- implementation and with measuring return.

Digital-only applications are outside the scope of this report except where they provide a useful comparison. Chatbots and web self-service are covered in the companion Inner Circle Guides.

SOURCES

Business-side findings are drawn from multiple recent surveys of over 500 US contact center operations. Consumer findings are drawn from surveying 1,000 US consumers.

Where a figure has been modeled rather than measured, the calculation and its assumptions are set out in full so that readers can substitute their own.

HOW TO READ THE SEGMENTATIONS

Contact center size bands are used throughout: small operations have fewer than 50 agent positions, medium 50 to 200, and large more than 200.

Vertical market breakdowns are shown where the variation is material and the base size supports it.

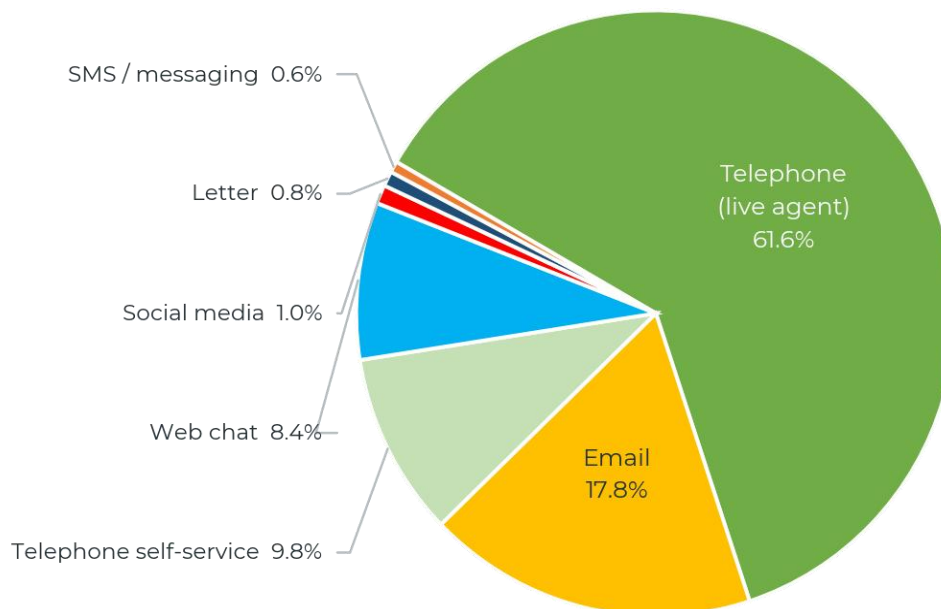
THE VOICE CHANNEL AND WHAT "VOICE AI" MEANS

Voice AI is among the most heavily marketed categories in contact center technology, and among the least precisely defined, which makes it worth establishing what is actually being sold before considering what it can do. This chapter sets out why the voice channel continues to attract the investment it does, what the term covers once it is broken into its component parts, and where the distinctions matter to anyone comparing one proposal against another.

VOICE IS NOT GOING AWAY

For twenty years the received wisdom has been that telephony would gradually give way to digital channels, and while the shift has certainly happened, it has not gone nearly as far as most predictions assumed.

Figure 1: Inbound interactions by channel



Live agent telephony accounts for 61.6% of inbound interactions, having declined very slightly once again. The long-term trend is gentle rather than steep, and it shows little sign of accelerating despite the industry continuing to expect that phone volumes will fall. Email holds its place as the second most important channel at 17.8%, web chat has grown this year from 7.9% to 8.4%, and SMS, messaging and letters remain very minor channels for most organizations.

More significant for this report is telephony self-service, which accounts for 9.8% of inbound activity and has risen again, from 7.4% in 2023. Taken together with live agent telephony, more than 71% of all inbound interactions in the US now arrive by phone, and close to one in seven of those is already handled without an agent. That is a materially different starting position from most digital-first commentary on this mark

et, and it is one reason voice deserves the attention this report gives it.

THE COST OF A VOICE MINUTE

Voice is the most expensive channel a contact center operates, though by a narrower margin than might be expected.

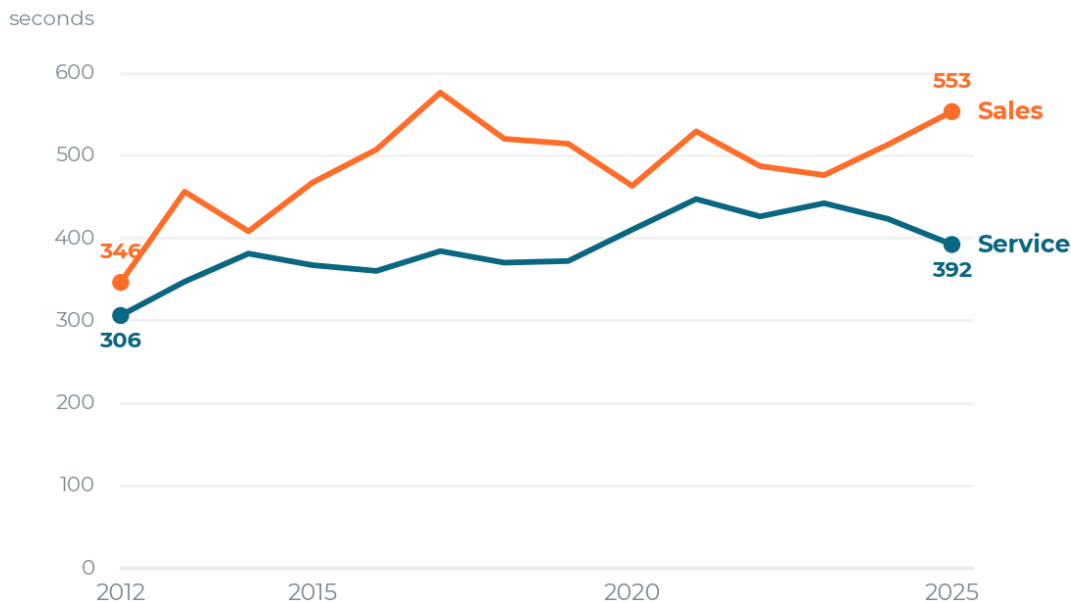
Figure 2: Cost per inbound interaction

Channel	Mean	1st quartile	Median	3rd quartile
Phone	\$7.20	\$8.00	\$7.35	\$4.00
Email	\$4.91	\$7.50	\$5.00	\$2.00
Web chat	\$5.84	\$7.75	\$5.00	\$3.20
Social media	\$5.50	\$8.00	\$5.00	\$3.00

The mean cost of handling an inbound call is \$7.20, against \$4.91 for email, \$5.84 for web chat and \$5.50 for social media.

Digital channels are somewhat cheaper than live voice, and this affects the business case for voice AI in a specific way: because a minute removed from a phone call is worth more than a minute removed anywhere else, identical automation applied to voice returns more than it does on chat or email.

The direction of travel makes the point sharper still.

Figure 3: Historical mean call duration (service & sales), 2012 - 2025

Mean service call duration has risen from 306 seconds in 2012 to 392 seconds in 2025, and sales calls from 346 seconds to 553 seconds over the same period. Service call duration has eased back from its 2023 peak of 442 seconds, which is the first sustained fall recorded in the series.

Part of this is a consequence of earlier automation succeeding rather than failing, since as short and simple enquiries moved to self-service and to digital channels, what remained on the phone became longer and more complex on average.

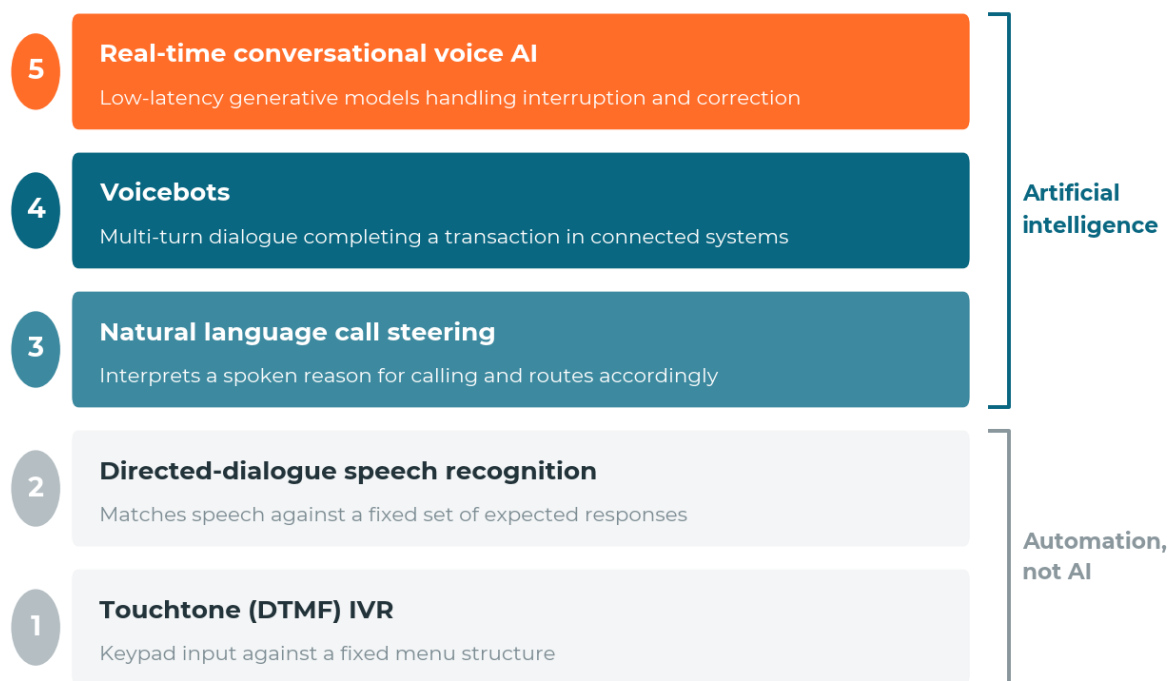
Queue times have moved in the same direction. Average speed to answer stood at 31 seconds in 2012 and 60 seconds in 2018 before reaching 101 seconds in 2021, and after easing back it stands at 74 seconds in 2025, with call abandonment at 7.3%. However, both remain well above where they were a decade ago.

Taken together, these figures describe a channel that is expensive, growing more expensive per contact, and carrying the great majority of inbound customer contact.

WHAT "VOICE AI" ACTUALLY COVERS

The term is applied to at least five distinct technologies which differ substantially in cost, capability and maturity.

Figure 4: The five levels of voice self-service technology



The distinction drawn on the right of the chart that matters commercially. Touchtone IVR is a keypad-driven decision tree that has been in service since the 1980s and contains no intelligence of any kind, while directed-dialogue speech recognition, although it accepts speech rather than keypresses, matches what it hears against a fixed grammar of expected responses and understands nothing that falls outside it. Neither can be described as artificial intelligence in any sense.

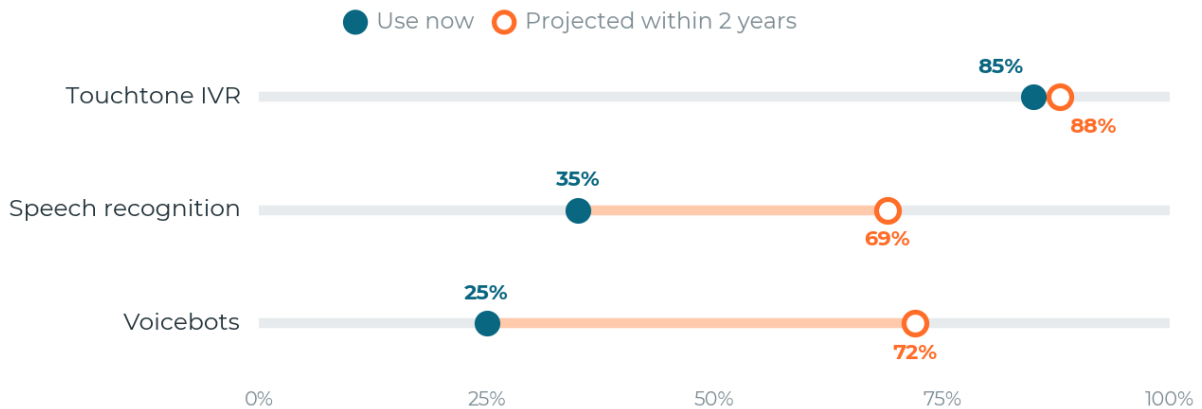
From the third level upwards the position changes, because a model is interpreting meaning, rather than matching a pattern.

Natural language call steering works out what a caller wants from an unconstrained sentence, voicebots carry that understanding through a multi-turn exchange and act on it in connected systems, and real-time conversational voice AI adds the ability to handle interruption, correction and the general untidiness of how people actually speak.

These three are genuinely AI, and they are also where the cost, the integration burden and the risk of the system doing something unintended all sit.

The practical consequence is that adoption figures for "voice AI" are close to meaningless unless the level is specified.

Figure 5: Current US adoption of voice self-service technologies



Touchtone IVR is used by 85% of US contact centers, rising to 100% of large operations, and a further 3% expect to implement it, which is to say that the overwhelming majority of what the industry currently operates under the heading of voice automation is a technology with no language understanding at all.

Automated speech recognition is used by 35% and by 55% of large operations, with a further 34% planning to implement, while voicebots are used by 25% against 72% projected within two years. There will be a considerable amount of cross-over in speech recognition and voicebots: few if any operations will be implementing keyword-based recognition.

It should be noted that these are not stages of a maturity path to be climbed in order. The lower levels are not obsolete and for a substantial proportion of contacts they remain the correct choice, since entering a sixteen-digit account number is faster and more accurate on a keypad than in conversation, and a yes-or-no confirmation gains nothing from generative reasoning. The relevant question when designing a journey is not how advanced a technology is, but whether it suits the contact it is being asked to handle.

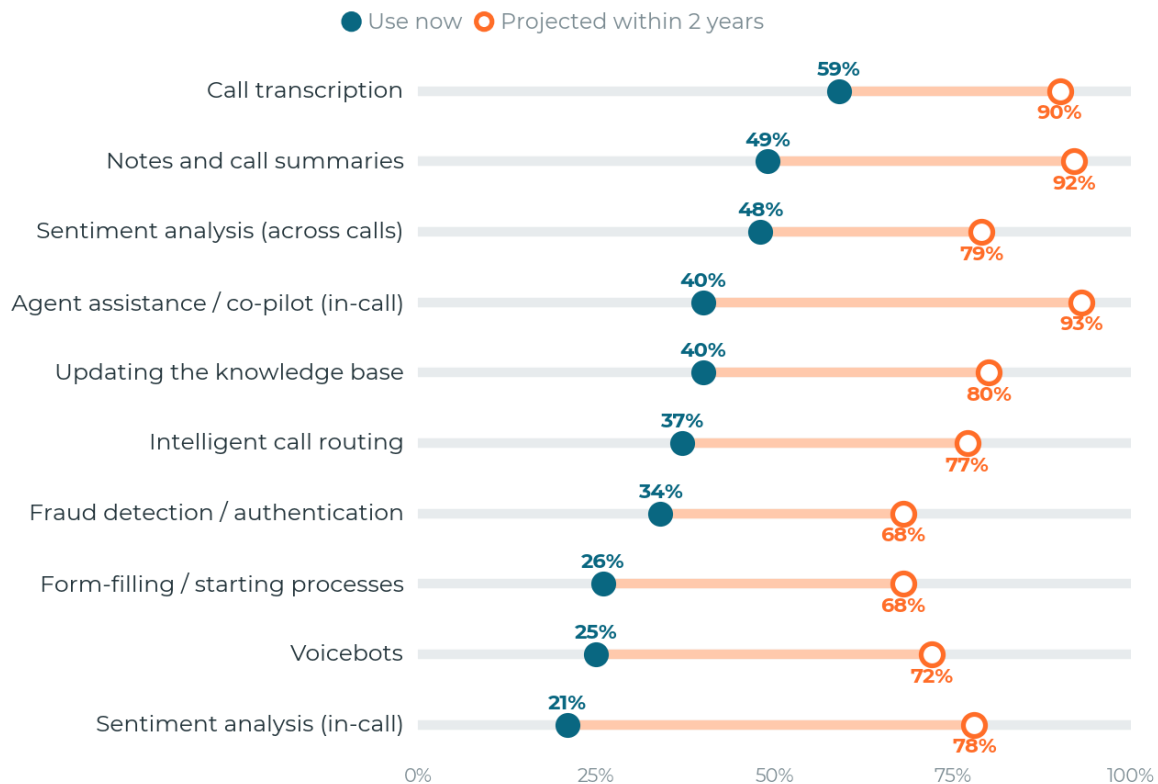
Everything discussed so far concerns voice AI used for self-service, which is to say systems that replace an agent for the whole interaction. While this is the most visible application and the one that attracts most of the marketing attention, it is a long way from being the only application of voice AI.

VOICE AI IS NOT ONLY SELF-SERVICE

A voicebot answering a call without a human being involved is only one of the things artificial intelligence now does to a voice interaction.

A far greater volume of voice AI runs alongside the agent rather than instead of them, or operates on the audio and the transcript after the conversation has finished, and in each case the customer may be entirely unaware that any AI was involved at all.

Figure 6: Current and planned use of voice-related AI applications



The pattern in the chart is consistent and worth dwelling on: the applications that assist rather than replace are already more widely deployed than voicebots.

Call transcription leads at 59%, followed by note and summary generation at 49%. Agent assistance has risen to 40%, level with knowledge base updating.

Voicebots are used by 25%, with 72% projected within two years.

These applications fall into five groups, each of which is treated at length later in this report.

- **Audio processing** improves what the system and the agent actually hear, through noise suppression, voice isolation, accent conversion and real-time translation. None of it is visible to the customer, but it determines how well everything downstream performs.
- **Call transcription** converts speech to text and is used by 59% with a further 31% planning to adopt. It is the foundation for analytics, summarization and quality assurance alike, which is why it is the most widely deployed voice AI application of any kind.
- **Agent assistance** supplies the agent with suggested responses, retrieved knowledge and next best action while the conversation is live, and is in place at 40% with 93% projected within two years. Sentiment analysis during the call, at 21%, does much the same job, and shows the largest gap between current and planned use of any application measured, at 57 points.
- **Summarization and wrap-up** generate call notes, disposition codes and CRM updates once the conversation ends, used by 49% and planned by a further 43%, taking projected adoption to 92%.
- **Authentication and fraud detection** use the voice itself as a credential and screen for synthetic or fraudulent callers, currently at 34% and projected to reach 68%.

Analytics belongs alongside these, though it operates on the aggregate rather than on the individual call. Sentiment analysis across multiple conversations is used by 48%, making the retrospective examination of voice one of the most established applications of AI in the contact center and one of the least discussed under the voice AI heading.

Intelligent call routing belongs in this list because it is the first point at which a model interprets what a caller has said. Rather than presenting a menu, it takes an unconstrained spoken sentence, works out the reason for the call and directs it accordingly, which makes it voice AI acting on speech even though no part of the enquiry is automated.

Framed this way, the market looks rather different from the way it is usually presented. Self-service automation is the most conspicuous form of voice AI and the one buyers are most often asked to consider first, yet it is currently one of the least adopted of all the applications measured.

THE FOUR LAYERS OF A VOICE AI DEPLOYMENT

Whichever of these applications is under consideration, and whatever the supplier chooses to call it, voice AI deployments can be assembled from the same four layers.

Figure 7: The four layers of a voice AI deployment



The weakest layer sets the effective performance of the whole.

These layers are examined in turn through the remainder of this report. They frequently originate from different suppliers, they are often procured at different times and by different people, and the weakest of the four determines the effective performance of everything above it. A voicebot with excellent intent detection will still fail if the audio reaching it is degraded, and an analytics platform can only be as accurate as the transcription it is reading.

Definitions matter commercially: loose terminology is not simply an irritation for the pedantic, but has a direct and measurable effect on buying decisions in at least three ways.

The first is that comparison becomes impossible. Two suppliers quoting a containment rate may be measuring entirely different things, since one may count any call that did not reach an agent, including those the customer abandoned in frustration, while the other counts only calls where the customer's stated goal was actually achieved. Without a common definition the two numbers cannot be placed side by side, yet they routinely are.

The second is that scope drifts upwards. A project justified on the basis of handling routine balance enquiries can end up specified as a fully conversational system carrying the corresponding cost, integration burden and risk, simply because the distinction between the levels was never made explicit at the point the business case was written.

The third is that failure gets misattributed. When a deployment underperforms, the conclusion drawn is usually that voice AI does not work when the fault may lie in a different layer altogether. Buyers would do well to establish which level and which layers a proposal actually addresses before comparing one price against another.

Affordable Care Drives 90%+ FCR Saves 2,000+ Hours/Month

Affordable Care is a U.S. dental support organization founded in 1975, focused on tooth replacement solutions. Based in Morrisville, North Carolina, it supports dentists through its Affordable Dentures & Implants brand. The network spans about 425 practices across roughly 40 states, many with on-site labs for same-day service now.

By implementing VoiceAI Agent, Affordable Care was able to reduce transfers, improve customer satisfaction, and elevate agent impact.

2026

Case Study.

95% containment rate**90% first-call resolution****2K agent-hours saved per month**

Challenges

Affordable Care was looking to upgrade its current interactive voice response (IVR) solution to a more advanced voice AI agent option but was struggling to find the right partner that offered flexible, reliable, and cost-effective deployment options that meet their unique needs.

Solution

Based on the long-term partnership with the Observe.AI team and the track record of success using the Observe.AI conversation intelligence platform, Affordable Care decided to become one of the first adopters of VoiceAI Agents—successfully deploying its first AI agent in 2 weeks.

Outcomes

- Beth, Affordable Care's AI Agent, handled FAQs on office locations, hours, and appointments, driving 95% containment.
- 90% of callers resolved inquiries on the first call without transfer, freeing agents to focus on higher-value conversations.
- Beth freed up 2,000 agent hours per month across the organization.

GETTING THE AUDIO RIGHT

The bottom layer of the stack set out in the previous chapter is the one that receives least attention, partly because it is invisible to the customer and partly because it has historically been treated as a hardware question rather than a software one.

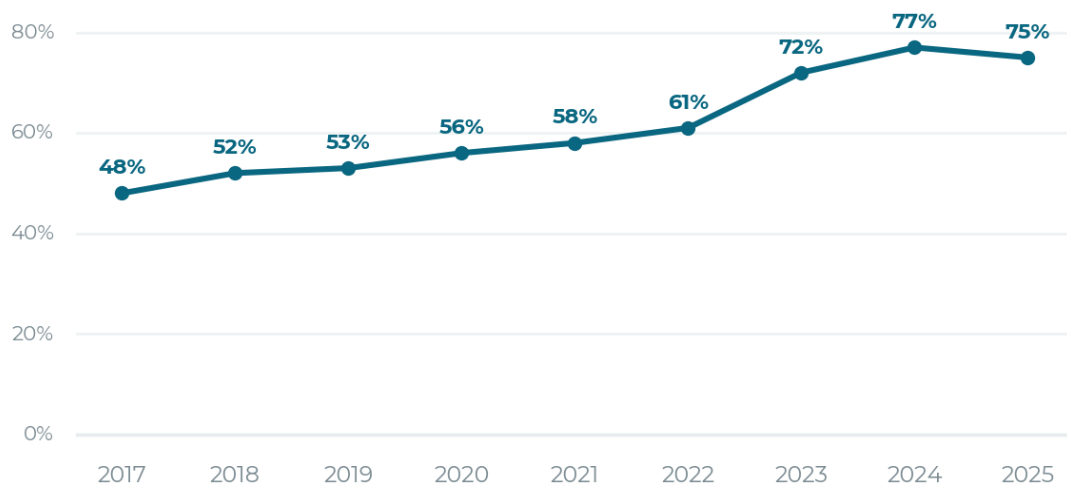
It determines what every other layer has to work with, and this section looks at why audio quality on a contact center call has been getting harder to control, what it costs when it goes wrong, and what can now be done about it.

TWO TRENDS WORKING AGAINST CLEAR AUDIO

Two structural changes over the past decade have made the typical call noisier at both ends, and neither shows any sign of reversing.

The first is the move to cell phones. On average, 75% of customers now use a cell phone rather than a landline to call the contact center, offering considerable potential for value-add services such as video and visual IVR, but also placing an uncontrolled acoustic environment at the customer end of most calls.

Figure 8: What % of your customers use a cell phone to call a contact center? 2017-25

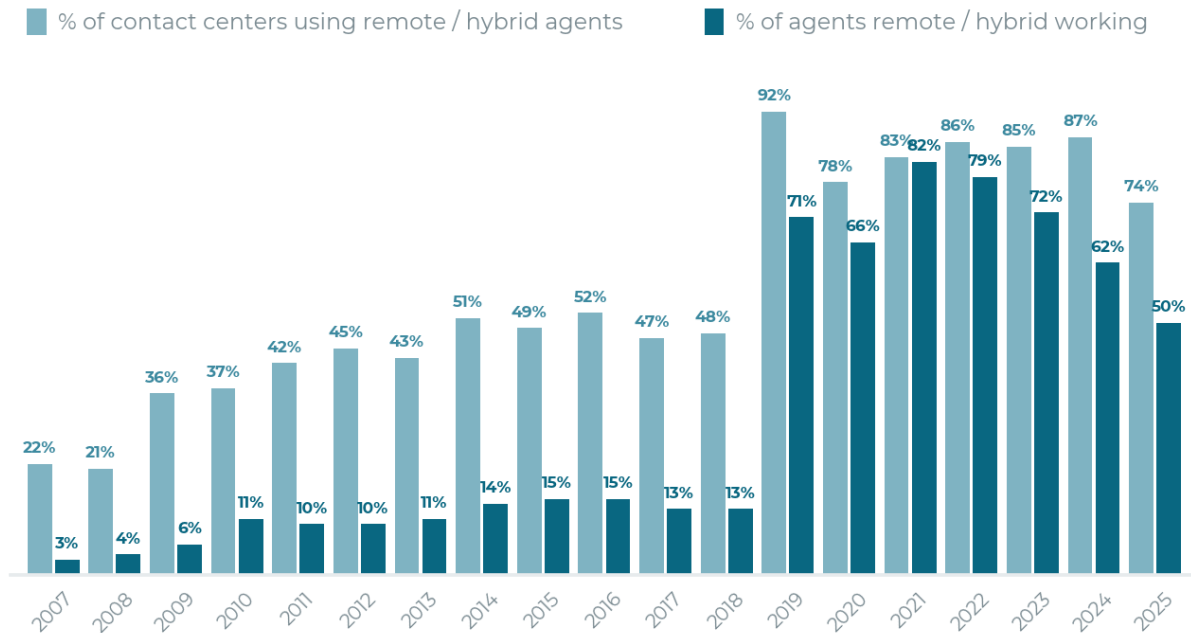


The shift has been steady rather than sudden. The figure stood at 48% in 2017 and rose gradually to 61% by 2022 before jumping to 72% in 2023 and 77% in 2024, easing back to 75% in 2025.

The practical consequence is that the typical call now carries considerably more background noise from the customer's end than it did, and the trend is still running in one direction.

The second change is on the contact center's own side of the line.

Figure 9: The rise of remote & hybrid contact center working (2007-2025)



Before 2019, remote working was a minority practice: no more than 15% of US agents worked remotely in any year between 2007 and 2018, even though around half of operations had some remote capability by the end of that period.

The pandemic moved the industry to over 80% of agents working remotely or in a hybrid manner in 2021, and while the figure has drifted down from its peak, it still stood at 50% in 2025 across 74% of contact centers. Remote and hybrid working is a smaller share of the US industry than of the UK, but it remains the arrangement for half of all agents.

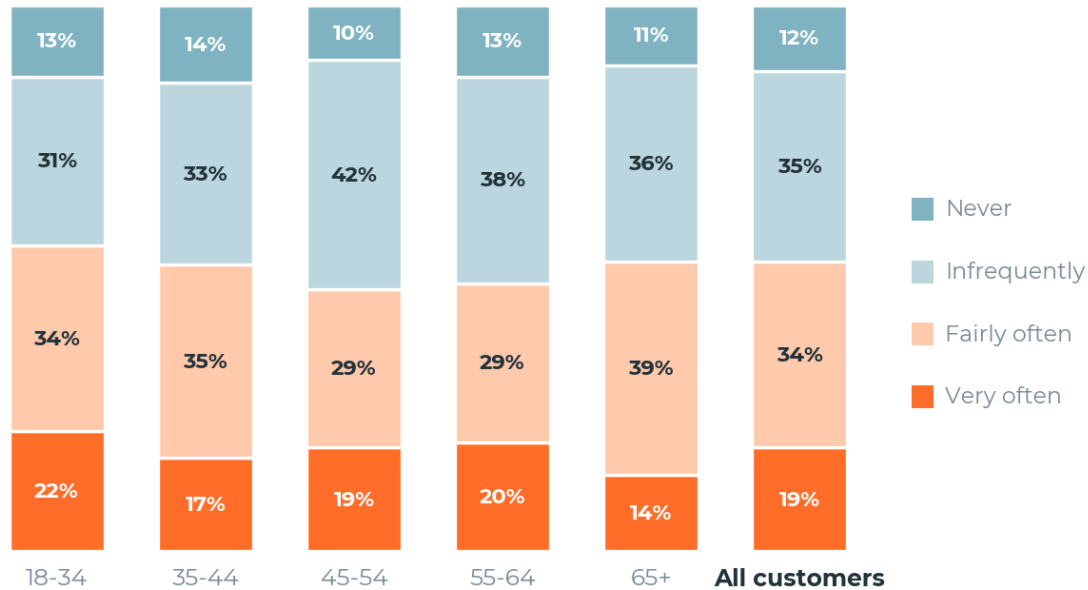
There is no expectation within the industry that a substantial return to a fully centralized model will happen.

This matters for audio because a remote or hybrid operation has to manage an agent working environment it does not own. Traffic, children, dogs and everyday domestic sound are less easily managed and less predictable than an office, although it should be noted that many large open-plan contact centers carry even higher levels of ambient noise than a spare bedroom does.

WHAT POOR AUDIO COSTS

Until recently there has been very little research quantifying the effect that poor audio quality has on the efficiency of the contact center or on the customer's experience of it, largely because time lost to mishearing is rarely tracked and therefore rarely noticed.

Figure 10: Customer issue: "I can't hear clearly and/or I have to repeat myself as they can't hear me" (by age range)



In ContactBabel's survey of US consumers, 19% reported that they very often had problems hearing the agent or were asked to repeat themselves, and a further 34% said it happened fairly often. Just over half of customers therefore experience the problem regularly, and only 12% say it never happens at all.

The age breakdown is the interesting part, because it does not follow the pattern most people would predict. It might have been expected that older customers would report the problem most, yet the reverse is true.

The 65+ group records the lowest "very often" figure of any age band at 14%, while the youngest cohort records the highest at 22%, with 56% of 18-34s saying it happens either very often or fairly often. Whatever is happening here, it is not simply a hearing-loss effect, and the likeliest explanation lies in how and where younger customers make their calls rather than in how well they hear.

Hearing loss remains a real factor in the older cohorts, and its prevalence will rise as the population ages. It cannot, though, be the explanation for a pattern in which the youngest customers report the problem most severely and the oldest the least.

Business cases need quantities rather than impressions, so ContactBabel has modeled the industry-wide cost of the time lost.

Figure 11: The unnecessary cost of mishearing



The calculation runs as follows. US contact centers handle an estimated 27.3bn agent-handled inbound calls a year. If 30% of those calls require a sentence to be repeated, that is 8.2bn calls. Repetition accounts for around 3.5% of the call; and at a mean cost per inbound call of \$7.20, that is 25.2 cents per affected call and \$2.1bn across the industry each year.

Applying the same assumptions to a 250-seat operation handling 12,688 calls per agent position per year gives an annual cost of \$194,418.

Some operations will sit well above these levels. Those with a high proportion of remote agents and those whose customers are predominantly on cell phones will experience more repetition than the average, and the model can be re-run with local assumptions.

WHAT CAN BE DONE ABOUT IT

The traditional answer has been hardware: noise-cancelling microphones filter out a proportion of unwanted background noise, and voice tubes allow more flexible microphone positioning with some improvement in sound quality, but even the most expensive and sophisticated headsets can only go so far in cleaning up background noise, and they still need to be replaced and repaired over time.

There is evidence that the hardware approach does produce results. A Spanish contact center issued one group of employees with headsets containing digital audio processors while a control group used traditional headsets, and the group with the processed audio handled calls more quickly, made fewer mistakes when collecting data, generated fewer repeat calls, and handled around 10% more calls per day than the control group.

The software approach works differently. AI-enabled voice isolation removes background noise from both sides of the conversation rather than only the agent's, and it operates both in real time, to keep the conversation flowing accurately, and on the recording afterwards, where it improves the quality of post-call analytics and speech-to-text transcription. As it is applied through software rather than in the headset, it also reduces the need to keep upgrading and replacing top-of-the-line hardware.

Voicebots depend on speech-to-text before anything else can happen, and carrying out that conversion is made materially more difficult by background noise.

This is most acute when the system is capturing names, addresses, payment card details and account numbers, which are precisely the elements a customer is most likely to be asked to repeat and precisely the elements where an error is most expensive.

Analytics, automated quality assurance and summarization all read the same transcript, so the accuracy of that first conversion impacts on everything after it.

ACCENT, LANGUAGE AND COMPREHENSION

Background noise is not the only thing that degrades understanding on a voice call. Two further issues affect both human agents and automated systems, and both have attracted specific AI products in the past two years.

The first is accent: while accent recognition by automated systems has improved greatly, it is still not infallible, and similar-sounding or identical-sounding words remain a recognized failure mode, with the examples given being “one” against “won”, “four” against “for”, and “two” against “to”, and these are exactly the sort of tokens that appear in account numbers, dates and amounts.

Accent conversion technology, sometimes marketed as accent localization or accent softening, is now sold in both directions: applied to the agent's voice so the customer understands more easily, and applied to the customer's voice so the agent does.

The second is language. Real-time translation converts speech between languages during the call, with the stated aim of allowing an operation to serve customers in languages its agents do not speak.

25% of US contact centers use real-time translation or multilingual support, and a further 27% plan to implement it. 45% have no plans at all. Adoption is markedly higher than in the UK, where the equivalent current-use figure is 18%, which is what a larger Spanish-speaking customer base would lead one to expect.

Figure 12: Current and planned use of real-time translation or multilingual support

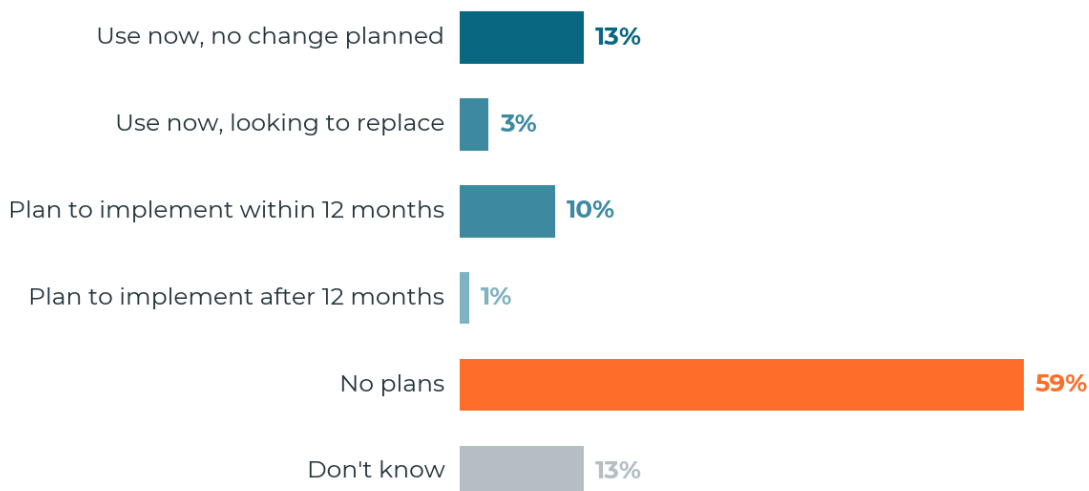
Response	% of respondents
Use now, no change planned	21%
Use now, looking to replace	4%
Plan to implement within 12 months	18%
Plan to implement after 12 months	10%
No plans	45%
Don't know	3%

Where translation is deployed it tends to be settled rather than experimental, with 21% using it now with no change planned against only 4% looking to replace what they have. That is the profile of a capability that works adequately for the organizations that need it, rather than one being actively iterated.

It should be noted that the question covers real-time translation and multilingual support together, so the figure includes organizations using AI to route a caller to a language-capable agent as well as those translating speech during the call. Those are different capabilities with very different costs, and the survey does not separate them.

Accent conversion is among the least adopted voice AI applications measured anywhere in this report. 16% use it now, 10% plan to implement within twelve months and 1% after that, while 59% have no plans at all and 13% do not know.

Figure 13: Current and planned use of accent localization or accent softening



Accent localization and real-time translation are both real-time speech processing applied at the same point in the call, although they solve different problems. Translation addresses a problem an organization can point to on a spreadsheet, namely customers it cannot currently serve. Accent conversion addresses comprehension within a language both parties already speak, which is harder to size. Even in the US, 59% have no plans at all, against 45% for translation.

However, it is worth setting that 59% against what accent is understood to cost operations that already take it seriously. Offshore and nearshore outsourcers treat accent as a recruitment criterion rather than a training afterthought, rejecting a substantial proportion of otherwise capable applicants at interview on that basis alone, which narrows an already shallow hiring pool. Those that then provide accent training carry an onboarding cost and a recurring monthly cost per agent thereafter, plus the cost of the trainer, and agents are taken off calls to attend.

There is also a gap between what businesses believe is happening on their calls and what customers report. Operations consistently estimate that their agents are asked to repeat themselves because of accent far less often than customers say they ask, and customers themselves report being uncomfortable about asking at all, which means some calls end with the customer holding an answer they did not fully understand.

The consequences run beyond handling time into confidence in the agent, willingness to buy, and how the customer feels about the organization afterwards.

Accent conversion is therefore competing with an established alternative rather than filling a vacuum, which may explain the 59% with no plans better than indifference does. The comparison a buyer is actually making is against recruitment screening and a training program they already run, and against the agent morale cost of asking people to change how they speak.

We suspect that in many cases, contact centers simply underestimate the issue from both the agents' and customers' perspectives, that this accounts for many of the 59%, and as the solution's full benefit is better understood it will become more popular.

VOICE SELF-SERVICE: FROM CONTAINMENT TO RESOLUTION

Voicebots occupy an odd position in the market. They attract more attention than any other application of voice AI, they are the thing most buyers mean when they use the term, and they are amongst the least deployed of everything measured in the previous sections.

This section sets out what telephony self-service currently achieves in US contact centers, why sessions are abandoned, and what customers themselves say about being handled by a machine. The answers place a firm ceiling on what any containment target can reasonably assume.

HOW MUCH SELF-SERVICE ACTUALLY HANDLES

Among contact centers that offer telephony self-service at all, a mean of 33% of calls are handled entirely without an agent. The distribution matters more than the average.

Figure 14: Overall proportion of calls handled entirely through self-service (only in respondents which offer telephony self-service)

	Proportion of calls handled entirely through self-service, if offered
1st quartile	60%
Median	15%
3rd quartile	8%
Mean	33%

The median of 15% sits far below the mean of 33%, which tells us that a minority of operations achieve a great deal through self-service and pull the average upwards, while the typical operation achieves considerably less.

A quarter of respondents handle 60% or more of calls this way, and a quarter handle 8% or less. That top quartile is the widest spread recorded in either market.

There are sound reasons why the figure is not higher. Many calls are simply unsuitable, because they contain multiple requests within the same conversation, because they are complex, or because the caller has decided they want to speak to a person. Some smaller operations have call volumes too low to justify implementing self-service at all.

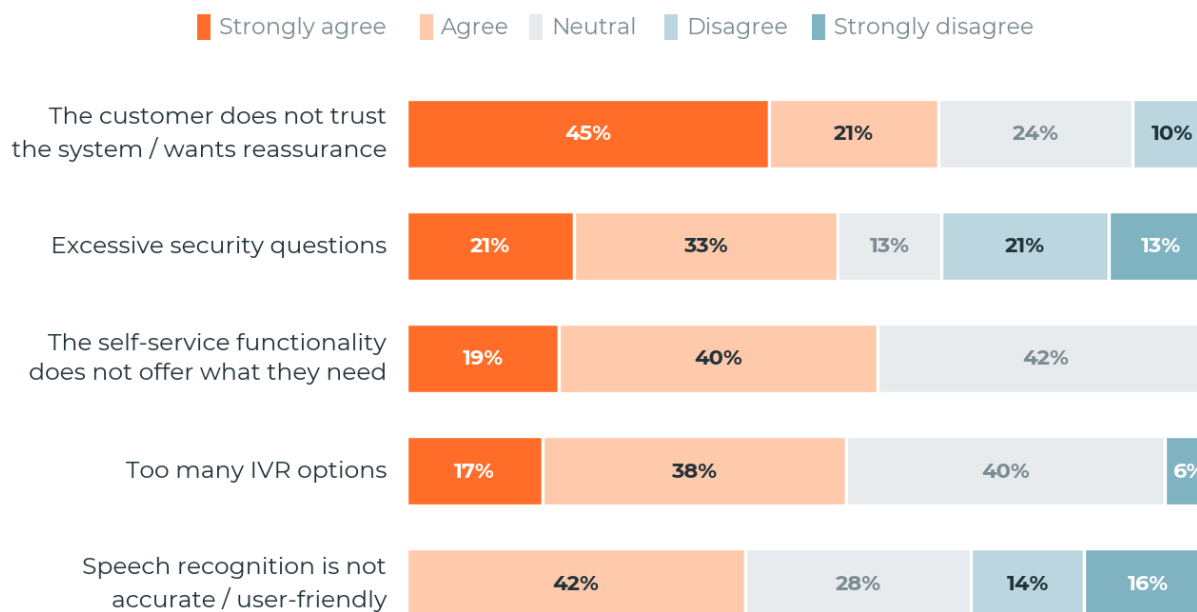
Customers who decline to use it have a straightforward remedy: a mean of 29% of calls that enter the self-service option are zeroed out, meaning the customer presses zero or finds some other shortcut to reach an operator. It varies considerably by operation size, with large contact centers seeing a mean of 34% zeroing out, against 14% in small operations.

The most plausible explanation is that larger operations attempt more. Respondents from larger organizations consistently report more options in their auto-attendant menus, and that tendency to offer a great deal of functionality is likely to extend to the self-service application as well. Overly complex or long-winded structures encourage abandonment, and that is a reasonable reading of what the size correlation shows.

WHY SESSIONS ARE ABANDONED

Contact centers were asked to what extent they agreed with five explanations for customers abandoning telephony self-service.

Figure 15: Reasons for abandoning telephony self-service sessions



66% agree to some extent that customers abandon because they do not trust the system and want human reassurance, with 45% agreeing strongly. 59% agree that the self-service function does not offer what the customer wants. Trust is by some distance the most strongly agreed-with explanation.

Among those using automated speech recognition, 42% agree that it is unpopular with customers because it lacks accuracy and user-friendliness. Given that speech recognition is the input layer for anything more sophisticated, that is a material figure, and it connects directly to the audio quality material in the previous section.

Only 17% strongly agree that too many options presented to customers is a major reason for seeking human assistance, and excessive security questions attracts the highest disagreement of all, with 34% disagreeing or strongly disagreeing.

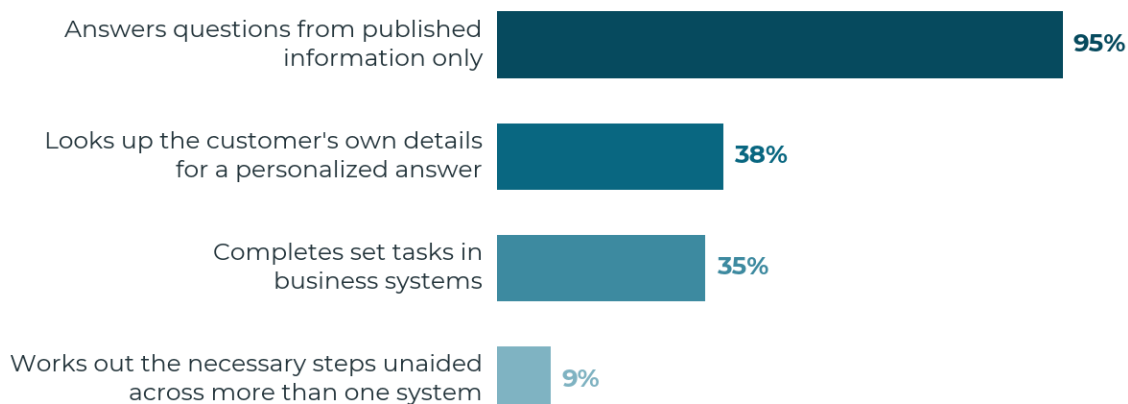
A customer abandoning because the system genuinely cannot do what they need is exercising a reasonable choice, and provided a clear exit to a live agent is marked early in the process, this does not have to be negative. Operations hiding their agents and sending customers around IVR loops does reputational damage to self-service.

WHAT VOICEBOTS CAN CURRENTLY DO

The figures above describe telephony self-service as a whole, most of which is touchtone IVR and directed-dialogue speech recognition rather than anything conversational.

What a modern voicebot can accomplish unaided, and what proportion of its interactions it resolves without a human, are separate questions.

Figure 16: Capability of the most capable chatbot or voicebot in the organization



95% of bot users – taking chatbots and voicebots together – report that their most capable bot answers questions from published information. 38% can look up the customer’s own details to give a personalized answer, 35% can complete set tasks in business systems, and 9% can work out the necessary steps unaided across more than one system.

Answering from published content is close to universal, acting in business systems is a minority capability, and the autonomous multi-system reasoning that dominates the marketing of this category exists in one operation in twenty.

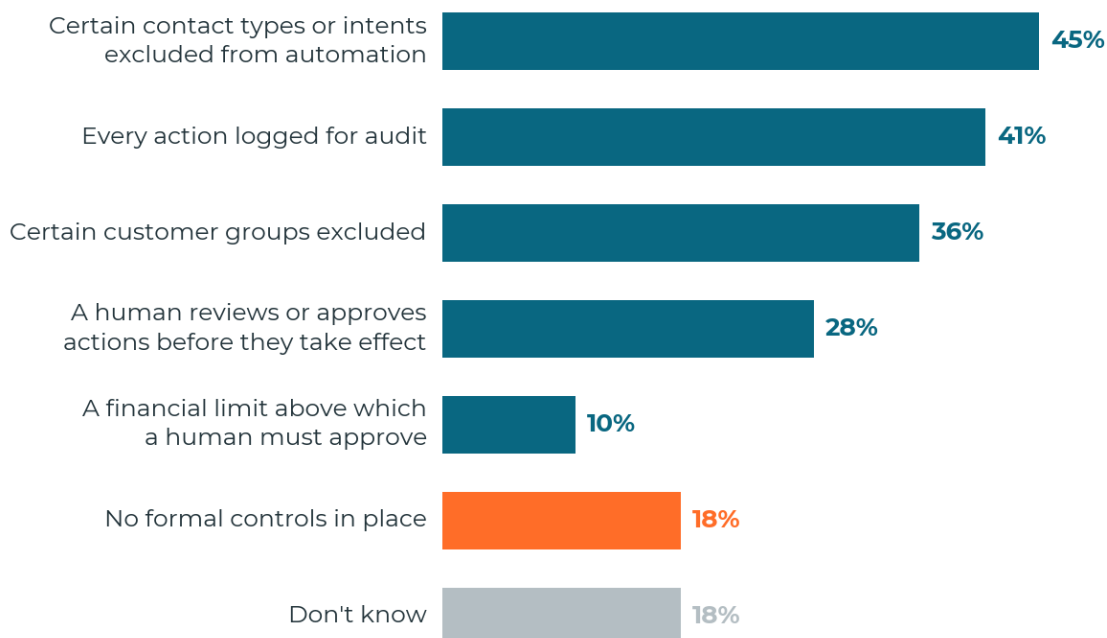
Where bots can act rather than merely answer, the question of what constrains them becomes material.

The controls shown in the following chart apply to bots of both kinds, chat and voice, since many organizations running a voicebot are running a chatbot alongside it and govern the two together.

45% exclude certain contact types or intents from automation, which is the most common control and also the crudest: it works by keeping the bot away from anything difficult rather than by governing what it does.

41% log every action for audit, 36% exclude certain customer groups, and 28% require a human to review or approve actions before they take effect.

Figure 17: Controls applied to the actions bots can take in business systems



The figure that should give any board pause is at the other end. 18% report no formal controls in place at all, and a further 18% do not know what controls exist.

Taken together, more than a third of organizations either have effectively little or no formal governance over what their bots do in live business systems or cannot say whether they have. Only 10% operate a financial limit above which a human must approve, which is the single most standard control in any other form of automated transaction processing.

Set against the capability data, the position is less alarming than it first appears, because a bot that only answers published questions has little to govern. It becomes considerably more alarming for the 35% whose bots complete tasks in business systems, although the data show that these are far more likely to have multiple controls attached to them.

HOW MUCH GETS RESOLVED WITHOUT A HUMAN

Among organizations running them, voicebots handle 22% of enquiries entirely without human involvement. The equivalent figure for chatbots is 31%.

The received view is that voice is intrinsically harder to contain than text, and on the evidence of these two numbers, the gap is not as large as might be expected, at least among organizations that have got as far as deploying a voicebot at all.

Two data points to note: firstly, selection, in that only 25% of respondents currently run a voicebot, and those that do are self-selected as the organizations with the appetite, the budget and the use cases to make one work.

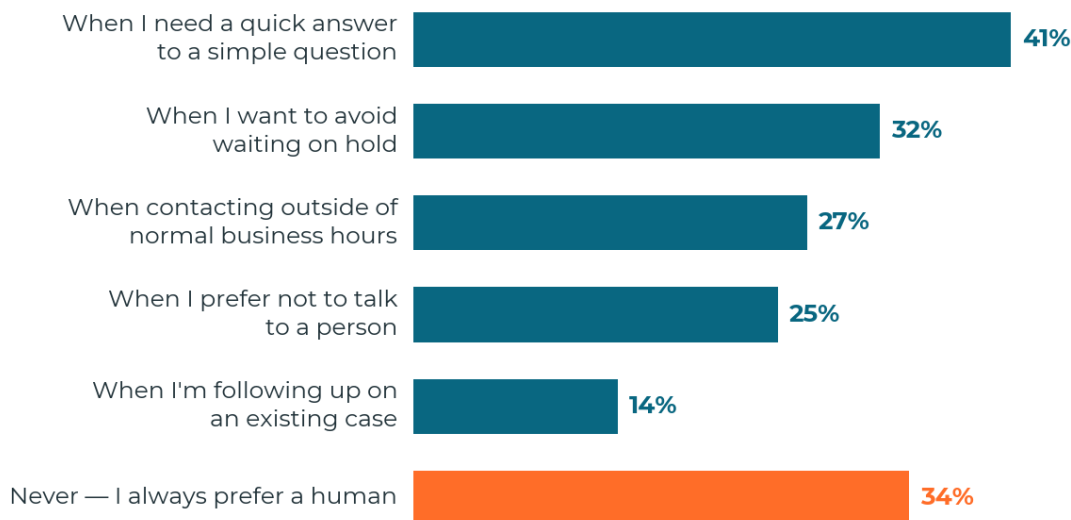
The second is that while this 22% figure sits above the 15% median for telephony self-service as a whole, it is still lower than the 33% mean across the industry, which is driven heavily by very simple touchtone IVR. However, as voicebots are being used for more complex tasks, this is not necessarily a concern.

WHEN CUSTOMERS PREFER TO DEAL WITH AI

The business view is only half of the picture, and the consumer research produces some encouraging findings.

66% of US consumers state that there are circumstances in which they would **prefer** to use AI rather than speak to a live agent. That is worth reading carefully, because it is not customers saying they would tolerate AI or would not object to it: they would actively prefer it.

Figure 18: When contacting an organization, in which situations would you prefer to deal with AI rather than a human?



41% would prefer AI when they need a quick answer to a simple question, which is by some distance the most popular situation.

32% would prefer it to avoid waiting in a queue, which is unsurprising given that queue length and first-contact resolution are the main drivers of customer experience, and the absence of a queue is one of the genuine structural advantages of voice self-service. Contacting outside business hours accounts for 27%.

25% say that there are circumstances when they would prefer not to talk to someone, perhaps about a sensitive issue, or simply that they do not like using the telephone.

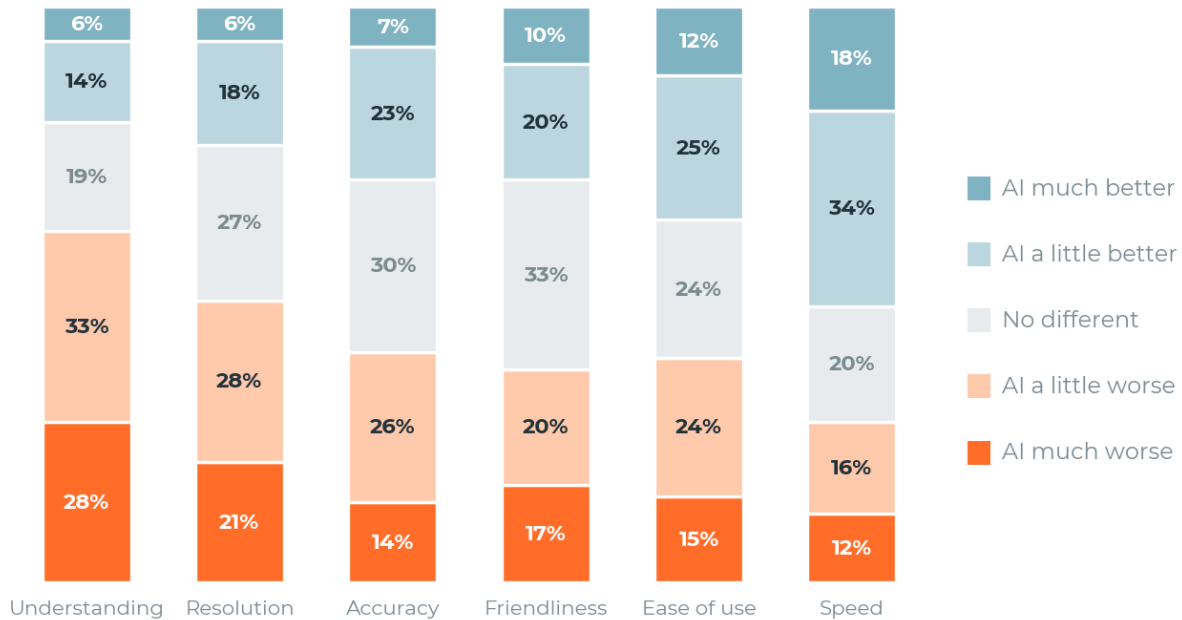
Against that, 34% of customers currently say they would never prefer AI and always want a human, with the figure skewed towards older customers.

That figure is a practical constraint on any containment ambition, and it does not appear to be movable through better design alone.

HOW CUSTOMERS RATE AI AGAINST HUMAN AGENTS

Consumers who have dealt with AI in customer service were asked to compare it directly with human agents across six dimensions.

Figure 19: How does AI compare to human agents in your experience?



The pattern points in one direction on most of the six measures, though less emphatically than in the UK.

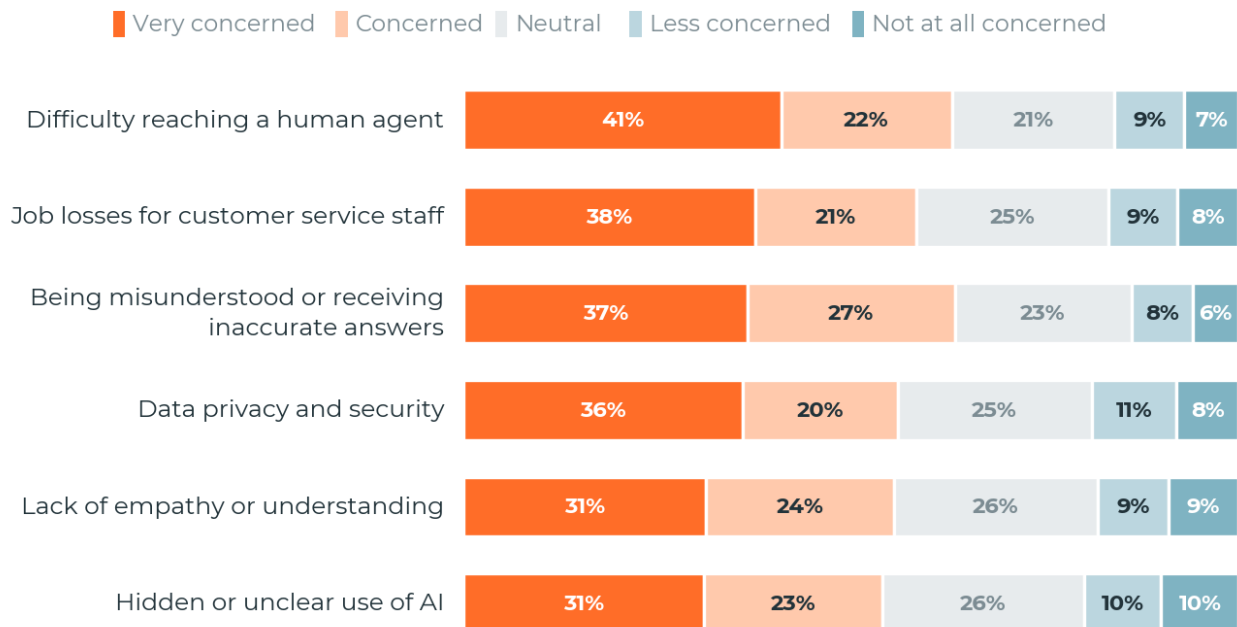
61% rate AI worse than a human for understanding the issue, against 20% who rate it better. 49% rate it worse for overall resolution, 40% worse for accuracy of information and 37% worse for friendliness of tone. On ease of use the market is close to evenly divided, at 39% worse against 37% better.

Speed of response is the single exception and the only measure where AI comes out clearly ahead, with 52% rating it better against 28% worse. That is the basis of the current consumer case for voice AI, and it explains why the situations in which customers say they would prefer AI are precisely those where speed is the thing they want: a quick answer, no queue, and service outside business hours.

THE CONCERN THAT SHAPES EVERYTHING ELSE

Consumers were also asked how concerned they are about various aspects of AI in customer service.

Figure 20: To what extent are you concerned about each of the following when dealing with AI in customer service?

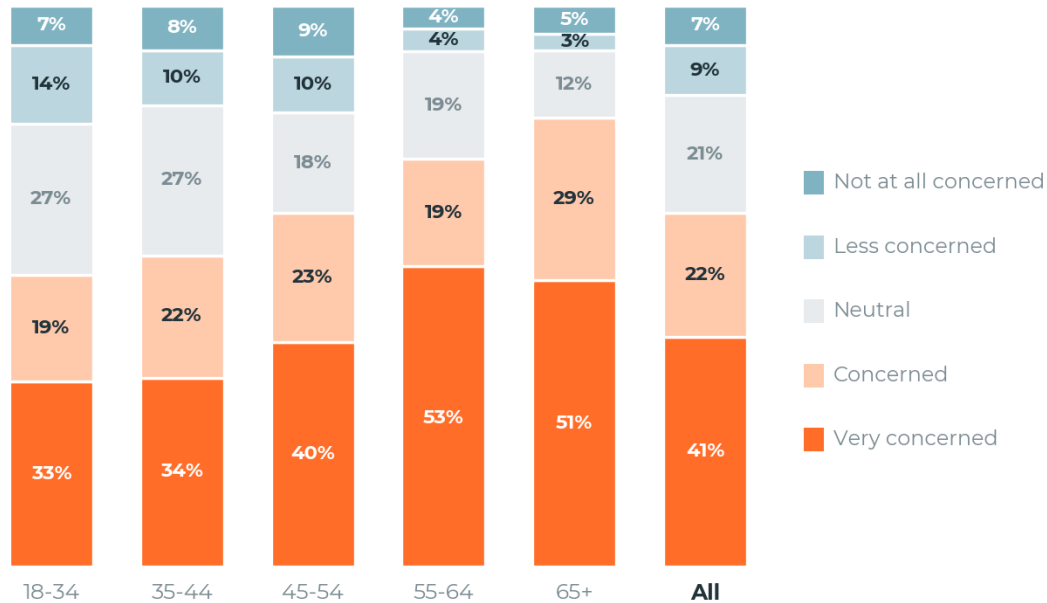


Difficulty reaching a human agent is the leading concern at 63% when very concerned and concerned are combined, followed by being misunderstood or given inaccurate answers at 64%, job losses at 59%, data privacy and security at 56%, lack of empathy at 55% and hidden or unclear use of AI at 54%.

The spread between first and last is narrow, which suggests a general unease rather than one dominant objection.

The first of those is worth examining by age, because the distribution is unusually wide.

Figure 21: Concerns about difficulty reaching a human agent when dealing with AI in customer service (by age range)



51% of customers aged 65 and over are very concerned about being unable to reach a human, and 80% are concerned to some extent. Among 18 to 34 year-olds the very concerned figure is 33%. Concern rises steadily through every age band in between.

This cohort of older customers has watched self-service tools, and touchtone IVR in particular, appear to be implemented in order to keep them away from live agents. Whether that was ever the intention is beside the point, because many customers experience self-service as denying them the right to speak to someone, and they respond accordingly.

The practical implication runs through everything in this chapter. Forcing customers to remain on a self-service channel is self-defeating: in the case of IVR they will zero out or select irrelevant options simply to reach a person, and any savings have to be weighed against the effect on long-term loyalty and revenue. Any self-service implementation, AI-enabled or not, needs a clear and easy route to a live agent.

WHEN SELF-SERVICE DOES NOT SUCCEED

Voice AI has one significant advantage over digital channels at the point where self-service fails, which is that the transition to an agent can be made without the customer starting again.

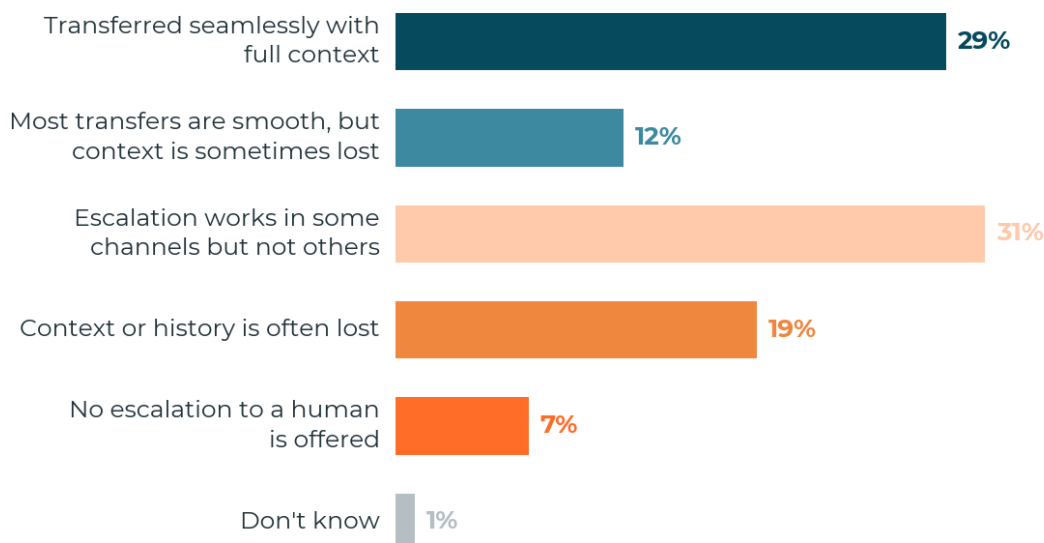
It can identify and authenticate the customer before passing the call on, which saves an average of 42 seconds per authenticated call (across the US contact center industry, agent-handled authentication accounts for over \$12bn a year).

Additionally, a summary of what the customer has been attempting can be passed to the agent, so the customer does not have to repeat themselves.

Together these make a second contact feel like a first-time resolution, which matters given that first-contact resolution is consistently among the top drivers of customer experience.

How well this works in practice is a separate question from whether the capability exists.

Figure 22: Quality of escalation from bot to human



29% report that customers are transferred seamlessly with full context, and 12% that most transfers are smooth but context is sometimes lost. 31% say escalation works in some channels but not others, which is a description of a partial integration rather than a design.

The tail is considerable, with 19% saying context or history is often lost, and 7% offer no escalation to a human at all. That last figure should be read against the consumer data in Figures 20 and 21, where difficulty reaching a human agent is the leading concern and 80% of customers aged 65 and over are worried about it. Only 29% of survey respondents can say the customer never has to repeat themselves, and repetition at the point of escalation is exactly the moment at which a customer concludes that the automation wasted their time.

WHAT THIS MEANS FOR CONTAINMENT TARGETS

The figures in this section support voice self-service, but as they stand, also constrain what can sensibly be promised for it.

Looking at current levels of voice self-service, a mean of 33% of calls handled end to end where telephony self-service is offered looks good at first glance, but a median of only 15% is not particularly impressive. Voicebots are currently reported to deliver 22% containment, and while this is still not spectacular, it will certainly improve as functionality and customer confidence and familiarity increase.

Note that 34% of consumers say they would never prefer AI under any circumstances, and this predicates a realistic containment ceiling that is a good deal lower than some business cases assume: 30-40% would seem a reasonable target to aim for initially.

The consumer research also indicates where the remaining headroom lies. Customers will choose automation for speed, and they judge it worse than a human on almost everything else.

A voicebot deployed against simple, quick, high-volume enquiries is working with the grain of what customers actually want. One deployed against complex or emotionally loaded contacts is working against 61% of consumers who consider AI worse at understanding the issue, and against the 63% who are worried about being unable to reach a person.

A containment target set without reference to these numbers is not a target but a wish.

VOICE AI ALONGSIDE THE AGENT

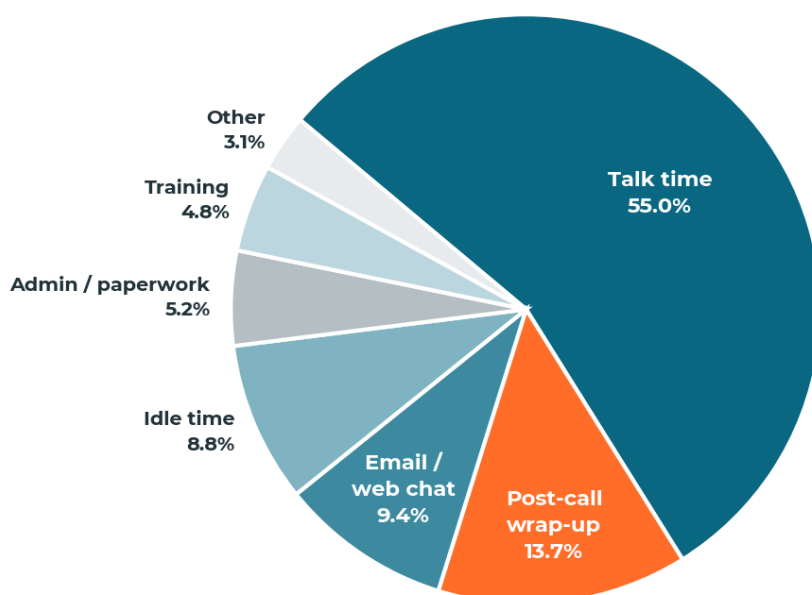
The voice AI applications in this chapter share a characteristic that separates them from everything in the previous one: the customer is still speaking to a person.

AI is listening, retrieving, drafting and summarizing in the background, and the value is measured in what the agent no longer has to do rather than in calls avoided. As shown earlier, these applications are already more widely deployed than voicebots, in several cases by a factor of three.

WHAT AGENTS ACTUALLY SPEND THEIR TIME ON

The case for assisting the agent rests on how a contact center day is currently spent.

Figure 23: Agent activity

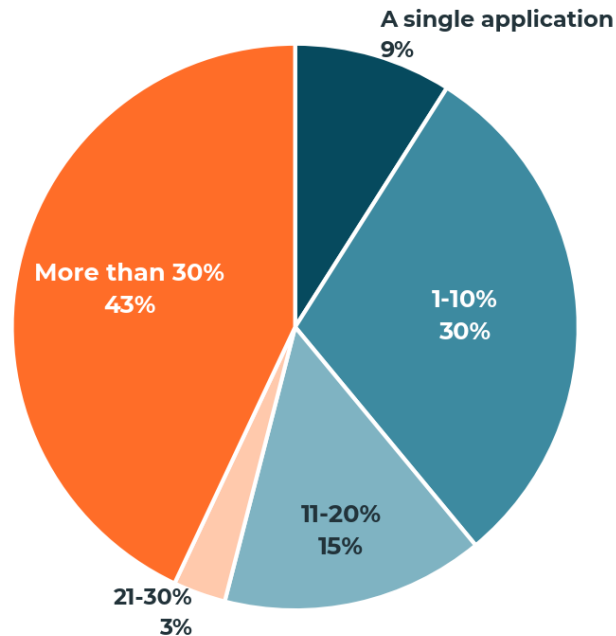


Talk time accounts for 55% of agent activity, with a further 9.4% on email and web chat. The largest remaining block is post-call wrap-up at 13.7%, which is time spent on work the customer never sees: writing notes, coding dispositions and updating systems after the conversation has ended. Idle time takes 8.8%, administration 5.2% and training 4.8%.

Post-call wrap-up is the single most addressable item on that list, because it is repetitive, it follows a predictable structure, and the raw material for automating it has already been captured. At an average length of 101 seconds, the average cost of post-call work is \$1.27 per call.

Time inside the call is less visible but no less significant.

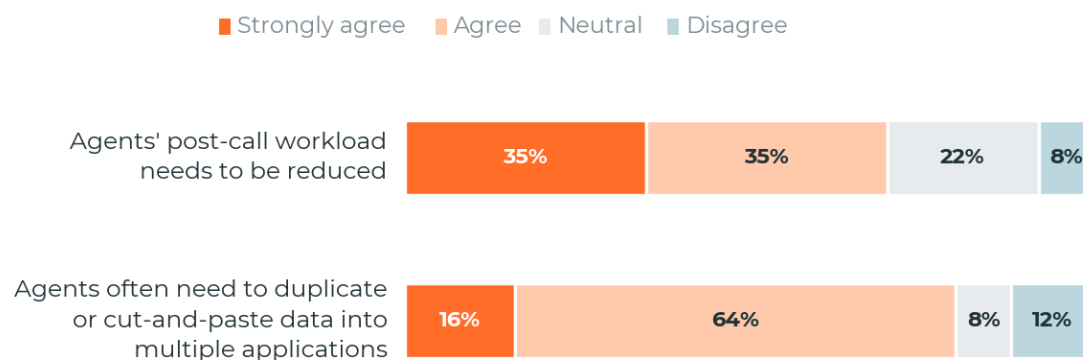
Figure 24: Proportion of inbound call spent navigating between screens / applications



Only 9% of contact centers report that agents handle inbound calls within a single application. At the other extreme, 43% report that more than 30% of the call is spent navigating between screens. For a substantial share of US operations, something approaching a third of every call is spent on system navigation rather than on the customer.

Respondents recognize the problem and its consequences. 80% agree or strongly agree that agents often need to duplicate or cut and paste data into multiple applications, a pattern that produces transcription errors as well as wasted time. 70% agree that agents' post-call workload needs to be reduced, with only 8% disagreeing

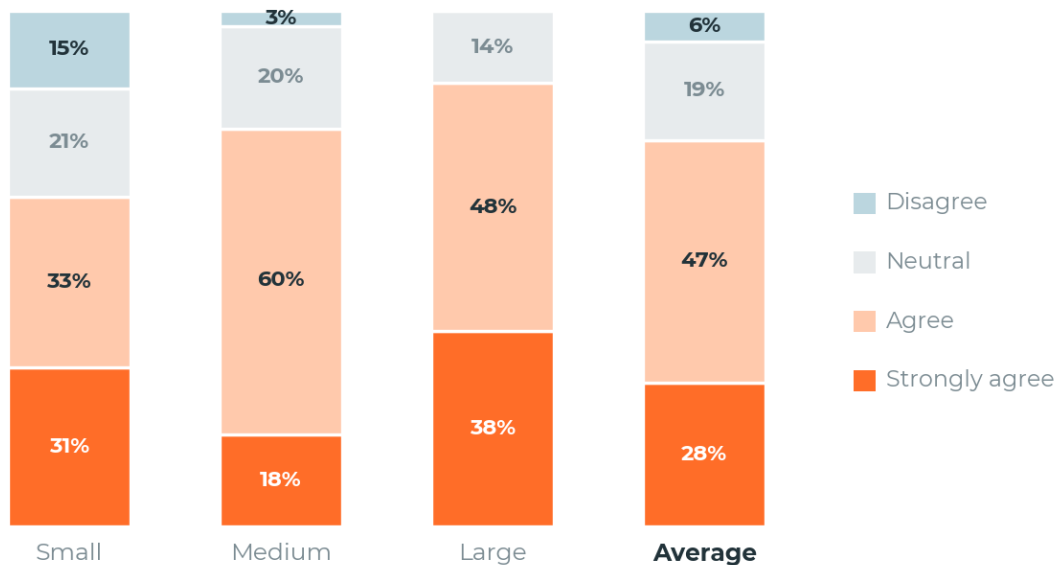
Figure 25: Requirements for in-call and post-call AI automation



THE ONBOARDING PROBLEM

System complexity has a second cost that is easy to overlook, which falls on new starters rather than on established agents.

Figure 26: "New agents find it difficult to familiarize themselves with our systems", by contact center size



75% agree or strongly agree that new agents struggle to familiarize themselves with the systems. Large operations report the greatest difficulty at 86% agreement, against 78% in medium-sized operations and 64% in small ones.

This matters for the business case because AI that surfaces the right answer at the right moment reduces what a new agent has to memorize before becoming productive. Whether it does so in practice is a measurable question rather than a theoretical one.

WHAT REAL-TIME ASSISTANCE IS MEASURED TO DELIVER

Average handling time is where real-time assistance most clearly delivers. 65% report a positive impact on AHT, of which 39% describe it as slight and 26% as significant, against 17% reporting no real impact and 4% a negative one.

Figure 27: Impact of real-time AI agent assistance on selected metrics



First contact resolution is close behind, with 54% reporting a positive impact. Customer satisfaction is the weakest of the four: 33% report no real impact, 28% a positive one, and 37% do not know. Agent onboarding shows the strongest slight-positive reading at 50%.

That pattern deserves attention. Assistance is making calls shorter and is also credited with improving first contact resolution. What it is not yet doing, on these figures, is moving customer satisfaction much.

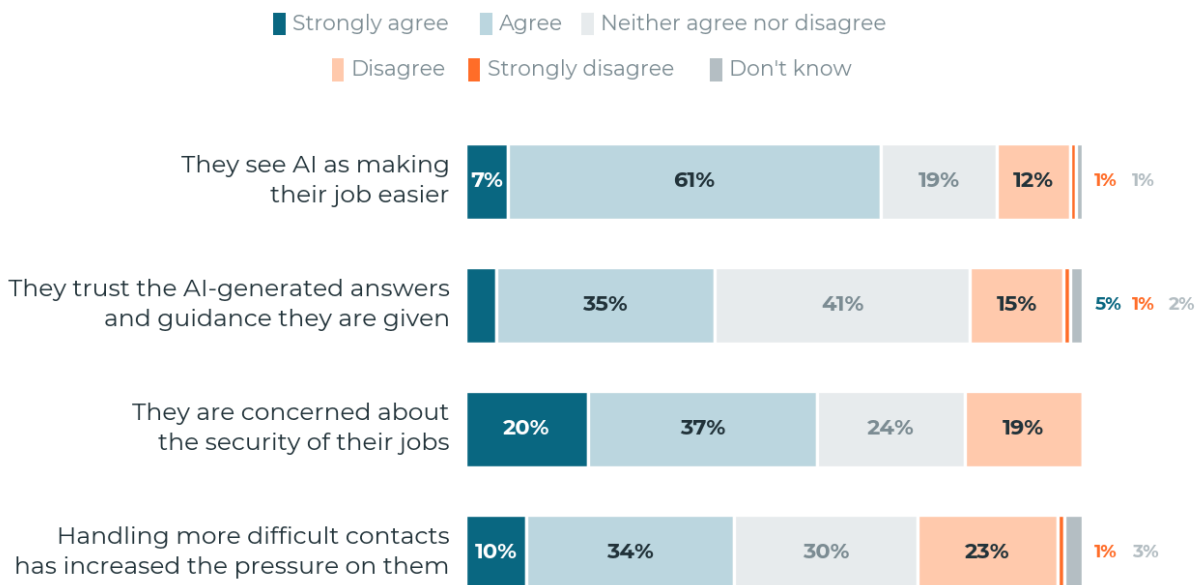
Across all four measures the don't-know proportion runs from 15% to 37%, highest on customer satisfaction. Time to competency in particular is rarely quantified exactly, so some of that is likely to reflect measurement methods rather than an absence of effect.

These figures come only from organizations already running real-time assistance, so the base is relatively small and the percentages should be read as indicative of direction rather than as precise quantities.

WHAT AGENTS MAKE OF IT

Almost all published research on agent assistance measures the effect on operational metrics. Very little measures what the agents themselves think, and the two are not the same thing: a tool that reduces handling time while being distrusted by the people using it is not a stable deployment.

Figure 28: Agreement with statements about how agents feel about AI



68% agree or strongly agree that their agents see AI as making the job easier, against 13% who disagree. That is a markedly more positive reading than the UK equivalent of 49%.

40% agree that agents trust the AI-generated answers and guidance they are given, 16% disagree, and 41% sit in the middle. Guidance that agents do not act on delivers nothing regardless of how accurate it is.

57% agree that agents are concerned about the security of their jobs, which is materially higher than the UK figure of 42% and sits alongside headcount data showing that US layoffs attributable to AI, while still uncommon, are more than twice as frequent as in the UK.

44% agree that handling more difficult contacts has increased the pressure on agents, which may well be storing up absence, burnout and attrition problems for the near future.

These are employers' readings of agent sentiment rather than responses from agents themselves, and should be treated as such. Managers who have invested in a technology are not disinterested observers of how their staff feel about it.

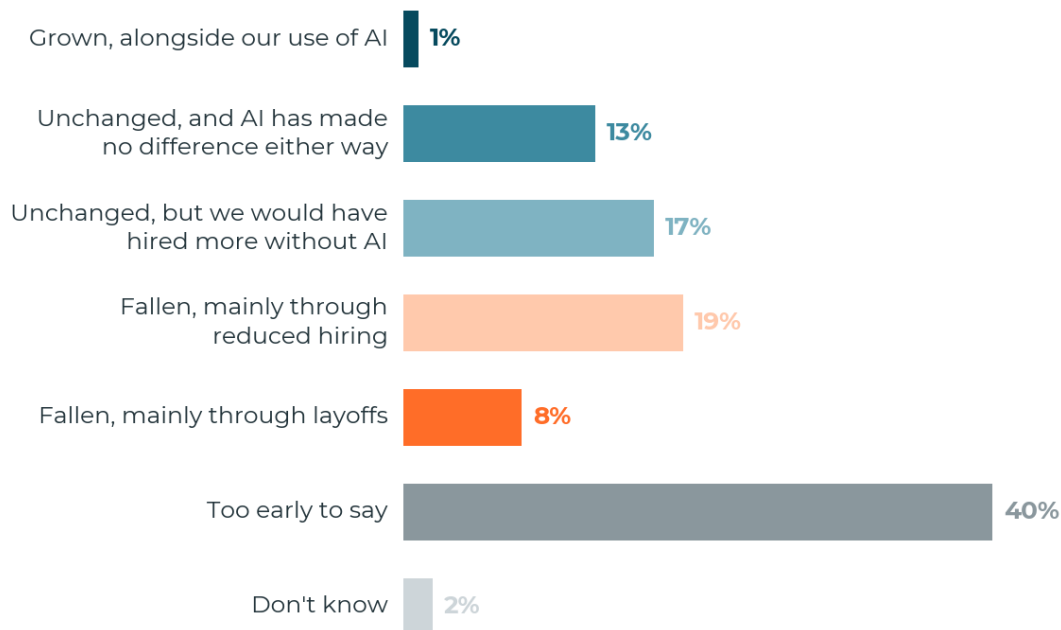
WHAT AI HAS DONE TO HEADCOUNT

Looking at the current effect of AI on agent headcount, the single largest response, at 40%, is that it is too early to say.

Among those willing and able to answer, the pattern is consistent and does not match the public discussion of this subject. 17% report headcount unchanged but state that they would have recruited more without AI, and 19% report headcount falling mainly through reduced recruitment. 13% report unchanged headcount with AI seen as irrelevant to this.

Only 8% report headcount falling mainly through layoffs, and 1% report growth alongside AI.

Figure 29: Effect of AI on the number of people employed in the contact center



So 36% describe an effect that operates through hiring rather than firing. This is consistent with the attrition data published elsewhere by ContactBabel, where an average US contact center attrition rate of 29% can therefore reduce headcount substantially without a layoff program.

The measures organizations took when introducing AI are covered in the survey, and formal training and consultation remain the most common at 72%.

25% changed agent targets or performance measures to reflect the new way of working, which is a low figure given that 44% believe the residual work has become harder.

19% did none of these things.

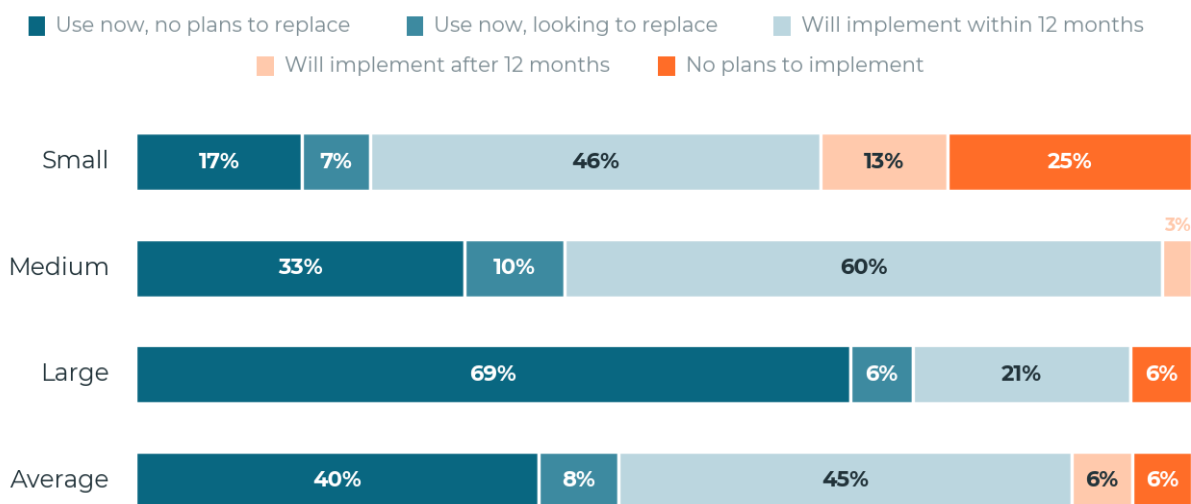
ANALYTICS, QUALITY AND ASSURANCE

This section is about what is done with the transcripts created through voice AI, how widely interaction analytics is deployed, what organizations use it for, how well quality assurance currently works, and the question that is rarely asked: how AI-handled conversations are themselves quality assured.

HOW WIDELY ANALYTICS IS DEPLOYED

48% of US contact centers use interaction analytics, comprising 40% with no plans to replace their system and 8% currently using but looking to replace or upgrade. A further 45% expect to implement within 12 months and 6% after that, leaving only 6% with no plans at all.

Figure 30: Use of interaction analytics, by contact center size



The size effect is pronounced, with large operations reporting 75% current usage against 43% in medium-sized operations and 24% in small ones. A quarter of small operations have no plans at all, while among large operations that figure is 6%.

Among those who use analytics, the functionality split is uneven.

Figure 31: Use of various interaction analytics functionality (from only those respondents who use analytics)

Interaction analytics type	% of users
Historical post-call speech analytics	86%
Multichannel analytics (email, web chat, social media)	53%
Desktop analytics	51%
Back office analytics	49%
Customer journey analytics	47%
Real-time speech analytics	28%

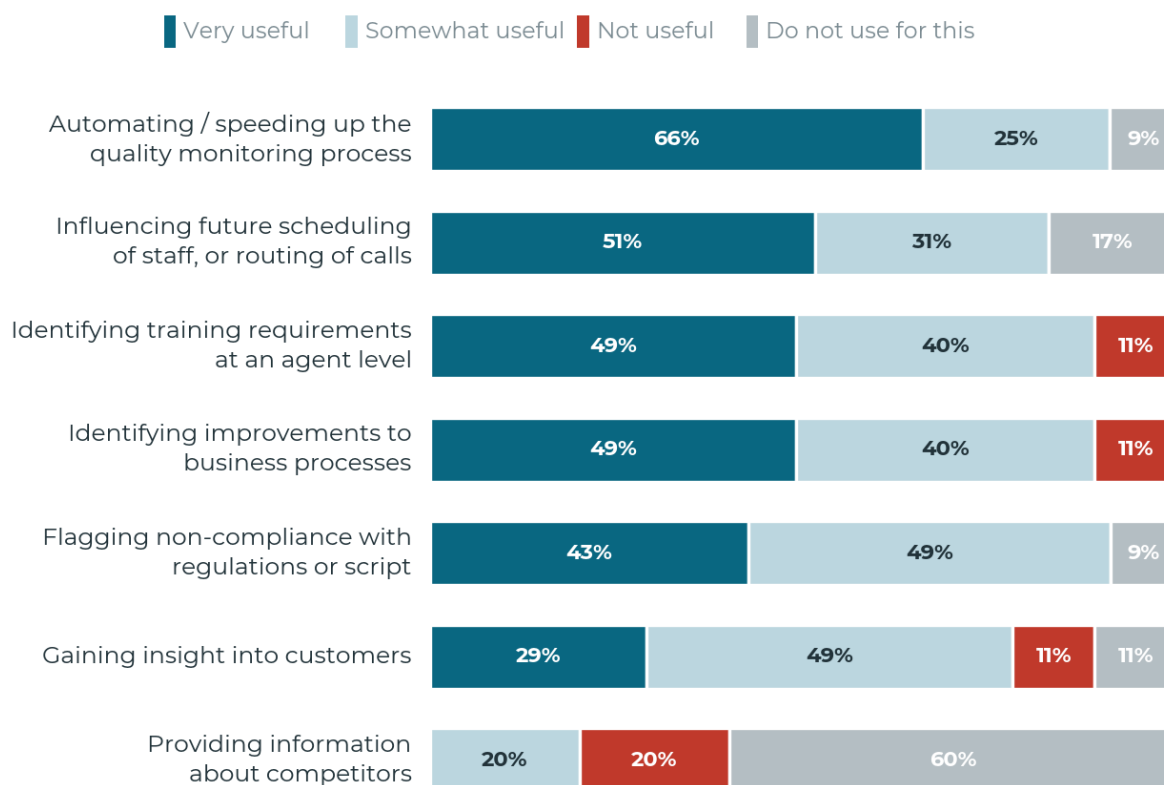
Post-call analysis remains the dominant use at 86%, while real-time analytics, which is the capability most closely associated with voice AI in supplier marketing, is used by only 28% of analytics users.

WHAT ANALYTICS IS USED FOR

Automating and speeding up the quality monitoring process is rated very useful by 66%, the highest figure for any application in either market. Influencing future scheduling or routing follows at 51%, then identifying training requirements and business process improvements at 49% each.

43% say the same of flagging non-compliance with regulations or script, and 42% of identifying training requirements at an agent level. Gaining insight into customers is less enthusiastically received, although is still positive.

Figure 32: Usefulness of post-call analytics



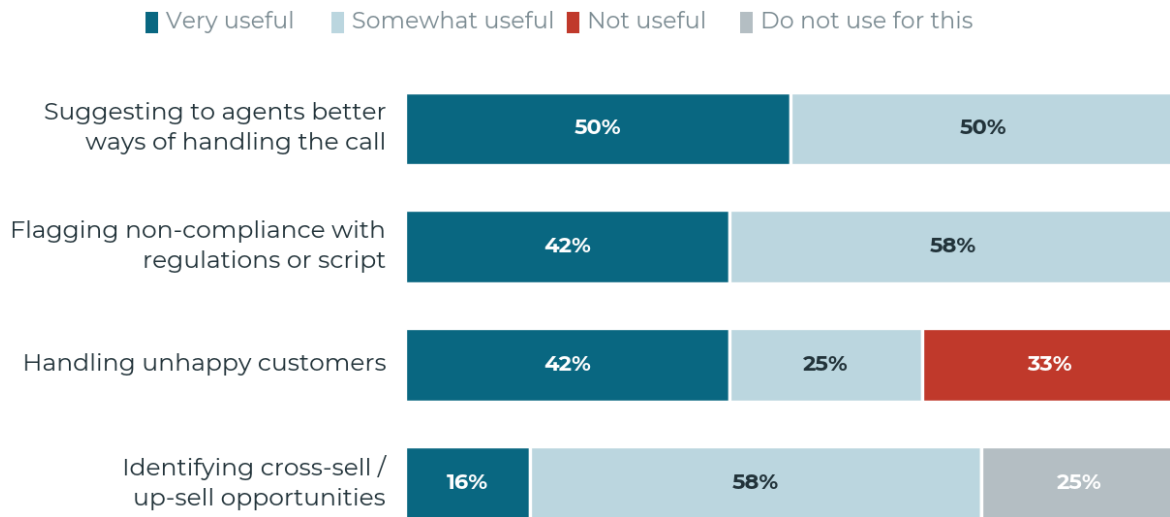
At the other end, providing information about competitors is not used by 60% of analytics users, and generally is not seen as particularly useful by those who do.

Among those who use real-time analytics, suggesting to agents better ways of handling the call is rated very useful by 50%, ahead of flagging non-compliance and handling unhappy customers at 42% each.

Identifying cross-sell and up-sell opportunities is rated very useful by 16%.

It should be noted that using analytics to handle unhappy customers in real time is reported to not be useful by 33%, suggesting that agents' flexibility and empathy is still heavily leant upon.

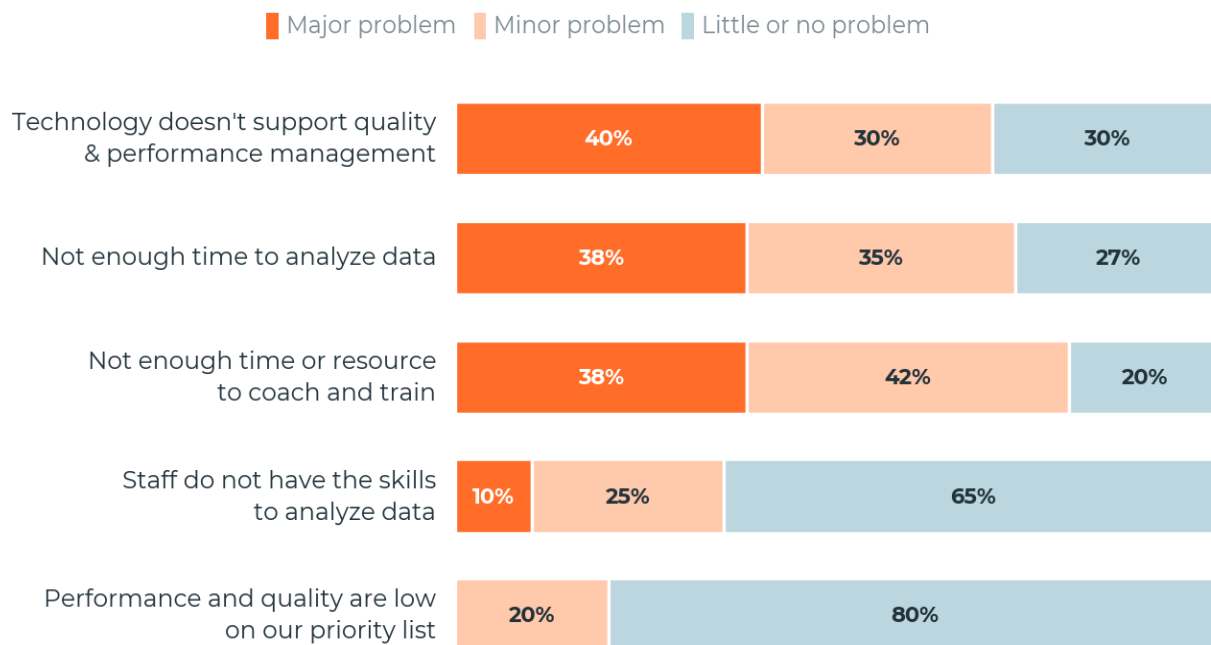
Figure 33: Usefulness of real-time analytics



WHY THE QUALITY PROCESS NEEDS THE HELP

The enthusiasm for automating quality monitoring makes more sense against the current state of quality assurance.

Figure 34: Challenges to managing performance & quality



38% describe not having enough time or resource to coach and train as a major problem, and a further 42% as a minor one, so 80% experience it to some degree.

Not having enough time to analyze data reaches 73%, with 38% calling it major. Technology not supporting quality and performance management is the single largest major problem at 40%.

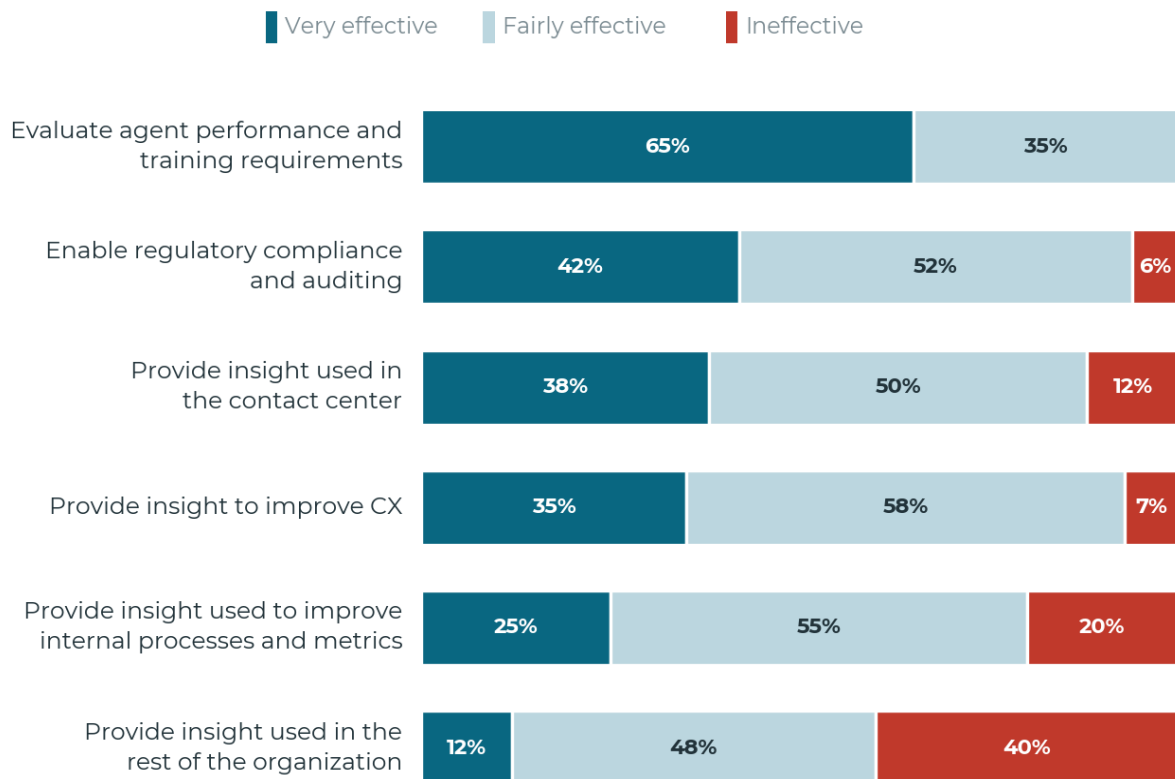
The constraints are different in kind and matter for what AI can realistically fix. Analysis capacity is a computation problem, and automated scoring addresses it directly. Coaching capacity is a people problem, and automating the analysis does not create the supervisor hours needed to act on it. An organization that automates scoring without freeing up coaching time will simply generate findings faster than it can respond to them.

It is worth noting what respondents do not report as a problem. 80% say that performance and quality being low on the priority list is little or no problem, and 65% say the same of staff lacking the skills to analyze data. The constraint is time and tooling, not intent or capability.

Where quality assurance is applied to its traditional purpose it works well: 65% rate it very effective for evaluating agent performance and training requirements, and 42% for enabling regulatory compliance and auditing.

Effectiveness falls away as the insight travels further from the contact center floor, reaching only 12% very effective for insight used in the rest of the organization, where 40% call it ineffective.

Figure 35: Effectiveness of quality assurance

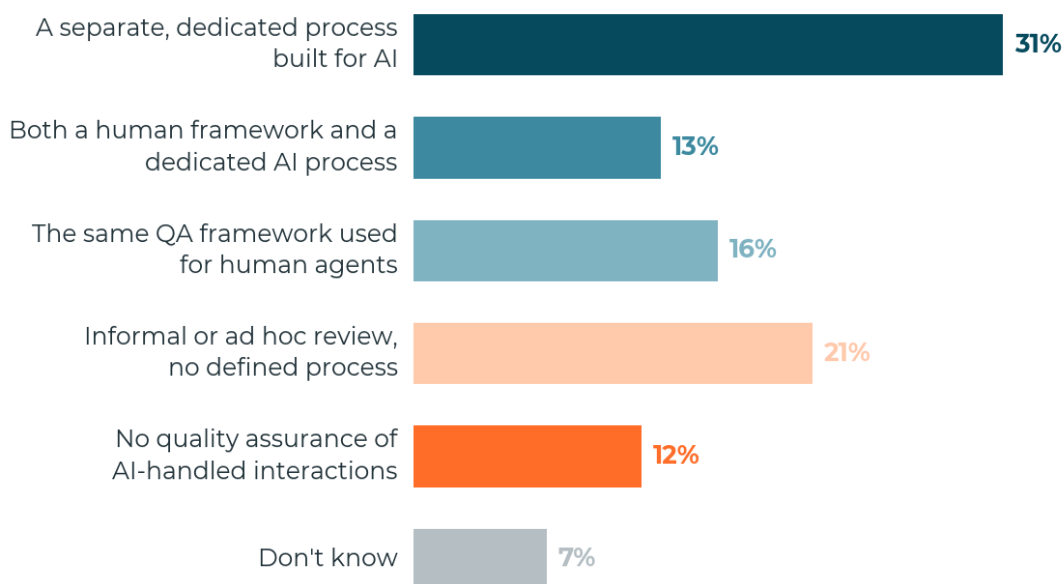


ASSURING THE AI ITSELF

The previous figures concern the quality assurance of human agents.

As voicebots and AI-drafted responses take on a larger share of customer contact, the same questions apply to them: whether the answers were correct, whether the tone was appropriate, whether regulatory requirements were met. Whether organizations are actually asking those questions of their AI is not something published research has often measured.

Figure 36: How organizations quality assure interactions handled by AI



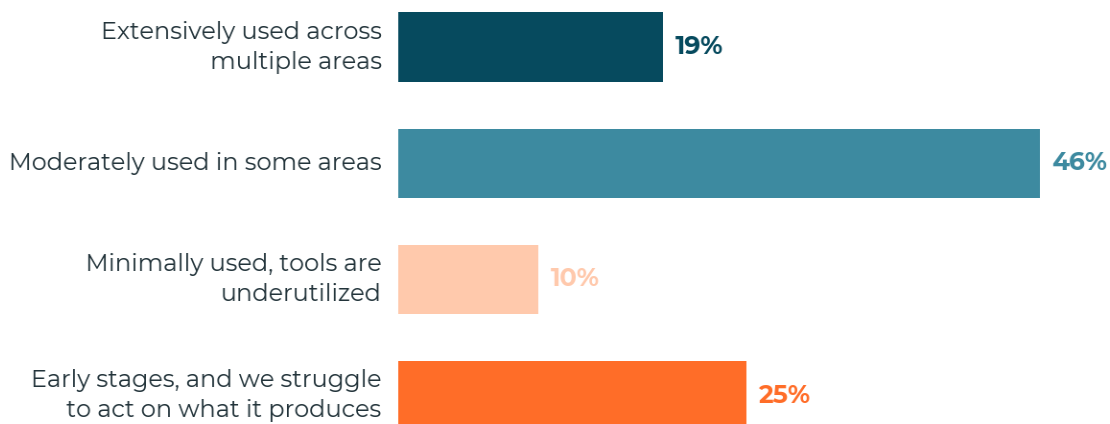
31% have built a separate, dedicated process for assuring AI-handled interactions, and a further 13% run both that and their human framework in parallel. 16% apply the same QA framework they use for human agents.

On the other hand, 12% carry out no quality assurance of AI-handled interactions at all, 21% review informally with no defined process, and 7% do not know. Two in five organizations therefore have no systematic way of establishing whether the conversations their AI handles are correct, compliant or acceptable in tone.

The earlier controls data showed that 18% have no formal controls over what bots may do and 18% do not know. An organization which lets a bot act in live business systems, applies no defined controls to those actions, and performs no quality assurance on the resulting conversations has automated a part of its operation that it cannot see.

Only 19% report that AI-driven analytics is used extensively across multiple areas. 46% describe moderate use in some areas, 25% say they are at an early stage and struggle to act on what it produces, and 10% say the tools are underutilized.

Figure 37: Extent to which AI-driven analytics is generating actionable operational insight



The US position is better than the UK here, where only 11% reported extensive use. Even so, the coaching bottleneck shown earlier reappears one layer up: 80% of operations report insufficient time to coach and train, and automating the analysis moves that constraint rather than removing it. An organization that buys analytics without also creating the supervisory capacity to respond will have exchanged one queue for another.

Automating the analysis moves the constraint rather than removing it, and an organization that buys analytics without also creating the supervisory capacity to respond will have exchanged one queue for another.

VOICE, IDENTITY AND FRAUD

Before most US calls can proceed, the caller has to prove who they are. This chapter covers how much of the call that consumes, what it costs, why voice biometrics has remained a marginal technology despite an obvious fit, and the reason that position may now be changing in both directions at once.

THE SCALE OF THE AUTHENTICATION PROBLEM

73% of US inbound calls require a security and identification process to be completed first, and 84% of those are handled by an agent.

The average agent-handled check takes 42 seconds, which against a mean call duration of 411 seconds means that just over 10% of the call is spent establishing who is on the line.

44% of respondents state that all callers go through identity verification, and only 8% that they never do.

Figure 38: Caller identity authentication methods (only those contact centers which authenticate some or all calls)

Identification method	Proportion of callers identified using this method
Agent	84%
DTMF IVR (touchtone)	14%
Speech recognition	4%
Voice biometrics	Under 1%
NB: figures add to more than 100% as multiple authentication methods may be used on a call	

ContactBabel estimates the cost of time spent on agent-handled security and identification checking at 73.4 cents per call, and \$12.3bn a year across the US industry.

This is the largest single quantified inefficiency in this report, and unlike most of the others it applies to a task that is entirely mechanical and which adds nothing positive to the customer or agent experience.

FRAUD, AND THE PROBLEM WITH USING A VOICE AS A CREDENTIAL

Voice biometrics generates a voiceprint from spoken words, calculating vocal tract measurements alongside pronunciation, emphasis, accent and speech rate. The resulting template is not affected by the caller having a cold, using a different handset, or ageing. On paper it removes most of a 42-second process.

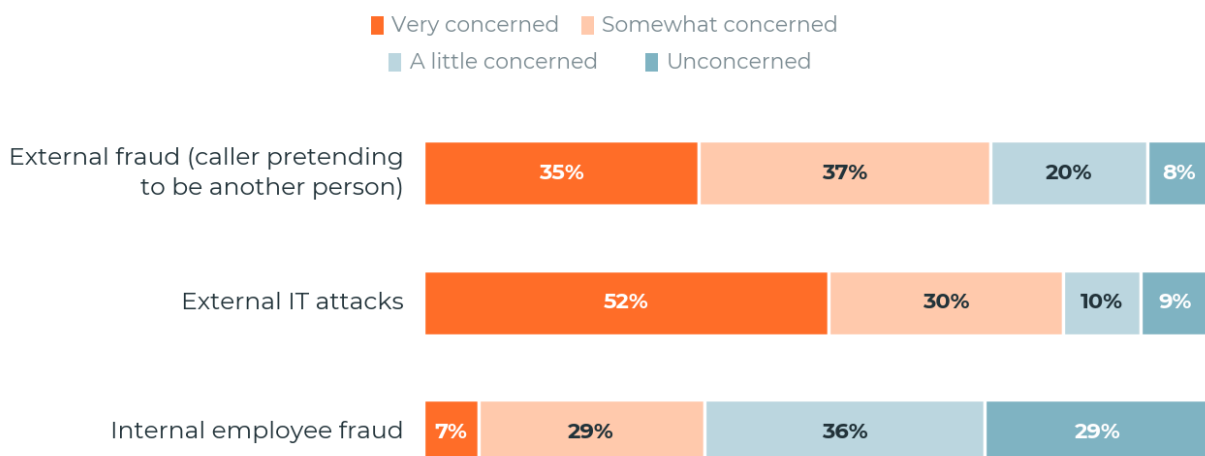
It has nonetheless remained under 1% of authentications for as long as ContactBabel has measured it. Investment intent, however, has increased and is seen as a top 5 investment priority by 7% of US contact centers, from a list of 27.

There are two things to note: first, that passive authentication placed in front of a voicebot removes the 42 seconds before a human is involved at all, which strengthens the case for voice self-service independently of whether the bot resolves the enquiry.

72% of contact centers are very or somewhat concerned about external fraud from callers pretending to be someone else, and 82% about external IT attacks, where 52% are very concerned.

Internal employee fraud attracts markedly less concern, with 29% unconcerned and only 7% very concerned which is notable given that around half of US agents work remotely or in a hybrid pattern.

Figure 39: Concerns about security

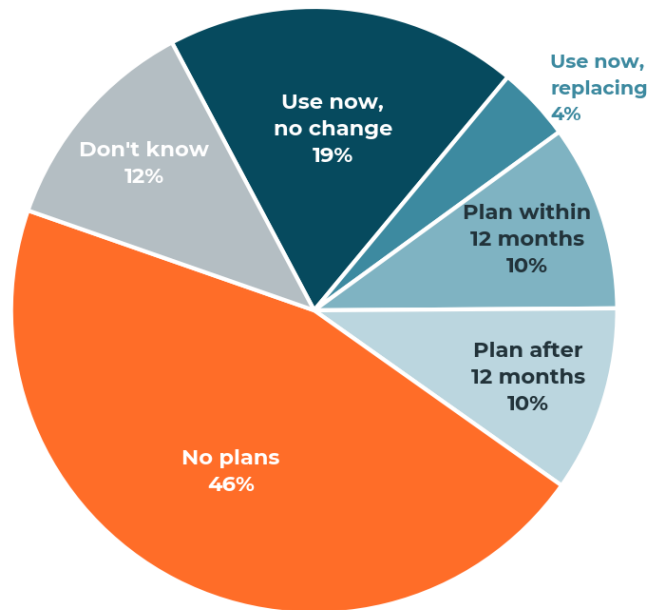


The difficulty for voice biometrics specifically is that generative voice models can now synthesize a convincing imitation of a particular speaker from a short sample. A credential that can be produced from a recording of the customer speaking is a weaker credential than it was when the technology was designed.

Figures circulating in the market on the growth of deepfake fraud attempts originate largely from vendor and identity-provider research rather than from independent survey work.

What can be said is that the direction of the risk is not disputed by anyone selling into this space, and that suppliers have responded by treating the voice as one signal among several rather than as a single factor.

Figure 40: Current and planned use of voice biometrics or AI-based authentication and fraud detection



23% report using voice biometrics or AI-based authentication and fraud detection, with a further 20% planning to implement. 46% have no plans at all and 12% do not know. Adoption is more than twice the UK level.

While the previous finding that voice biometrics accounts for under 1% of actual authentications might make this 23% figure seem strange, it can be reconciled as the survey question asked covers both AI-based authentication and fraud detection as well as voice biometrics proper.

It captures organizations running behavioral or device-based risk scoring, transcript-based fraud detection, or anomaly detection on call patterns, none of which identify anyone by voice. The 23% is therefore AI applied to the fraud and identity problem generally, and the under 1% figure is how many callers are actually identified solely by their voice.

The more useful reading is that 46% of the market has no plans in this area at all. Authentication is simultaneously the largest single mechanical inefficiency in the US contact center, at \$12.3bn a year, and an application close to half the market is not addressing.

Where that reluctance comes from is not something this survey establishes, though the arrival of synthetic voice as a credible attack on the obvious solution is unlikely to have helped.

PROACTIVE OUTBOUND AND THE US RULES

Everything so far in this report concerns calls the customer makes. This section concerns calls the organization makes, which is where voice AI meets a body of regulation written three decades before the technology existed and since applied to it directly.

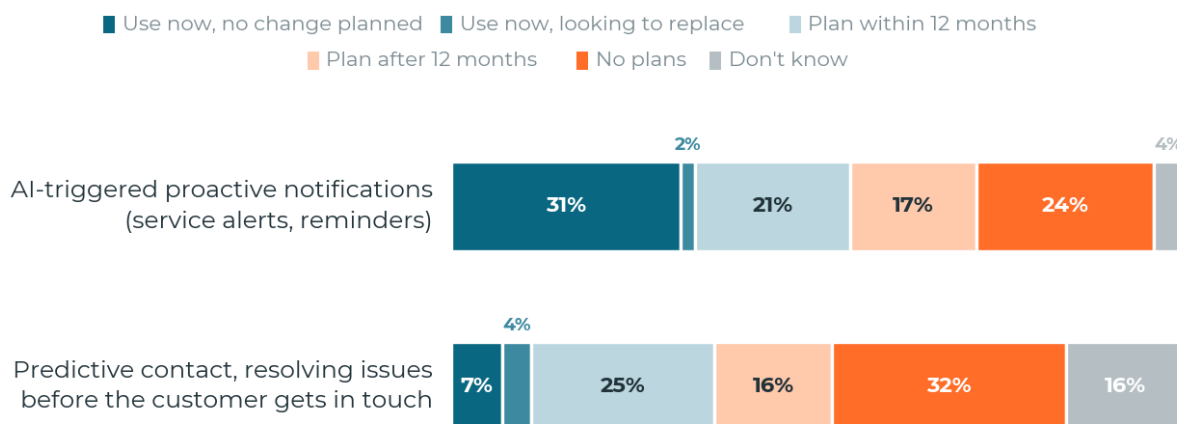
While the commercial case is quite straightforward, the compliance position is not, and in the US the principal exposure is not the regulator.

THE COMMERCIAL CASE

Proactive voice covers service alerts, appointment reminders, delivery and payment follow-ups, and resolving issues before the customer becomes aware of them.

The attraction is that it converts an inbound call the organization would have paid \$7.20 to handle into a short outbound contact it controls the timing of, and in many cases prevents the enquiry arising at all.

Figure 41: Current and planned use of proactive and predictive outbound AI



33% already use AI-triggered proactive notifications such as service alerts and appointment reminders, and 11% use predictive contact to resolve issues before the customer gets in touch. A further 38% and 41% respectively plan to implement.

It is worth being clear that neither application is confined to voice. A proactive notification is at least as likely to be sent by SMS or email as spoken down a telephone line, and the analysis below concerns the voice element specifically. Text messages carry their own TCPA exposure, which is beyond the scope of this report.

A KEY DISTINCTION TO UNDERSTAND

Please note that the material in this section is our current best understanding of the US position and is offered as general guidance rather than as legal advice. Regulation in this area is moving quickly, and several of the developments described below are recent enough that readers should verify their current status. Nothing here should be acted upon without first taking advice from a qualified US telecoms attorney.

The Telephone Consumer Protection Act of 1991, and the FCC rules implementing it at 47 C.F.R. § 64.1200, draw their sharpest line between telemarketing and everything else. Marketing calls attract the most demanding consent standard; informational and transactional calls attract a lower one, though as set out below they are not unregulated.

An appointment reminder, an outage alert, a fraud alert or a delivery update relating to an existing relationship is generally informational rather than telemarketing. A call that promotes a product or service is telemarketing whatever it is called internally, and our understanding is that the test is the content and purpose of the call rather than the label on the campaign.

Most legitimate proactive voice AI in a contact center appears to us to sit on the informational side of that line. Organizations whose outbound carries any promotional element may wish to assume the stricter standard applies.

WHAT THE RULES REQUIRE

The pivotal development for this report is the FCC's Declaratory Ruling of 8 February 2024, which confirmed that a voice generated by artificial intelligence is an "artificial or prerecorded voice" within the meaning of the TCPA. The Commission stated that the statute does not allow a carve-out for technologies that purport to provide the equivalent of a live agent. The ruling took effect on release.

The practical effect is that a conversational voice agent is treated exactly as a pre-recorded message is, however lifelike it sounds and however capable it is of responding in real time. Where the UK position on the same question remains unsettled, the US has answered it explicitly, and answered it in the way that imposes the higher burden.

The prohibitions sit in two places. Calls to a residential line using an artificial or prerecorded voice fall under 47 U.S.C. § 227(b)(1)(B). Calls to a cell phone using an automatic telephone dialing system or an artificial or prerecorded voice fall under § 227(b)(1)(A)(iii). The implementing rules are at 47 C.F.R. § 64.1200(a)(1) to (a)(3).

Consent follows from that, and differs by line type. Prior express written consent is required before an artificial or prerecorded voice call may be made for telemarketing purposes, whether to a cell phone or a residential line. For non-telemarketing calls to a cell phone, prior express consent is required but need not be written. Screening against the National Do Not Call Registry does not substitute for either: consent is the test.

Residential lines are treated differently again, and this is the point most likely to catch out a service notification program. Our understanding is that certain non-telemarketing prerecorded calls to residential lines are exempt from the consent requirement, including specified healthcare, package delivery and financial institution fraud messages, but that the exemptions carry conditions. Since July 2023 those conditions have included a numeric cap, which we understand to be three calls within any consecutive 30-day period, together with an opt-out mechanism. An outage alert or appointment reminder program running at higher frequency than that may need consent even though the calls are plainly not marketing.

On whether a human placing the call changes the position, the Commission has addressed the point directly. In its 2020 Soundboard ruling, relied on in the AI ruling itself, the FCC held that the presence of a live agent selecting which prerecorded message to play does not negate the statutory prohibition against initiating a call using a prerecorded or artificial voice. On that reasoning, a campaign is unlikely to be brought outside the rules by having a person dial and an AI conduct the conversation.

The FCC's telemarketing rules also require that a prerecorded telemarketing call state the identity of the business responsible for initiating it at the beginning of the message, give a telephone number at which that business can be reached, and provide an automated, interactive opt-out mechanism. The abandoned call rules permit no more than 3% of live calls answered by a person to be abandoned, measured across a single calling campaign and recalculated for each successive 30-day period where a campaign runs longer. A call counts as abandoned if it is not connected to a live representative within two seconds of the called person's completed greeting.

One detail in that rule bears directly on system design. Where the subscriber has given prior express written consent, our understanding is that a prerecorded telemarketing call is not treated as abandoned provided the message begins within two seconds of the customer's greeting. For a consented AI campaign, that turns opening latency from a matter of good manners into a compliance parameter.

A further Notice of Proposed Rulemaking issued in 2024 proposed defining an "AI-generated call" and requiring disclosure at the start of such calls. Our understanding is that these proposals had not been finalized at the time of writing, and readers should check their current status.

WHERE THE EXPOSURE COMES FROM

The TCPA carries statutory damages of \$500 per violation, rising to \$1,500 where the violation is willful or knowing, with no aggregate cap. Because those damages attach per call and the statute provides a private right of action, the practical exposure for a campaign at scale comes from class action litigation rather than from regulatory penalty. An organization placing a million calls without adequate consent is not facing a fine in the ordinary sense.

Two recent developments that have been reported are worth noting, both of which we understand to be current but neither of which we have been able to verify against primary sources.

The first is that a federal appeals court is reported to have vacated the FCC's so-called one-to-one consent rule in early 2025, with the Commission rescinding it later that year, leaving the earlier written consent standards in place. Any organization that has been building its consent capture around the one-to-one rule may wish to check where that now stands.

The second is that revocation rules finalized in 2024 are understood to have taken effect in 2025, and to give a revocation received through one channel effect across other communications from the same sender.

State law sits on top of the federal position and is fragmenting. Several states operate their own mini-TCPA statutes with their own consent standards and private rights of action, and a number have introduced AI-specific requirements covering disclosure and automated decision-making. A national campaign built only to the federal floor is unlikely to satisfy every state it reaches.

Illinois deserves separate mention for anything touching voice biometrics. The Biometric Information Privacy Act treats a voiceprint as biometric identifier data and requires written notice and consent before collection, with a private right of action. That bears directly on the authentication material earlier in this report as well as on outbound.

Federal and state privacy law sits over all of it, including transparency about what the call is doing, the retention of call audio and transcripts, and the treatment of any automated decision the call produces.

ENFORCEMENT AND THE DIRECTION OF TRAVEL

The February 2024 ruling followed a series of incidents in which cloned voices were used in robocalls, including one impersonating a sitting president ahead of a state primary, and the Commission was explicit that voice cloning was the conduct it had in mind. We understand that the carrier which transmitted those particular calls subsequently settled with the FCC for \$1m.

Two points transfer directly to legitimate deployments. The first is that the ruling turns on how the voice is produced rather than on how human it sounds, so a well-engineered conversational agent is in scope precisely because it is convincing. The second is that presenting an automated system as a named human, without disclosure, appears to us to compound the exposure rather than reduce it. The identification requirements that apply to prerecorded telemarketing calls are unlikely to be satisfied by a system that introduces itself with a first name and nothing else.

It seems clear to us that an organization running automated outbound at scale in the US carries a materially different risk profile from one running the same campaign in the UK, and that the difference is the availability of class action rather than the size of any regulatory fine.

DISCLOSURE

There is no general federal statute requiring disclosure that a caller is speaking to a machine, although the FCC has proposed one and several states have legislated in the area. There seem to us to be three reasons to disclose anyway.

The first is that the identification requirements already applying to prerecorded telemarketing calls sit uncomfortably with an undisclosed AI persona, and the FCC has shown where its concerns lie.

The second is that customers say they want it: as shown earlier, 54% of US consumers are concerned about hidden or unclear use of AI, meaning not knowing whether they are dealing with a person or a system.

The third is that state AI disclosure requirements appear to be proliferating, with several states having legislated on in-call disclosure, and an organization operating nationally will end up disclosing in some states regardless. Building disclosure in once may prove simpler than maintaining a state-by-state script.

For institutions regulated at federal or state level, including financial services and healthcare, sector rules on fair treatment, vulnerable customers and record-keeping apply to an automated outbound call as they would to a human one.

POINTS A BUYER MAY WISH TO CONSIDER

The following are offered as questions worth asking rather than as a compliance procedure. What is appropriate will depend on the campaign, the customer base and the states involved, and none of it substitutes for advice.

- Whether each campaign is telemarketing or informational, judged on content rather than label, and whether that reasoning has been written down anywhere.
- Whether anything with a promotional element has prior express written consent behind it, and whether the evidence of that consent could be produced if asked.
- Whether consent provenance is recorded against each number, including when and how it was obtained and what it covered.
- Whether informational campaigns to residential lines are running within whatever numeric limits apply to the exemption being relied on, since frequency rather than content may be the binding constraint.
- Whether National Do Not Call and applicable state registries are screened regardless, and whether suppression operates across campaigns rather than within them.
- What the opening latency actually is under load. Where consent is in place, the two-second window before the message begins is a defined threshold rather than a quality target.
- Whether the call identifies the calling business at the outset, discloses that it is automated in plain terms, and avoids presenting itself as a named individual.
- Whether an automated opt-out is offered and honored, and whether revocation is handled within the conversation rather than only at a closing menu.
- Whether the abandoned call rate is monitored against the 3% allowance, and on what basis the campaign period is defined.
- Whether calling line identification is accurate and dialable.
- How long call audio, transcripts and any voiceprints are retained and on what basis, with particular care where Illinois residents may be called.

None of this is unique to AI. What is new is the volume such systems make possible, and the fact that statutory damages attach to every call.

IMPLEMENTATION: WHAT BREAKS

This chapter deals with the practical failure modes of voice AI: the technical characteristics that make voice harder than text, the integration burden, where organizations are sourcing the capability, and what has been paused or abandoned after deployment.

WHY VOICE IS HARDER THAN TEXT

A chatbot receives clean, discrete, already-textual input. A voicebot receives none of those things, and the differences matter.

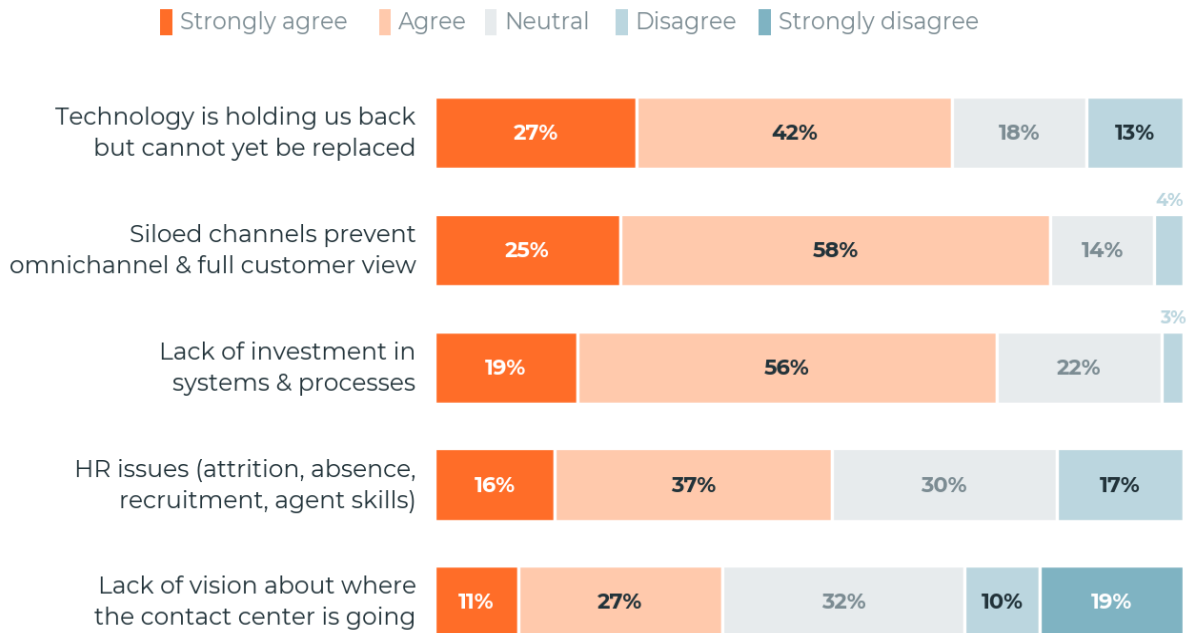
- **Non-linear speech.** Conversations do not follow linear paths. People repeat themselves, rephrase halfway through a sentence, and interrupt. A text interface never has to cope with any of that.
- **Latency.** Users of automated speech recognition have long been accustomed to pauses of several seconds while the system processes and replies. Suppliers state that around half a second is now achievable with a well-configured voicebot, but latency remains something to establish before committing rather than after.
- **Intent detection.** People are more long-winded when speaking and include information that is not relevant. Systems also have to separate multiple intents within a single utterance, such as a request to book tickets that specifies route, number of travelers, date and arrival time in one sentence.
- **Recognition accuracy.** Accent recognition has improved substantially but is not infallible, and similar-sounding words remain a documented failure mode.
- **Pacing.** A pause in a text conversation is expected and gives the customer time to think. In speech it does not, and customers can feel rushed, give the wrong information, or want to start again.
- **No visual channel.** A voicebot cannot show the customer anything, so implementations are generally narrower in scope than the equivalent chatbot.
- **Voice persona.** Many voice assistants sound robotic or over-produced, or do not share the national accent of the typical customer. Some organizations vary the persona by customer segment.

The dependency on multiple components creates a compound risk that is easy to underestimate. Speech-to-text, natural language processing and noise cancellation all have to work together, quickly and reliably, and some of those components are external services. Each additional link is a further point of failure, and the resilience of the whole is not the resilience of the best part.

WHAT IS HOLDING CONTACT CENTERS BACK GENERALLY

Voice AI is being implemented into operations that already report significant constraints.

Figure 42: What is preventing the contact center from achieving its aims?



Siloed channels preventing omnichannel working and a full view of the customer is the most widely agreed constraint at 83% agreement, followed by lack of investment in systems and processes at 75% and technology holding the operation back at 69%. Lack of vision attracts the least agreement at 38%.

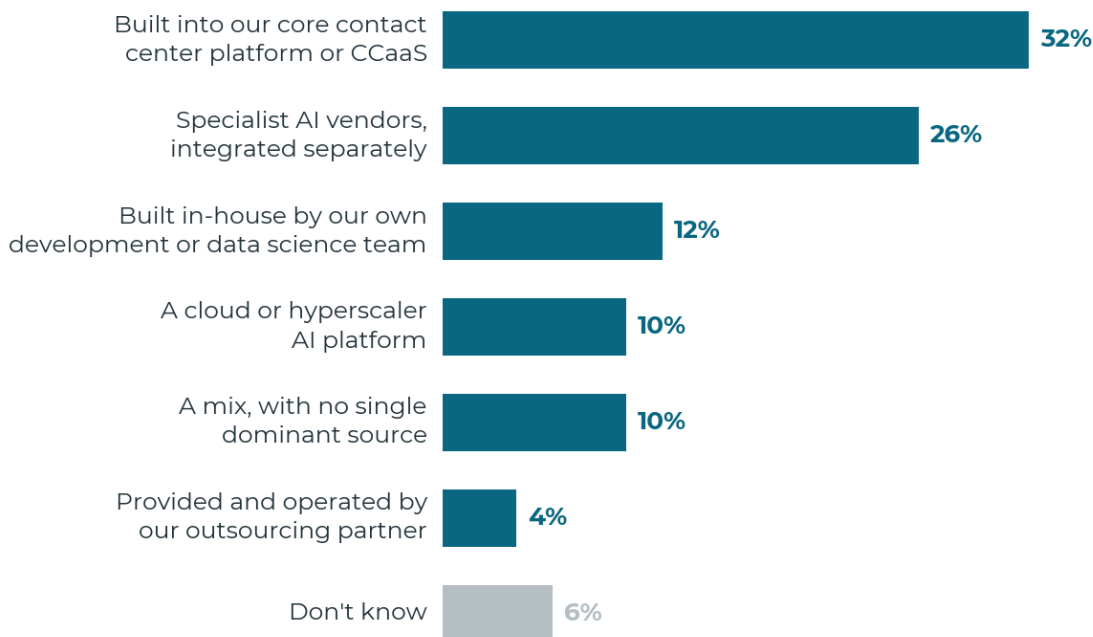
The first three of those bear directly on voice AI. An organization whose channels are siloed cannot easily give a voicebot the context it needs to resolve an enquiry, and one whose core technology cannot yet be replaced will be integrating AI into systems that were not designed to be integrated with.

Neither is an argument against proceeding, but both change what a realistic first deployment looks like.

SOURCING, AND WHAT GOT ROLLED BACK

Where voice AI capability comes from is the question with the greatest commercial consequence for both buyers and suppliers, and the one on which published evidence is thin.

Figure 43: Where the majority of AI capability in the contact center comes from



32% source the majority of their AI from the core contact center platform or CCaaS provider, which is the single largest category but well short of a majority. 26% use specialist AI vendors selected separately, 12% build in-house, and 10% each use a hyperscaler platform or describe a mix with no dominant source.

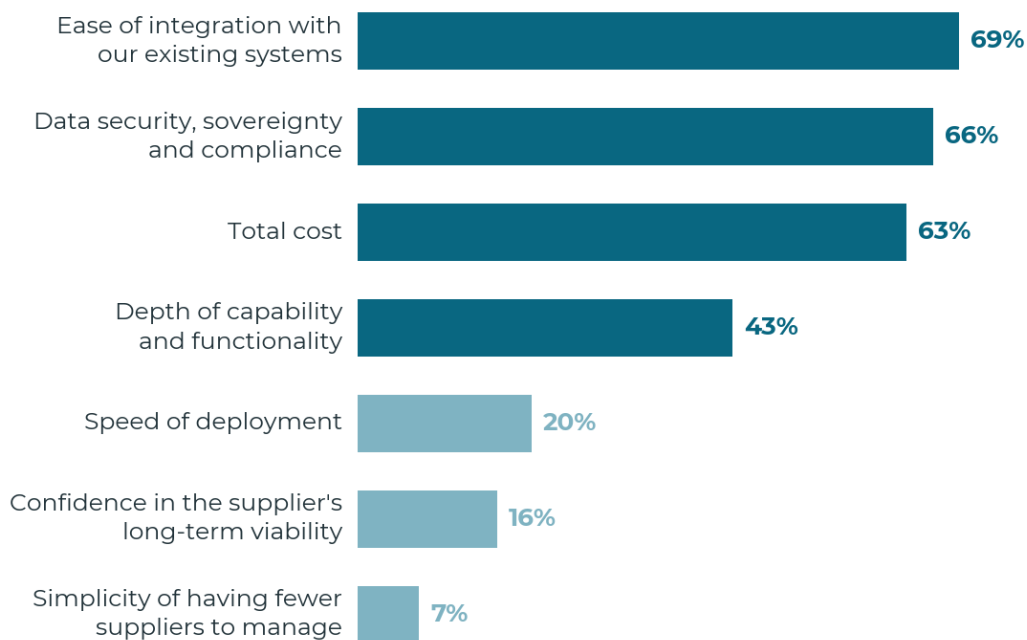
The forward view is similarly unresolved. 28% expect to build more in-house, 18% to move towards specialist vendors, and 31% to consolidate onto their core platform provider in one form or another. 14% expect no significant change.

There is no settled consensus about where this market is going. A platform vendor should note that fewer than a third of buyers intend to consolidate onto it; a specialist vendor should note that more organizations intend to build the capability themselves than to buy it from a specialist. The in-house figure is the one most likely to be revised downwards by experience.

When considering what matters most when deciding where to source AI, survey respondents could select up to three factors.

Ease of integration with existing systems leads at 69%, followed by data security and compliance at 66%, total cost at 63% and depth of capability at 43%. Those four account for almost the whole distribution.

Figure 44: What matters most when deciding where to source AI (top 3 answers)



Speed of deployment matters most to only 20% of buyers, and simplicity of having fewer suppliers to manage to 7%. “Deploy in days” is among the most common claims made for voice AI products, and it addresses a criterion that only one buyer in fourteen treats as a priority. Confidence in the supplier’s long-term viability, at 16%, is similarly low given how young many of the vendors in this category are.

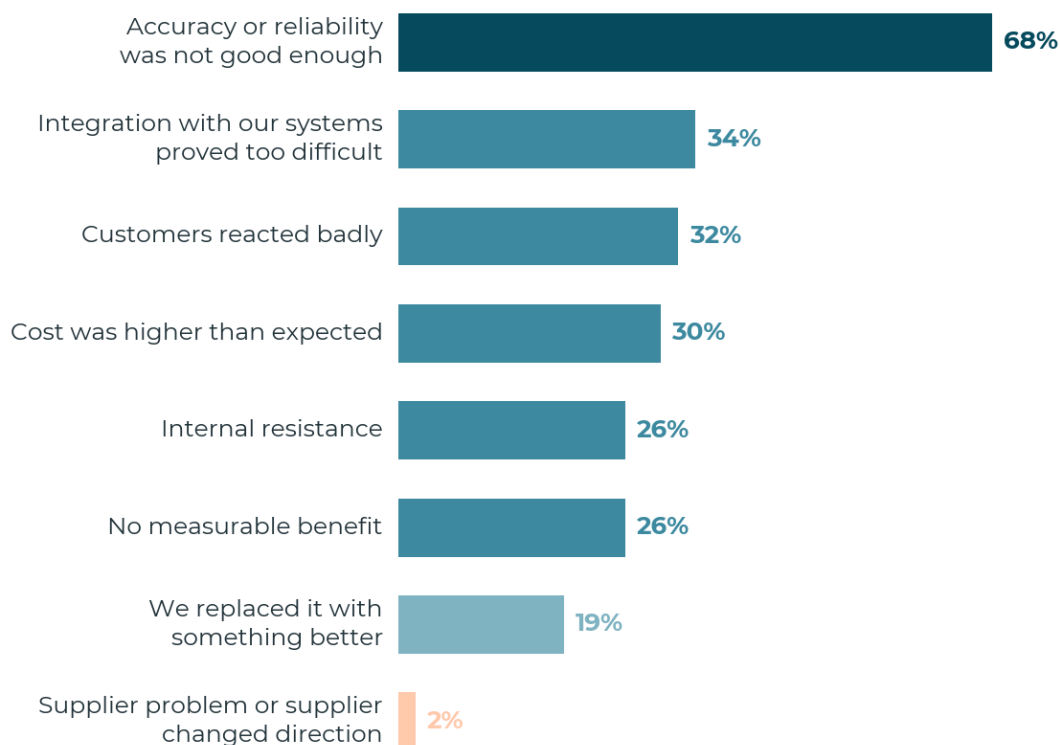
The practical implication for anyone building a business case is that a proposal which leads on speed and convenience is answering questions the buyer is not asking, while the questions they are asking, about data handling and integration, are the ones where evidence is hardest to produce.

WHAT GOT STOPPED

39% of respondents have stopped using, paused or rolled back an AI solution after deploying it. 50% have not and 11% do not know.

64% stopped at a limited pilot and 21% at proof of concept or demonstration, which together describe evaluation working as intended. 15% stopped in full production.

Figure 45: Main reasons for stopping, pausing or rolling back an AI solution



Accuracy or reliability not being good enough is the leading reason by a wide margin at 68%, with integration proving too difficult following at 34%, customers reacting badly at 32%, and cost being higher than expected at 30%.

Two figures are worth drawing out. Customers reacting badly accounts for 32%, which suggests many US deployments are reaching customers in volume before their weaknesses are found. 19% stopped because they replaced the solution with something better, which describes a market moving faster rather than retreating.

This is valuable data precisely because nobody has an incentive to publish it. Suppliers do not, and buyers rarely volunteer it. A 39% rate of abandonment, with accuracy the leading cause, is the appropriate context for any proposal that quotes a deployment timeline in days.

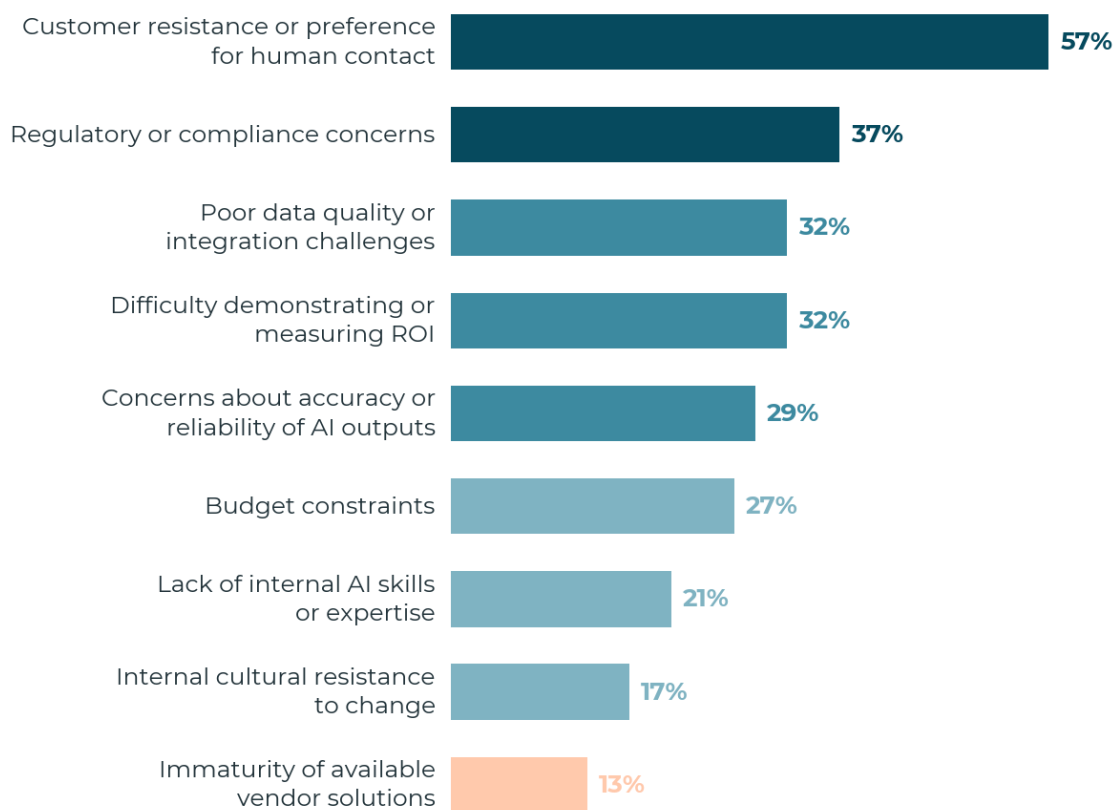
WHAT IS BLOCKING FURTHER ADOPTION

We asked survey respondents what would block further adoption of AI.

Customer resistance or preference for human contact is the leading barrier at 57%, well ahead of regulatory or compliance concerns at 37% and poor data quality at 32%.

That it heads the list is a key finding, and it is the one barrier on which this report has independent corroboration: 34% of US consumers say they would never prefer AI to a human, 63% are concerned about difficulty reaching a human agent, and consumers rate AI worse than a human on five of six service dimensions. Businesses and their customers agree, and they are describing the same constraint from opposite sides.

Figure 46: Most significant barriers to further AI adoption over the next two to three years



It is also the barrier least susceptible to the remedies usually proposed. Skills gaps at 21% and data quality at 32% are all addressable with money and time. A customer who does not want to speak to a machine is not addressable with either, and the organizations that make progress against this constraint will be those that automate the contacts customers are content to have automated rather than those that push hardest. Vendor immaturity, at 13%, is the lowest-ranked barrier: buyers do not believe the products are the problem.

MEASURING RETURN, AND WHAT TO DO NEXT

This chapter covers how voice AI has performed against the expectations set for it, how AI-handled contacts compare with human-handled ones, and what the evidence in this report supports as a sequence of investment.

THE BUILDING BLOCKS OF A BUSINESS CASE

Whatever a supplier quotes, a defensible business case for voice AI is built from a small number of measurable quantities, most of which appear earlier in this report. They are collected here so a reader can substitute their own figures.

Quantity	US figure
Mean cost per inbound call	\$7.20
Mean call duration	411 seconds
Calls requiring identity verification	73%
Authentication handled by an agent	84%
Mean agent authentication time	42 seconds
Calls affected by repetition (modeled)	30%
Time added per repetition event	15 seconds
Agent time on post-call wrap-up	13.7%
Self-service containment where offered (mean)	33%
Self-service containment where offered (median)	15%

These quantities can be combined to size each intervention on the same basis. The table below models the annual saving per 1,000 inbound calls from four changes, using a cost per second of agent-handled call of \$0.0175, derived from \$7.20 over a mean duration of 411 seconds.

Please note that these figures are for example only, make a number of assumptions and are presented to give an idea of the possible scale of savings achievable.

Figure 47: Modelled saving per 1,000 inbound calls, by intervention (as examples only)

Intervention	Calls affected	Saving each	Per 1,000 calls
Increase self-service containment, 33% to 50%	170	\$7.20	\$1,224
Automate 70% of post-call work	all agent-handled	\$0.89 (70% of \$1.27)	up to \$890
Passive authentication replaces agent ID&V	613	\$0.74	up to \$454
Remove repetition caused by poor audio	300	\$0.25	up to \$75

Containment is the largest prize by a clear margin. A contained call avoids the whole \$7.20 rather than a fraction of it, so even a modest improvement in containment outweighs a substantial improvement in anything that merely shortens a call the agent still handles.

Moving containment from 33% to 50% is worth roughly two and a half times what removing agent authentication entirely is worth, and sixteen times what eliminating repetition is worth.

Two qualifications apply: automation is not free, so the \$7.20 is a gross figure from which the per-interaction cost of the voicebot must be deducted along with all implementation, integration and data cleaning costs., and a “contained” call which produces a repeat contact has not been contained in any sense that matters, so the saving should be measured net of repeat contacts within a defined window.

There is also a serious qualification about certainty. The \$454 from authentication is more likely: the task is mechanical and the customer has no view on who performs it.

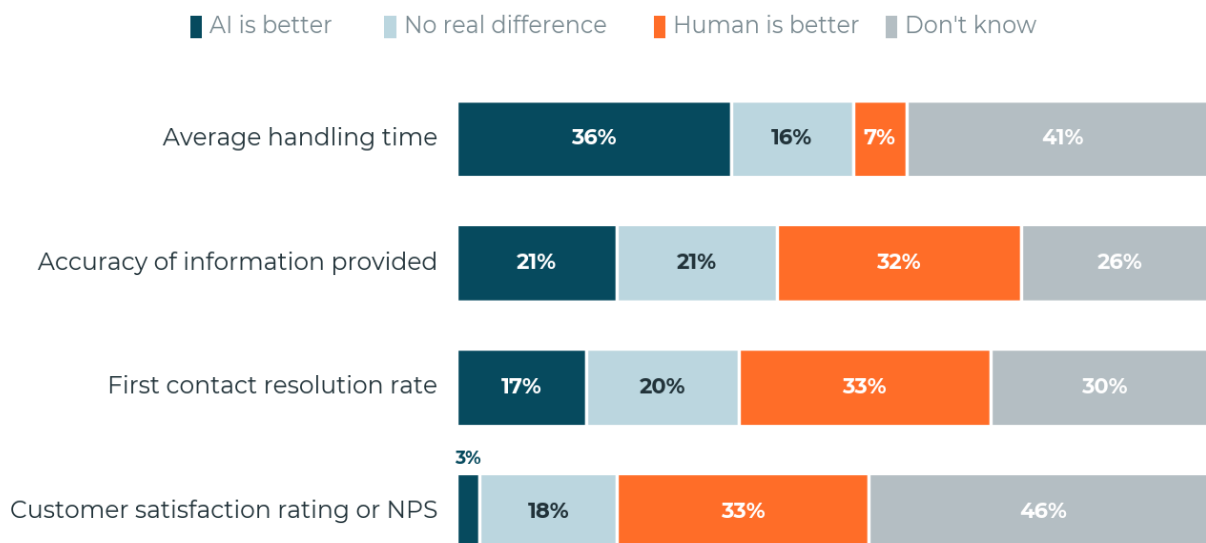
The \$1,224 from raising containment is less certain: the median US operation currently contains 15% of voice calls (against a mean of 33%), 34% of consumers say they would never prefer AI under any circumstances, and the leading reason for abandonment is trust rather than anything a better voice interface necessarily fixes: organizations have to earn the customers’ trust with their new voice AI systems.

PERFORMANCE AGAINST EXPECTATIONS

Looking at the outcomes of AI, average handling time is the one measure on which AI is judged clearly superior. 36% say AI is better, 7% say the human is better and 16% see no real difference.

On every other measure the balance runs the other way: 32% say the human is better for accuracy of information against 21% for AI, 33% against 17% for first contact resolution, and 33% against 3% for customer satisfaction rating or NPS.

Figure 48: Differences in performance between AI-handled and human-handled interactions



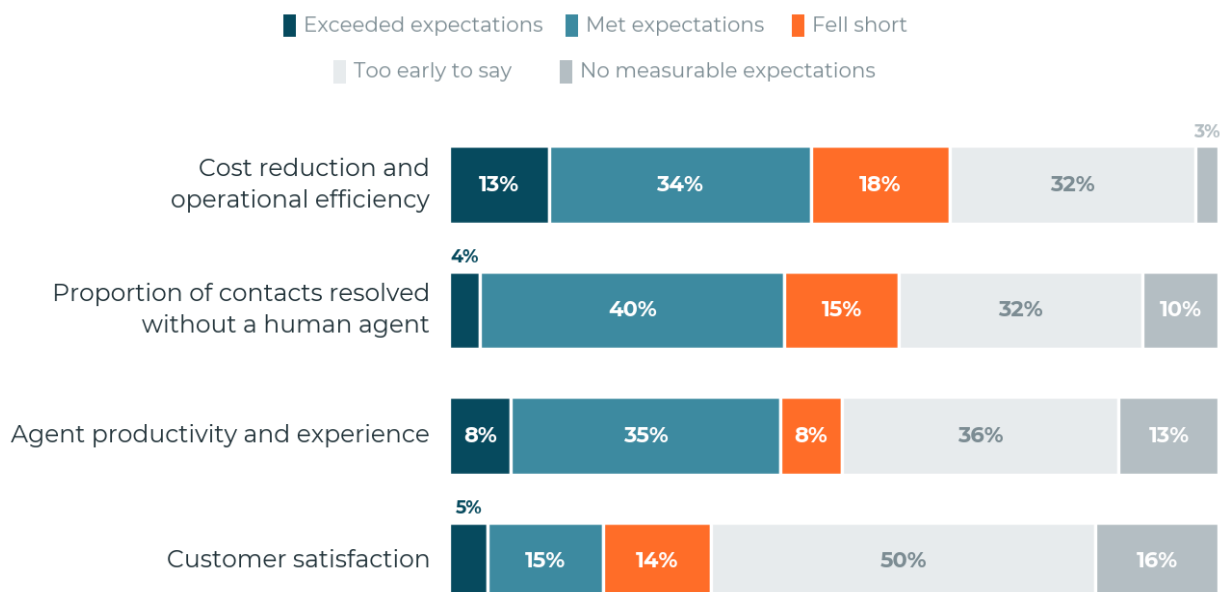
That is the same shape as the consumer data seen earlier, where customers rate AI worse than a human on five of six dimensions and better only on speed. Businesses and their customers are describing the same thing from opposite sides of the call, and the agreement between two independent samples is a useful corroboration.

The don't-know column is also a key finding: it runs from 26% to 46%, so close to half the market cannot say whether its AI produces better or worse satisfaction than its people for comparable contacts.

When considering performance against expectations, cost reduction and operational efficiency is where AI most often meets what was promised: 47% report it met or exceeded expectations, against 18% falling short.

Contacts resolved without a human is close behind at 44%, and agent productivity at 43%. Customer satisfaction is the weakest, at 20% met or exceeded and 14% falling short.

Figure 49: Performance against the expectations set in the original business case



“Too early to say” accounts for between 32% and 50% of responses, highest on customer satisfaction, which is consistent with a market in which most deployments are recent. Between 3% and 16% set no measurable expectations at all.

What the interim results do establish is the direction of spending and of expectation.

4% of respondents expect their contact center AI budget to fall over the next two years. 34% expect an increase of up to 25%, 22% an increase of 26% to 50%, and 18% an increase of more than 50%. 12% expect it to stay roughly the same and 9% don’t know.

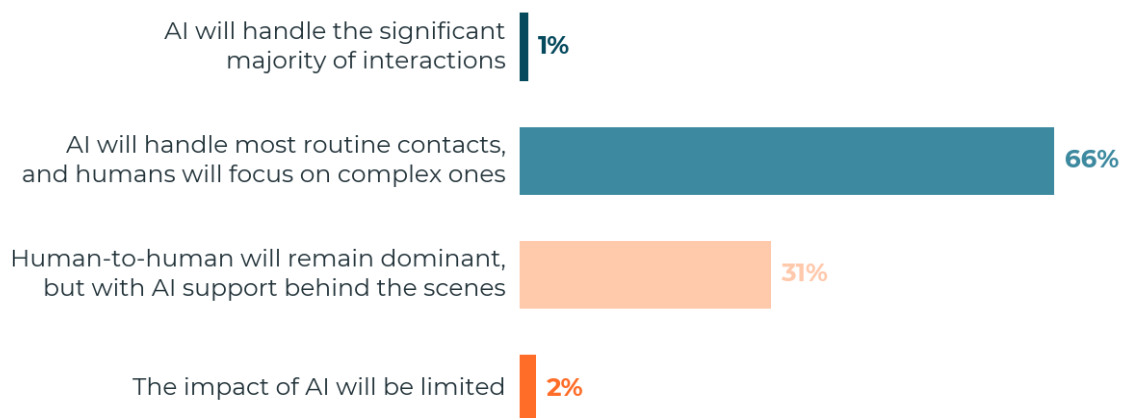
A market in which 74% of buyers expect to spend more and very few expect to spend less is not one that is reconsidering the technology, whatever the rollback figures in the previous section suggest about individual projects.

Our survey respondents were asked to consider what the future held for AI in contact centers.

Only 1% expect AI to handle the significant majority of interactions including complex ones, with 66% believing that AI will handle most routine contacts while humans focus on complex ones, and 31% that human-to-human interaction will remain dominant with AI supporting behind the scenes.

Only 2% expect the impact of AI to be limited on their operations.

Figure 50: Which statement is closest to the expected position in five years



The industry's own expectation is therefore a division of labor rather than a replacement, and the tails on both sides are negligible.

Almost nobody working in US contact centers believes AI will take over, and almost nobody believes it will amount to little.

The 66% and the 31% are separated less by disagreement about capability than by where each organization currently sits, and both are consistent with the evidence in this report: automation works on the contacts customers are content to have automated, and the residue is harder than what came before.

THE RIGHT TARGET, MEASURED THE WRONG WAY

Containment being the largest potential prize – in terms of cost savings – does not necessarily make containment rate a good metric to focus upon.

As a target it is well chosen, but as a number on a proposal it is unreliable, for two reasons this report has evidenced.

It is not defined consistently: one supplier may count every call that did not reach an agent, including those the customer abandoned, while another counts only calls where the customer's stated goal was achieved.

Containment may be measured as the proportion of contact sessions initiated that did not reach an agent, because that is the figure the platform can produce without integration: while it knows whether a call was transferred to an agent, but it does not always know whether the customer got what they wanted in the self-service session. A caller who gives up in the IVR and rings off is not transferred, so on that definition they can be counted as "contained".

That is why the wording of the metric matters more than it looks. "Not transferred to an agent" and "customer's stated goal achieved" are different quantities, and only the second is worth paying for. Worse, a contained call that produces a repeat contact two days later has cost more than a call handled once by an agent.

A better headline metric is cost per resolved contact, measured across the whole journey including any repeat contact within a defined window, reported separately for AI-handled and human-handled contacts of comparable type. Defined that way, containment remains a number worth maximizing.

A SEQUENCE THAT THE EVIDENCE SUPPORTS

Containment being the largest prize does not settle the order in which work should be done, and the evidence in this report points to a set of dependencies rather than a fixed sequence.

What follows is offered as considerations rather than as steps, because the right order depends on where an organization already is.

Audio quality and transcription accuracy are the foundation the rest of the stack is built on, and everything above them inherits their quality.

The direct saving from removing repetition is the smallest in the table, at \$75 per 1,000 calls, but that figure massively understates its importance because speech-to-text accuracy is what every layer above it consumes.

A transcript that is wrong produces a bot that misunderstands, an analytics platform that miscounts and an automated QA score that cannot be relied on. Superior transcription is the cheapest way to improve several things at once.

Post-call summarization is the second largest prize at up to \$890 per 1,000 calls and one of the most certain in the table. It applies to almost every call, and it fails safely, because a poor summary is corrected by the agent rather than experienced by the customer.

Authentication is worth up to \$454 per 1,000 calls in its own right and has a second property worth noting: a bot cannot resolve an account-specific enquiry without establishing who is calling, and 73% of US calls require that.

Analytics and automated scoring address the 73% of operations reporting insufficient time to analyze their data, and they are also how an organization finds out which intents its bot is failing on. The constraint identified earlier applies: automated analysis without additional coaching capacity produces findings nobody can act on.

Real-time assistance is the point at which AI output first reaches the customer, mediated by an agent who can override it, which makes it a useful place to find out how a model behaves on live calls before it is trusted unsupervised.

Self-service resolution is likely to be the largest prize for many organizations, and the one most affected by everything above. The evidence suggests targeting the enquiries customers say they would prefer to automate, being quick answers to simple questions, queue avoidance and out-of-hours contact, and being more cautious with complex or emotionally loaded contacts where 61% of consumers rate AI worse than a human at understanding the issue.

To summarize: containment through self-service is likely to be the largest prize that voice AI promises, but the quality of the transcript, the reliability of identity verification and the visibility provided by analytics all determine how much of that prize is reachable.

None of this argues for holding back self-service until everything else is finished, but it cannot be stated clearly enough that the quality of the transcript, the reliability of identity verification and the visibility provided by analytics all change what a self-service deployment can achieve, and that an organization which has attended to them fully will get more from the same investment than one which has not.

ABOUT CONTACTBABEL

ContactBabel is the contact center industry expert. If you have a question about how the industry works, or where it's heading, the chances are we have the answer.

We help contact centers compare themselves to their closest competitors so they can understand what they are doing well, what needs to improve and how they can do this.

The coverage provided by our massive and ongoing primary research projects is matched by our experience analyzing the contact center industry. We understand how technology, people and process best fit together and how they will work collectively in the future.

e: info@contactbabel.com | w: www.contactbabel.com | t: +44 (0)1434 682244

Free research reports available from www.contactbabel.com (UK and US versions):

- **The Inner Circle Guide to:**
 - Agent Engagement & Empowerment
 - Agentic AI
 - AI-Enabled Agent Assistance
 - Chatbots & Voicebots
 - Cloud-based Contact Center Solutions
 - Customer Engagement & Personalization
 - Customer Interaction Analytics
 - First-Contact Resolution
 - Fraud Reduction & PCI Compliance
 - Omnichannel Workforce Optimization
 - Remote & Hybrid Working Contact Center Solutions
 - Self-Service
 - South African Contact Centres
 - Voice AI
- **The UK Contact Centre Decision-Makers' Guide**
- **The UK Customer Experience Decision-Makers' Guide**
- **UK Contact Centre Verticals:** Communications; Finance; Insurance; Outsourcing; Retail & Distribution; Travel; Utilities
- **AI in UK Contact Centre Verticals:** Finance; Insurance; Retail & Distribution; Utilities
- **The US Contact Centre Decision-Makers' Guide**
- **The US Customer Experience Decision-Makers' Guide**
- **US Contact Centre Verticals:** Finance; Insurance; Outsourcing; Public Sector; Retail & Distribution
- **AI in US Contact Centre Verticals:** Finance; Insurance; Retail & Distribution
- **The AI Series:** how can AI can help contact centers' operational and commercial issues?
Research reports: First-Contact Resolution; Sales Growth; Workforce Engagement; Business Insights: Customer Insights; Agent Productivity; Digital Customer Contact; Contact Centre Cost Reduction; Customer Satisfaction.