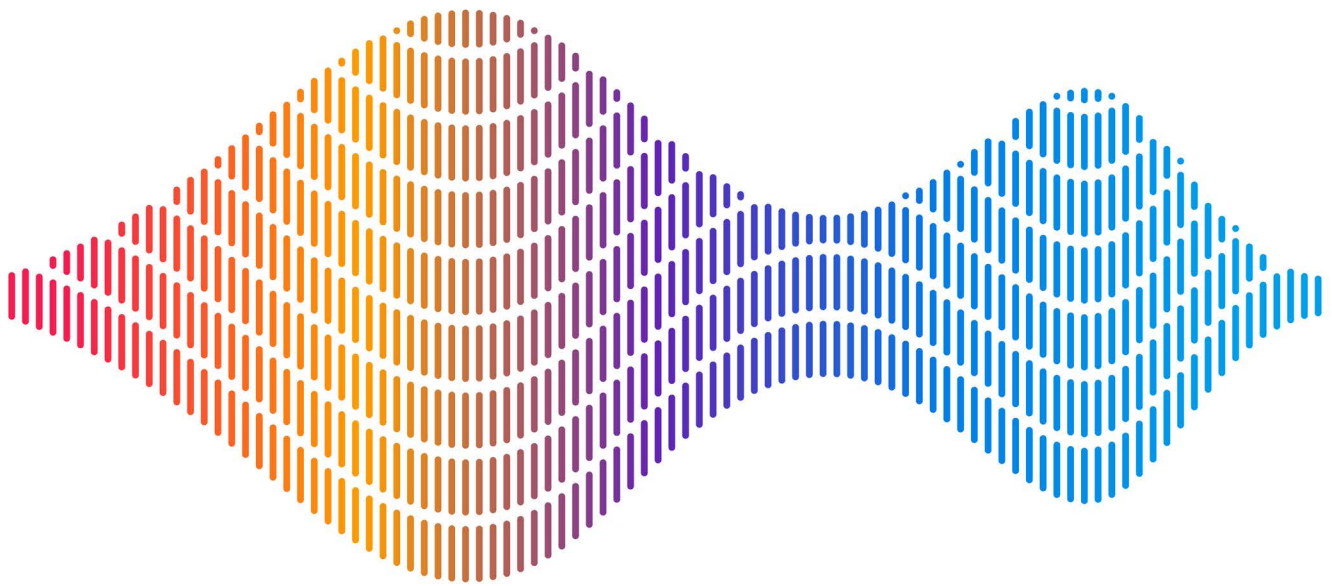




The Inner Circle Guide to Voice AI

Research supported by



Transforming IVR into Intelligent Conversation

Your first step toward AI infused customer experience.

What Geomant VoiceShift is:

VoiceShift is a structured discovery engagement that analyses your current IVR, identifies automation opportunities, and defines a clear roadmap to intelligent voicebots that deliver faster, more natural customer experiences.

What's Included:

IVR Performance & Experience Audit

Review of call flows, routing logic, customer effort and failure points.

AI & Automation Opportunity Map

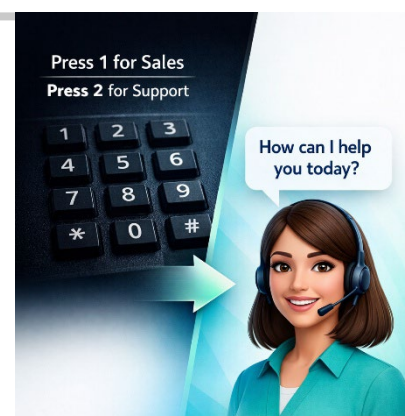
Identification of high-value use cases for voicebots and self-service.

Conversational Journey Redesign

Reimagined flows using natural language and intent recognition.

Migration Roadmap & Business Case

Phased plan, ROI modelling, and technology recommendations.



Live Voicebot Prototype (Included)

Geomant will build and deploy a **working AI voicebot** that emulates all or part of your current IVR using real intents and flows.

This gives your stakeholders a **hands-on, audible demonstration** of how AI transforms the customer experience:

- Hear the difference between menus and natural conversation
- Experience real-time intent recognition
- Test routing, FAQs, and self-service flows
- Validate feasibility before committing to full deployment

This is the “wow moment” that accelerates buy-in.

 VoiceShift

What you walk away with:

- Evidence based modernisation plan
- Prioritised automation roadmap
- Conversational flow concepts
- Integration recommendations
- Quantified CX + operational impact
- A working voicebot prototype you can demo internally

Why Geomant?

- 25 years of Contact Centre expertise
- Deep AI & integration capability
- Vendor-agnostic, business-first approach
- Proven IVR modernisation methodology

Ready to evolve your voice channel?

Geomant VoiceShift gives you clarity, confidence, and a practical plan to move from legacy IVR to next generation conversational automation — with a real voicebot you can test from day one.

Book your VoiceShift Discovery Session - **email us at hello@geomant.com**

The Inner Circle Guide to Voice AI (UK 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	14
Voice AI is not only self-service.....	16
The four layers of a voice AI deployment.....	18
Getting the Audio Right.....	20
Two trends working against clear audio.....	20
What poor audio costs.....	22
What can be done about it.....	23
Accent, language and comprehension	24
Voice Self-Service: From Containment to Resolution	27
How much self-service actually handles.....	27
Why sessions are abandoned	28
What voicebots can currently do.....	29
How much gets resolved without a human.....	30
When customers prefer to deal with AI	31
How customers rate AI against human agents	32
The concern that shapes everything else	33
When self-service does not succeed	35
What this means for containment targets.....	36
Voice AI Alongside the Agent.....	37

What agents actually spend their time on.....	37
The onboarding problem.....	39
What real-time assistance is measured to deliver	40
What agents make of it.....	41
What AI has done to headcount	42
Analytics, Quality and Assurance	43
How widely analytics is deployed	43
What analytics is used for	45
Why the quality process needs the help	46
Assuring the AI itself.....	48
Voice, Identity and Fraud	50
The scale of the authentication problem	50
Fraud, and the problem with using a voice as a credential	52
Proactive Outbound and the UK Rules	54
The commercial case.....	54
A key distinction to understand	55
What the rules require.....	55
The enforcement precedent.....	57
Disclosure.....	57
A practical checklist.....	58
Implementation: What Breaks.....	59
Why voice is harder than text	59
What is holding contact centres back generally.....	60
Sourcing, and what got rolled back.....	61
What got stopped.....	63
What is blocking further adoption	64
Measuring Return, and What to Do Next.....	65
The building blocks of a business case	65
Performance against expectations	67



The right target, measured the wrong way.....70

A sequence that the evidence supports 70

About ContactBabel..... 72

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), 2004 - 2025	13
Figure 4: The five levels of voice self-service technology	14
Figure 5: Current UK adoption of voice self-service technologies.....	15
Figure 6: Current and planned use of voice-related AI applications	16
Figure 7: The four layers of a voice AI deployment.....	18
Figure 8: What % of your customers use a mobile phone to call a contact centre? 2016-25	20
Figure 9: The rise of remote & hybrid contact centre working (2010-2026)	21
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)	22
Figure 11: The unnecessary cost of mishearing	23
Figure 12: Current and planned use of real-time translation or multilingual support.....	24
Figure 13: Current and planned use of accent localisation or accent softening	25
Figure 14: Overall proportion of calls handled entirely through self-service (only in respondents which offer telephony self-service)	27
Figure 15: Reasons for abandoning telephony self-service sessions.....	28
Figure 16: Capability of the most capable chatbot or voicebot in the organisation	29
Figure 17: Controls applied to the actions bots can take in business systems.....	30
Figure 18: When contacting an organisation, in which situations would you prefer to deal with AI rather than a human?.....	31
Figure 19: How does AI compare to human agents in your experience?	32
Figure 20: To what extent are you concerned about each of the following when dealing with AI in customer service?	33
Figure 21: Concerns about difficulty reaching a human agent when dealing with AI in customer service (by age range)	34
Figure 22: Quality of escalation from bot to human.....	35
Figure 23: Agent activity	37
Figure 24: Proportion of inbound call spent navigating between screens / applications....	38

Figure 25: Requirements for in-call and post-call AI automation.....	38
Figure 26: "New agents find it difficult to familiarise themselves with our systems", by contact centre size	39
Figure 27: Impact of real-time AI agent assistance on selected metrics.....	40
Figure 28: Agreement with statements about how agents feel about AI	41
Figure 29: Effect of AI on the number of people employed in the contact centre	42
Figure 30: Use of interaction analytics, by contact centre size	43
Figure 31: Use of various interaction analytics functionality (from only those respondents who use analytics).....	44
Figure 32: Usefulness of post-call analytics.....	45
Figure 33: Usefulness of real-time analytics.....	46
Figure 34: Challenges to managing performance & quality	46
Figure 35: Effectiveness of quality assurance	47
Figure 36: How organisations quality assure interactions handled by AI	48
Figure 37: Extent to which AI-driven analytics is generating actionable operational insight	49
Figure 38: Proportion of calls requiring caller identification & average time taken, 2010-25	50
Figure 39: Caller identity authentication methods (only those contact centres which authenticate some or all calls)	51
Figure 40: Time taken to authenticate caller identity using only an agent (seconds)	51
Figure 41: Concerns about security.....	52
Figure 42: Current and planned use of voice biometrics or AI-based authentication and fraud detection	53
Figure 43: Current and planned use of proactive and predictive outbound AI	54
Figure 44: What is preventing the contact centre from achieving its aims?	60
Figure 45: Where the majority of AI capability in the contact centre comes from	61
Figure 46: What matters most when deciding where to source AI (top 3 answers)	62
Figure 47: Main reasons for stopping, pausing or rolling back an AI solution	63
Figure 48: Most significant barriers to further AI adoption over the next two to three years	64
Figure 49: Modelled saving per 1,000 inbound calls, by intervention (as examples only)	66

Figure 50: Differences in performance between AI-handled and human-handled interactions..... 67

Figure 51: Performance against the expectations set in the original business case.....68

Figure 52: Which statement is closest to the expected position in five years.....69



Geomant is an AI systems integrator with 25 years of Contact Centre and CX experience, always focused on improving customer interactions. We design, integrate and operate customer interaction solutions that make enterprises more accessible to their customers — across contact centres, self-service, AI assistants, voice and chat bots, queue management and hybrid human-plus-AI service models.

As a systems integrator, Geomant is focused on delivering real improvements to customer experience through the practical use of AI - with a measurable ROI. We combine consulting depth, engineering muscle and operational discipline to deploy AI responsibly and measure it continuously — from conversation intelligence and agent assist, to fully autonomous voice agents, QA automation, real-time translation, chatbots and post-call summarisation.

We have our own AI and CCaaS software, but unlike pure software vendors, Geomant delivers end-to-end transformation across hybrid, cloud and legacy environments. Our services span implementation, custom development, support and maintenance and CX consulting. We support major platforms end-to-end, operate 24/7 managed services, and remain vendor-agnostic by design — so customers get the right solution, not just the vendor's product.

Trusted by leading enterprises across banking, telecoms, logistics, insurance and retail — including BNP Paribas, DPD, Vodafone, Qatar Airways, Generali and Unicredit — Geomant turns complex interaction ecosystems into practical, measurable and reliable outcomes.

If your organisation is looking for a Systems Integrator that can guide you through the practicalities of introducing AI – talk to Geomant. We'll be happy to show you just what can be achieved by incorporating AI into your customer interactions.

Find out more at www.geomant.com/#ai. Contact us via our website, or email us at hello@geomant.com.

Contact:

w: www.geomant.com/#ai

e: hello@geomant.com

INTRODUCTION AND METHODOLOGY

This report examines the use of artificial intelligence in the voice channel of UK contact centres: 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 UK 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 UK contact centre operations. Consumer findings are drawn from surveying 1,000 UK consumers.

Where a figure has been modelled 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 centre 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

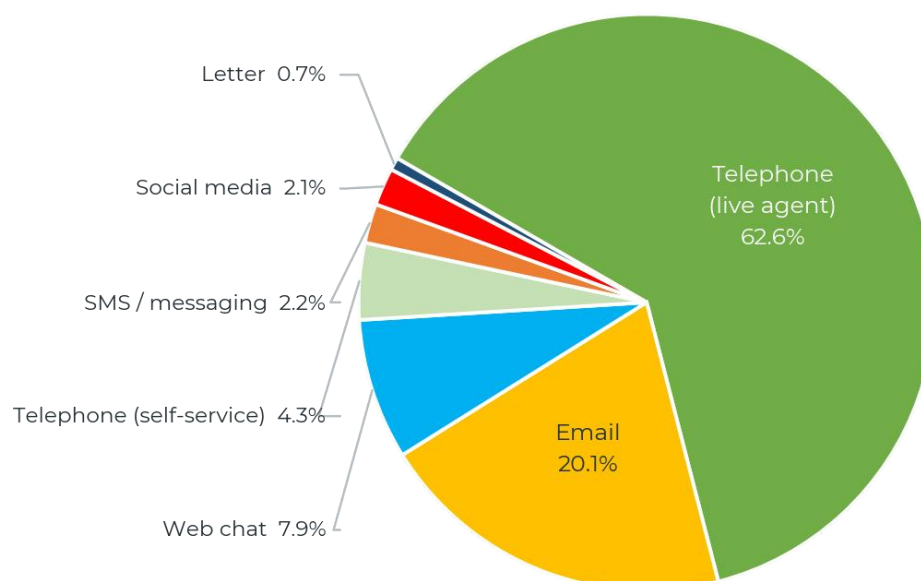
Voice AI is among the most heavily marketed categories in contact centre technology, and among the least precisely defined, which makes it worth establishing what is actually being sold before considering what it can do.

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 62.6% of inbound interactions. The long-term decline has been gentle rather than steep, and in recent years it appears to have stabilised altogether, despite the industry continuing to expect that phone volumes will fall. Email remains the largest digital channel at 20.1% and web chat has grown to 7.9%, while social media has settled back to around 2% after a brief period at a higher level.

More significant for this report is telephony self-service, which accounts for only 4.3% of inbound activity and has fallen since 2022, having previously risen from 4.9% in 2019 to 6.2% in both 2020 and 2021 in line with the growth of self-service elsewhere. This is a central fact of the voice AI market: automation has advanced across every other channel while the automated share of the voice channel has at the very best, remained stagnant.

THE COST OF A VOICE MINUTE

Voice is the most expensive channel a contact centre operates, and the gap has widened.

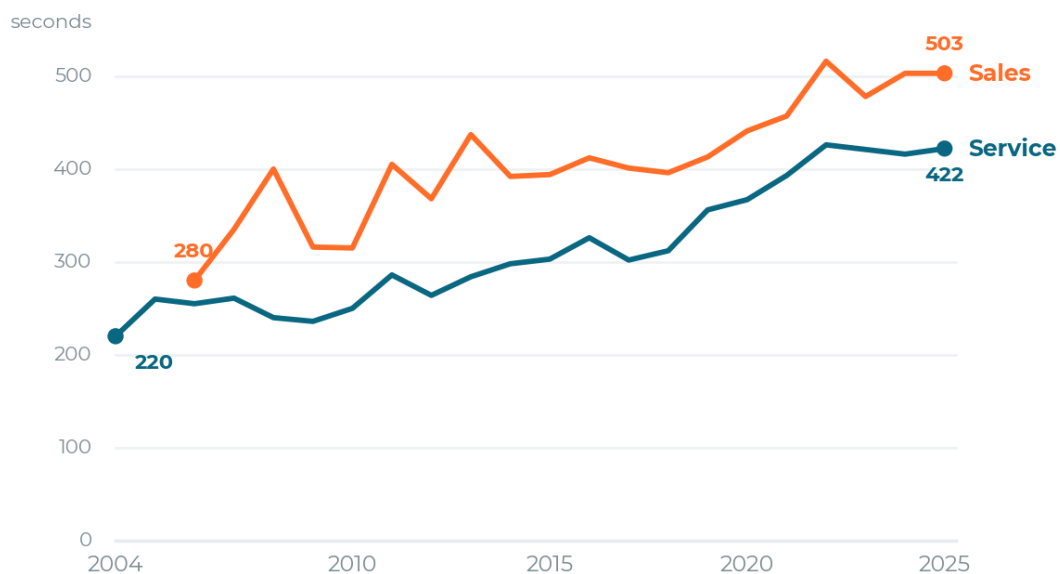
Figure 2: Cost per inbound interaction

Channel	Mean	1st quartile	Median	3rd quartile
Phone	£6.17	£7.00	£5.00	£4.00
Email	£4.92	£6.13	£5.00	£2.73
Web chat	£3.38	£5.00	£3.15	£1.11
Social media	£2.99	£3.75	£2.60	£2.00
IVR / speech recognition	c. 30-50p	-	-	-

After many years in which digital channels cost much the same as voice to handle, digital has become considerably cheaper, 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), 2004 - 2025



Mean service call duration has risen from 220 seconds in 2004 to 422 seconds in 2025, while sales calls have gone from 280 seconds in 2006 to 503 seconds, meaning that calls are now roughly twice as long as they were twenty years ago.

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.

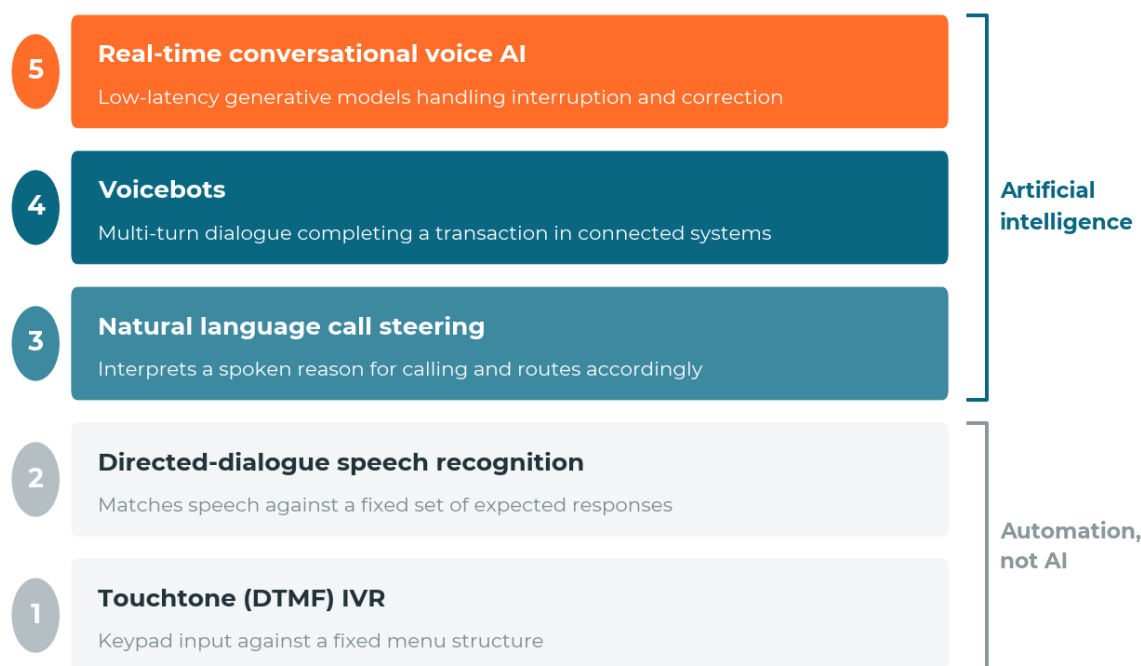
Queue times have moved in the same direction. Average speed to answer stood at 16 seconds in 2004 and 40 seconds in 2019 before rising to 82 seconds in 2020 and peaking at 120 seconds in 2022, and although it has eased slightly since, it remains at 111 seconds in 2025 with call abandonment at 6.8%.

Taken together, these figures describe a channel that is expensive, growing more expensive per contact, and under greater strain than at any point in the time we have been measuring this.

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 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.

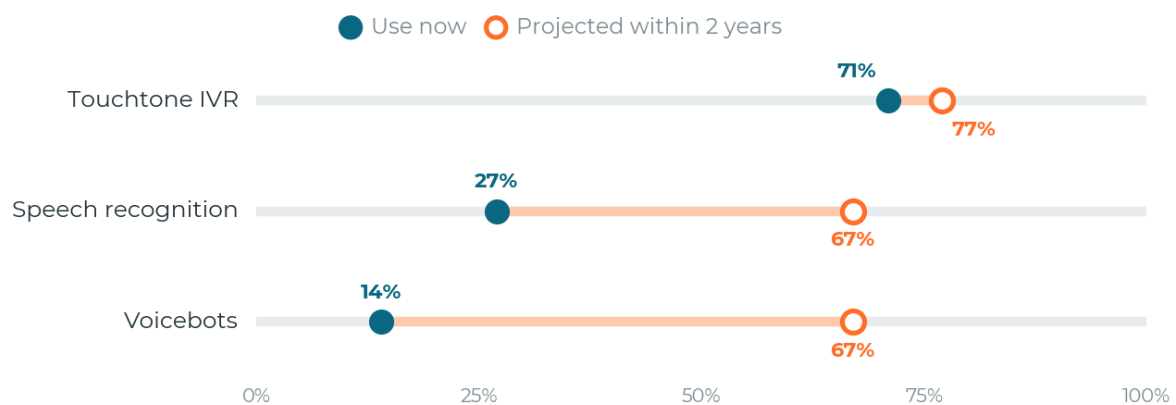
From the third level upwards the position changes, because a model is interpreting meaning, rather than matching a pattern.

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 UK adoption of voice self-service technologies



Touchtone IVR is used by 71% of UK contact centres, rising to 89% of large operations, and a further 6% 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 27% and by 37% of large operations, with a further 40% planning to implement, while voicebots are used by 14% against 67% projected within two years.

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.

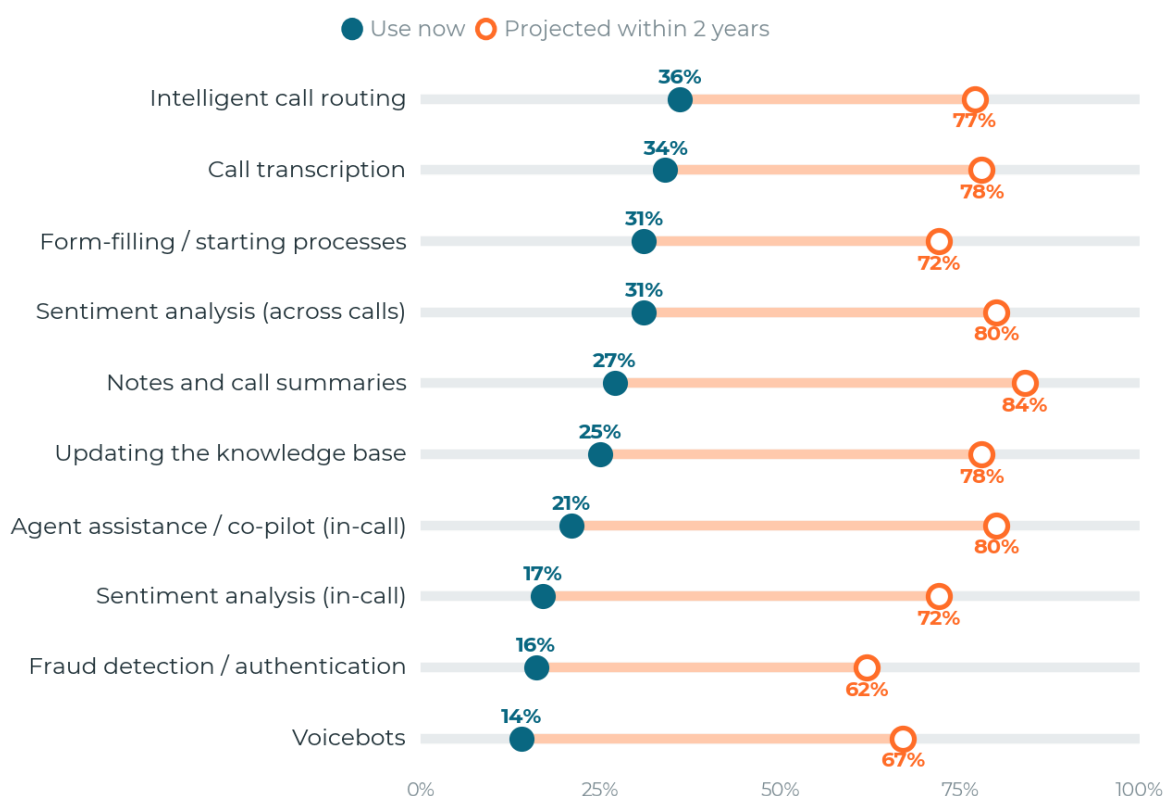
Everything discussed so far concerns voice AI used for self-service, which is to say systems that replace an agent for the whole interaction. While this is the most visible application and the one that attracts most of the marketing attention, it is a long way from being the only application of voice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

Voicebots themselves are used by 14%, with 67% projected within two years, which is one of the steepest expected rises of any application shown.

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 34% with a further 44% planning to adopt. It is the foundation for analytics, summarisation and quality assurance alike, which is why it is one of the most widely deployed voice AI applications.
- **Agent assistance** supplies the agent with suggested responses, retrieved knowledge and next best action while the conversation is live, and is in place at 21% with 59% planning to implement. Sentiment analysis during the call, at 17%, is linked to agent assistance.
- **Summarisation and wrap-up** generate call notes, disposition codes and CRM updates once the conversation ends, used by 27% and planned by a further 57%, which is the highest projected adoption of any application measured.
- **Authentication and fraud detection** use the voice itself as a credential and screen for synthetic or fraudulent callers, currently at 16% and planned by a further 46%.

Analytics belongs alongside these, though it operates on the aggregate rather than on the individual call. Sentiment analysis across multiple conversations is used by 31%, and interaction analytics more broadly by 58%, making the retrospective examination of voice one of the most established applications of AI in the contact centre 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 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.

From inbox to **voice**: phasing AI into customer service at a major European courier

How a leading European parcel and logistics provider absorbed rising contact volumes without adding headcount - by treating AI as a journey, not a switch.

227k+

AI-handled email enquiries within 11 months

16k+

Calls handled by the voicebot in first two months

54%

Automation rate, and increasing

8 FTE

Peak-month workload saved in a 25-agent team

THE CHALLENGE

The customer's national contact centre handled enquiries by phone and email, with calls routed through a traditional IVR to 25 agents. Parcel tracking, delivery changes, returns and complaints are repetitive, high-volume and spiky - and volumes were growing faster than the business wanted to grow the team.

The brief was explicit: absorb the growth with AI, keep the existing Avaya telephony and back-office estate, improve accuracy rather than merely deflect, and do it in local languages - not just English.

A PHASED JOURNEY, DESIGNED WITH THE CUSTOMER, DELIVERING REAL ROI

- 1 **Email first - fastest proof of value**
AI drafted answers; initially agents reviewed and sent. Handling time fell from 4-5 minutes to 2, with 92% of drafts approved unchanged. Confidence, not just savings.
- 2 **Automation without the human in the loop**
Once accuracy was proven, answers went out automatically. Bot efficiency rose from 35% to 54%; only genuinely complex cases reach an agent queue or trigger an automatic expert ticket.
- 3 **Voice added - deliberately simple at launch**
A focused "where is my parcel?" voice bot went live in June, identifying the caller from their number, retrieving live parcel status and speaking it back - averaging 355 calls per working day.
- 4 **Voice scale-up**
Next releases add spoken parcel-number capture, delivery date and address changes, returns, complaints and depot/parcel-point location - the same logic the email bot already uses.

"The numbers tell part of the story, but the bot's continued improvement in efficiency is also a testament to the exceptional professionalism, dedication and ongoing support of Geomant's consultants and developers"

CUSTOMER

WHY VOICE NEEDS MORE THAN A KNOWLEDGE BASE

- **Simple answers** - policy and "how do I..." questions are served from a curated knowledge base.
- **Real answers** - real-time information - parcel status, depot data and delivery windows require live interrogation of core business systems.
- **Real actions** - changing a delivery date or address, booking a return or raising a claim means initiating transactions, not just reading data. This systems-integration layer is where Geomant adds significant value.
- **One brain, many channels** - email and voice share the same email and voice share the same core system and integrations, with specific use-case logic ensuring consistent answers.
- **Safety by design** - a dedicated phishing use case recognises attempts to extract customer data and refuses them.

THE SOLUTION

Built on Geomant's **Graia AI contact centre platform**, wrapped around the customer's existing **Avaya IP Office** estate and middleware connecting to backend systems. Best-in-class AI services - Azure OpenAI for reasoning, Speechmatics for local language speech-to-text, Azure for speech synthesis - are combined with **custom Geomant middleware** integrating parcel, ticketing and email systems.

"Phasing was as important as the technology choice. Start where the benefit lands soonest, prove accuracy, build trust, then take the customer into voice with the evidence in hand."

GEOMANT

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.

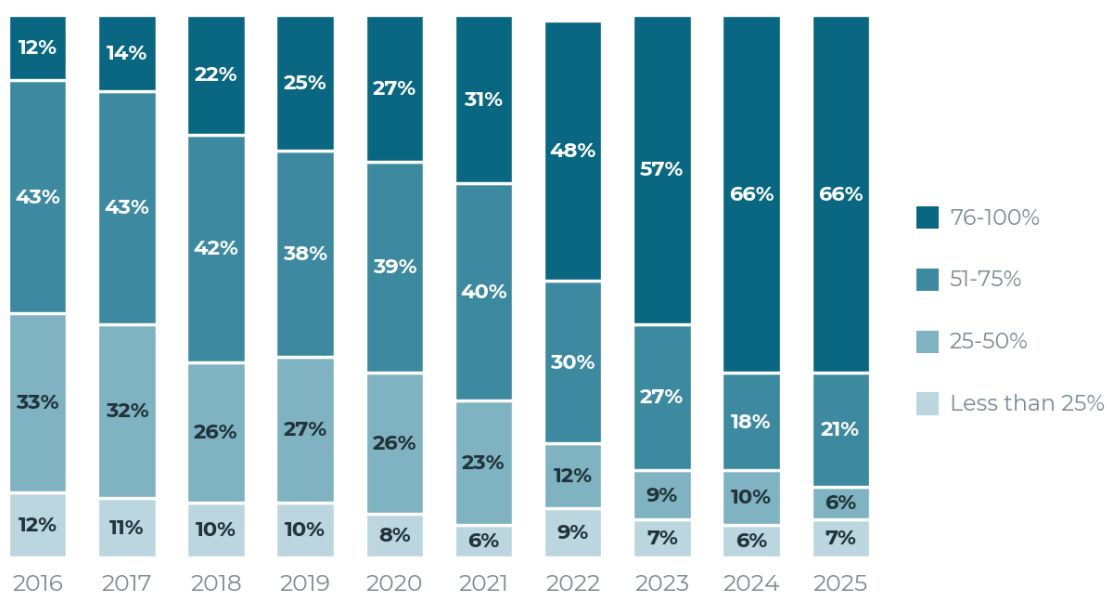
It determines what every other layer has to work with, and this section looks at why audio quality on a contact centre call has been getting harder to control, what it costs when it goes wrong, and what can now be done about it.

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 mobile. On average, 74% of customers now use a mobile phone rather than a landline to call the contact centre, and 87% of survey respondents state that more than half of their customers do so. The switch-off of the PSTN is accelerating this..

Figure 8: What % of your customers use a mobile phone to call a contact centre? 2016-25



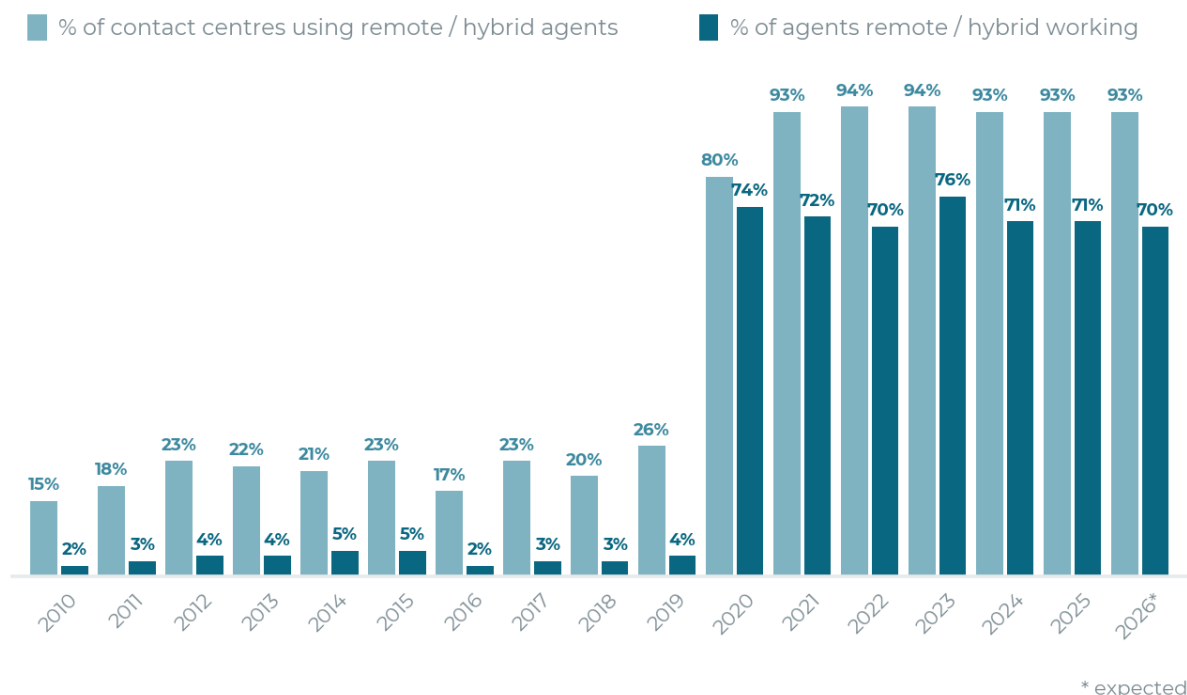
The shift has been rapid: in 2016, only 12% of contact centres reported that more than three-quarters of their customers called from a mobile, and that figure now stands at 66%, having more than doubled since 2021.

At the other end of the range, the proportion of operations where fewer than a quarter of customers use mobile has fallen from 12% to 7%.

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 centre's own side of the line.

Figure 9: The rise of remote & hybrid contact centre working (2010-2026)



Before 2020, remote working was a marginal practice: fewer than 4% of UK agents worked from home in any year between 2010 and 2019, even though around a fifth of operations had some remote capability.

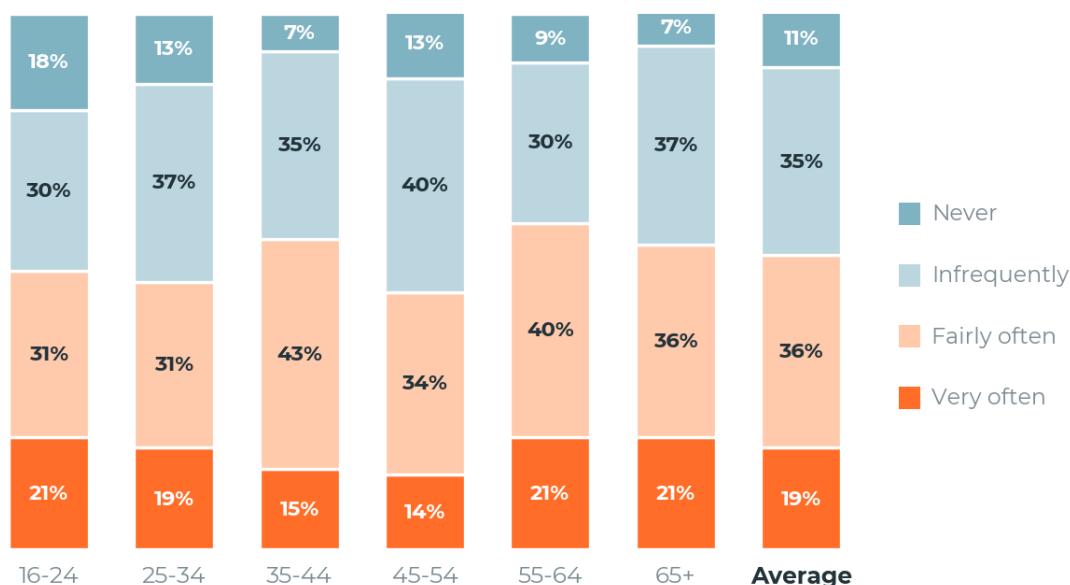
The pandemic moved the industry to 74% of agents working remotely or in a hybrid pattern within a single year, and the position has held ever since, with 71% in 2025 and 70% expected in 2026 across 93% of contact centres. There is no expectation within the industry that a substantial return to a fully centralised 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 centres 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 centre 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 1,000 UK consumers, 19% reported that they very often had problems hearing the agent or were asked to repeat themselves, and a further 36% said it happened fairly often. Only 11% said it never happened at all.

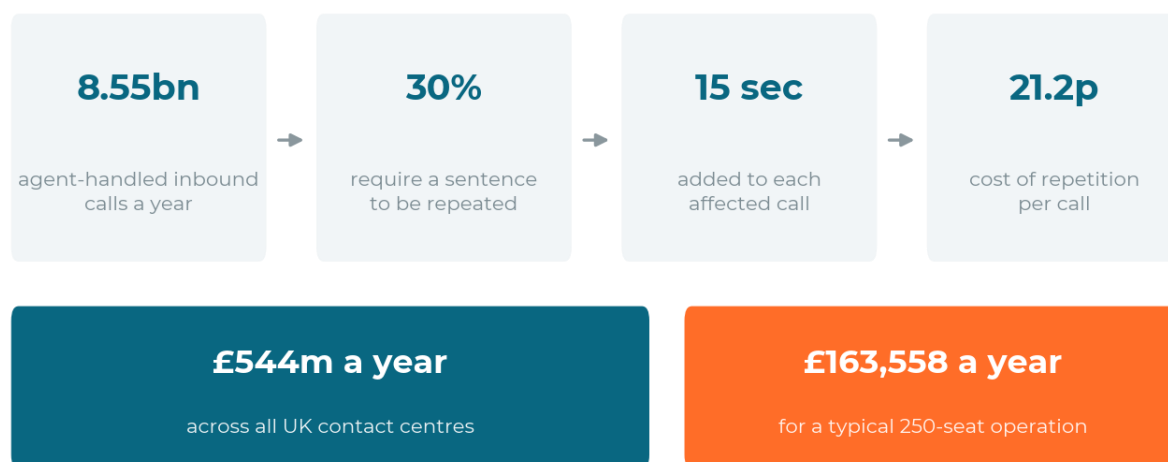
The age breakdown is the interesting part, because it does not follow the pattern most people would predict.

Older customers do report the problem regularly, with 21% of both the 55-64 and 65+ groups saying it happens very often, and the RNID reports that more than half of people aged 55 and almost 80% of those aged 70 and over have some hearing loss, a prevalence that will rise as the population ages.

Yet the youngest cohort reports it just as much: 21% of 16-24s say very often, and 52% say either very often or fairly often. Whatever is happening here, it is not simply a hearing-loss effect.

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

Figure 11: The unnecessary cost of mishearing



The calculation runs as follows: UK contact centres handle an estimated 8.55bn agent-handled inbound calls a year. If 30% of those calls require a sentence to be repeated, that is 2.57bn calls; if each repetition adds 15 seconds to a mean call duration of 437 seconds, repetition accounts for 3.4% of the call; and at a mean cost per inbound call of £6.17, that is 21.2p per affected call and £544m 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 £163,558.

Some operations will sit well above these levels. Those with a high proportion of remote agents, those whose customers are predominantly on mobiles, and those serving an older customer base will all 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 centre 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 summarisation 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 recognised 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 localisation 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.

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

Response	% of respondents
Use now, no change planned	16%
Use now, looking to replace	2%
Plan to implement within 12 months	13%
Plan to implement after 12 months	14%
No plans	50%
Don't know	6%

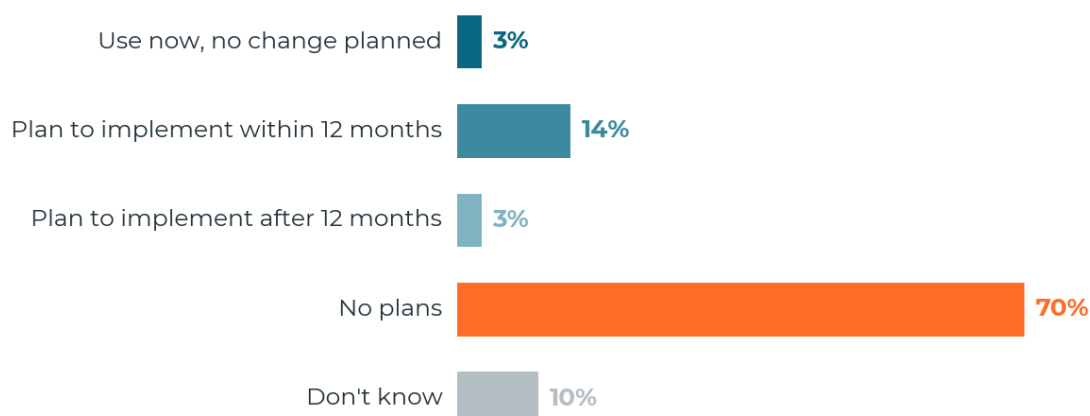
18% of UK contact centres use real-time translation or multilingual support, and a further 27% plan to implement it. Half of the market, at 50%, has no plans at all, which is the second highest no-plans figure of any application measured in the survey.

Where translation is deployed it tends to be settled rather than experimental, with 16% using it now with no change planned against only 2% looking to replace what they have. That is the profile of a capability that works adequately for the organisations 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 organisations 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 the least adopted voice AI application measured anywhere in this report. 3% use it now, 14% plan to implement within twelve months and 3% after that, while 70% have no plans at all and 10% do not know. No other application in the survey reaches a no-plans figure of 70%.

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



Accent localisation 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 organisation can point to on a spreadsheet, namely customers it cannot currently serve. Accent conversion addresses comprehension within a language both parties already speak, which while harder to size and justify, certainly impacts on agent and customer experience, quite extremely in some instances.

It is worth setting that 70% against what accent is understood to cost operations that already take it seriously. Offshore 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 organisation afterwards.

Accent conversion is therefore competing with an established alternative rather than filling a vacuum, which may explain the 70% with no plans better than indifference does. The comparison a buyer is actually making is against recruitment screening and a training programme they already run, and against the agent morale cost of asking people to change how they speak.

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 the least deployed of everything measured in the previous chapters.

This section sets out what telephony self-service currently achieves in UK contact centres, 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 centres that offer telephony self-service at all, a mean of 16% 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	23%
Median	10%
3rd quartile	5%
Mean	16%

The median of 10% sits well below the mean of 16%, 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 23% or more of calls this way, and a quarter handle 5% or less.

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.

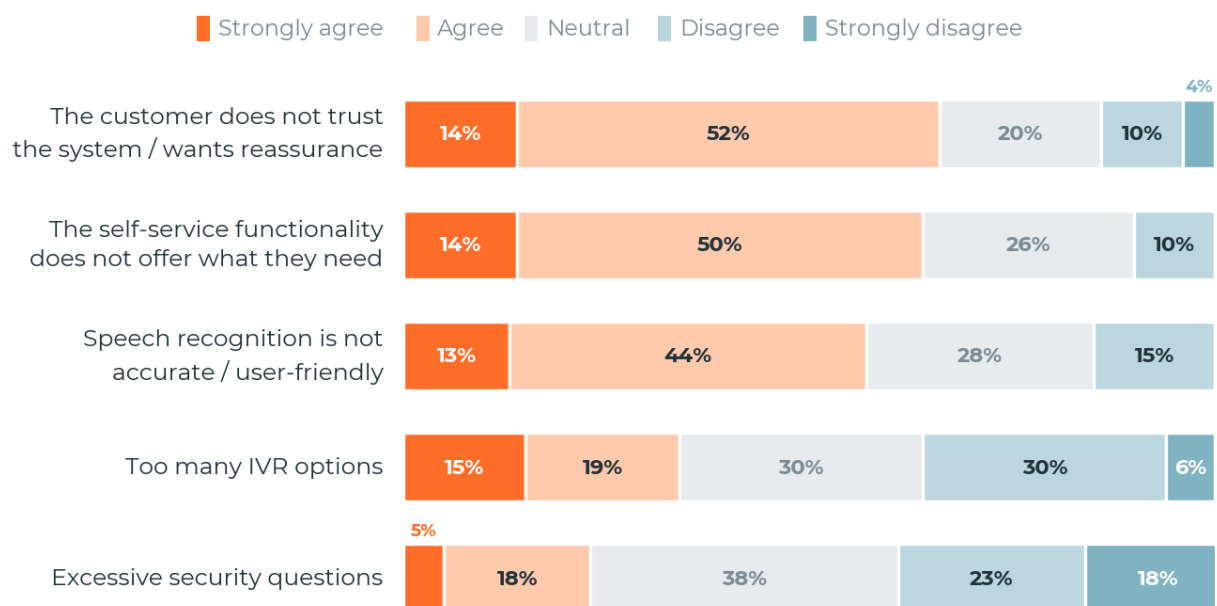
Customers who decline to use it have a straightforward remedy: a mean of 9% of calls that enter the self-service option are zeroed out, meaning the customer presses zero or finds some other shortcut to reach an operator. That figure has improved on previous years, which suggests voice self-service is becoming more acceptable, but it varies considerably by operation size. Large contact centres see a mean of over 13% zeroing out, against 8% to 9% in small and mid-sized operations.

The most plausible explanation is that larger operations attempt more. Respondents from larger organisations 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 centres 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



Two findings dominate: 64% agree to some extent that customers abandon because the self-service function does not offer what they want, and 66% agree that the customer simply does not trust the system, preferring human reassurance that the request has actually been carried out or the information is actually correct.

Among those using automated speech recognition, 57% 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 high figure, and it connects directly to the audio quality material in the previous section.

Only 15% 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 41% disagreeing or strongly disagreeing.

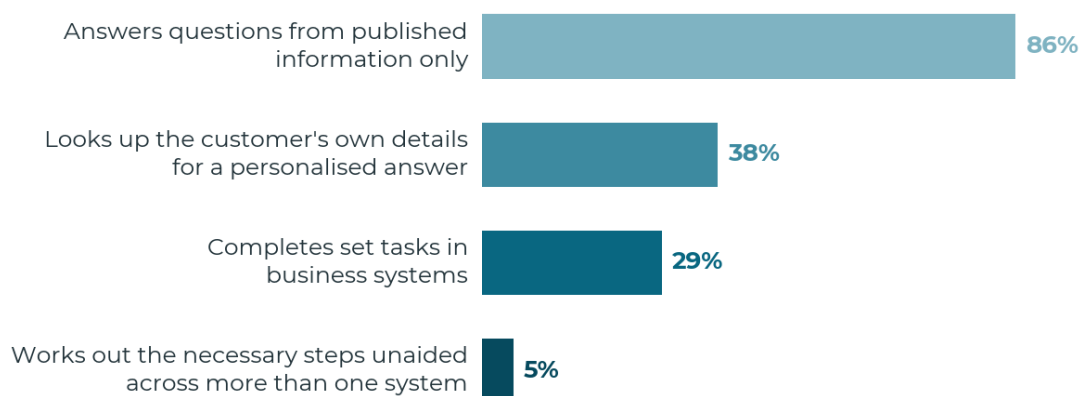
A customer abandoning because the system genuinely cannot do what they need is exercising a reasonable choice, and provided a clear exit to a live agent is marked early in the process, this does not have to be negative. Operations hiding their agents and sending customers around IVR loops does reputational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



86% of bots users (NB – this is from users of both chatbots and/or voicebots) report that their most capable bot answers questions from published information. 38% can look up the customer's own details to give a personalised answer, 29% can complete set tasks in business systems, and 5% can work out the necessary steps unaided across more than one system.

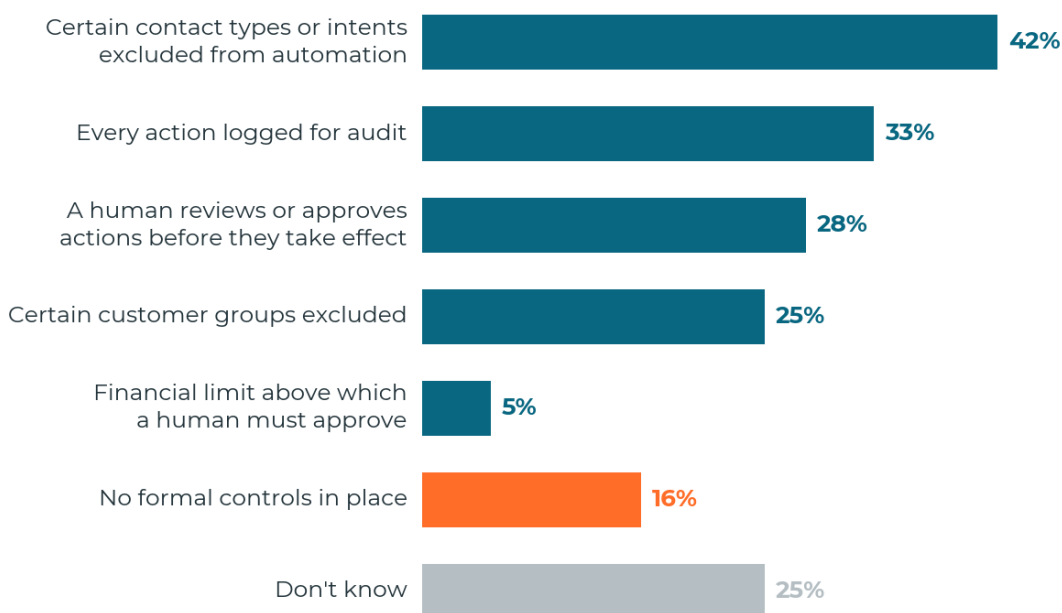
Answering from published content is close to universal, acting in business systems is a minority capability, and the autonomous multi-system reasoning that dominates the marketing of this category exists in one operation in twenty.

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 organisations running a voicebot are running a chatbot alongside it and govern the two together.

42% 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. 33% log every action for audit, 28% require a human to review or approve actions before they take effect, and 25% exclude certain customer groups, which may include customers flagged as vulnerable.

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



It is concerning that 16% report no formal controls in place at all, and a further 25% do not know what controls exist. Taken together, two in five organisations either have no governance over what their bots do in live business systems or cannot say whether they have. Only 5% 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 29% 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 organisations running them, voicebots handle 19% of enquiries entirely without human involvement. The equivalent figure for chatbots is 20%.

The near-identical voicebot and chatbot figures are a surprise. The received view is that voice is intrinsically harder to contain than text, and on the evidence of these two numbers it is not, at least among organisations that have got as far as deploying a voicebot at all.

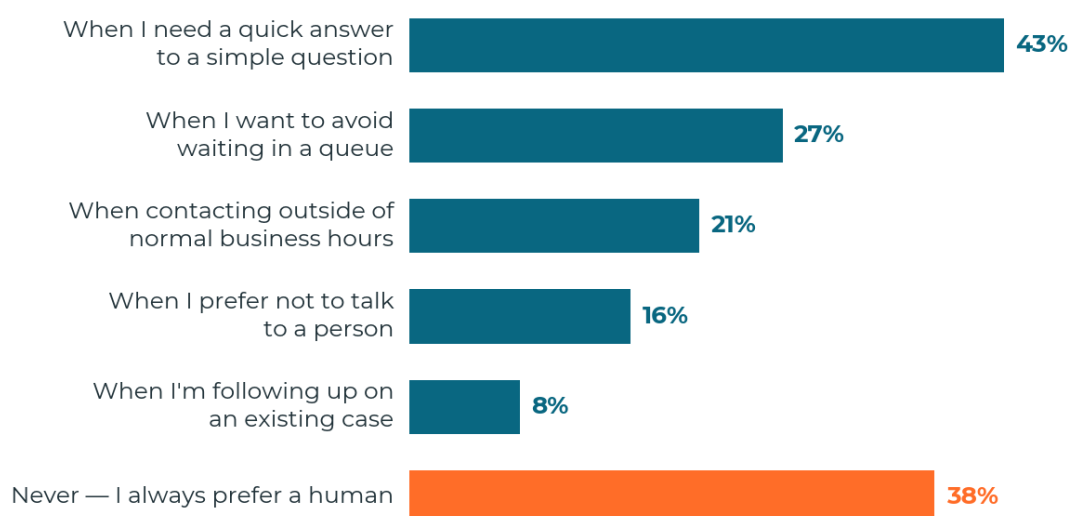
Two data points to note: firstly, selection, in that only 14% of respondents currently run a voicebot, and those that do are self-selected as the organisations with the appetite, the budget and the use cases to make one work. The second is that 19% sits above the 16% mean and well above the 10% median for telephony self-service as a whole, which is consistent with voicebots outperforming touchtone IVR although it is not spectacularly better yet.

WHEN CUSTOMERS PREFER TO DEAL WITH AI

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

62% of UK 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 organisation, in which situations would you prefer to deal with AI rather than a human?



43% would prefer AI when they need a quick answer to a simple question, which is by some distance the most popular situation and aligns closely with what most AI in customer service is currently used for.

27% 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 21%.

16% 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.

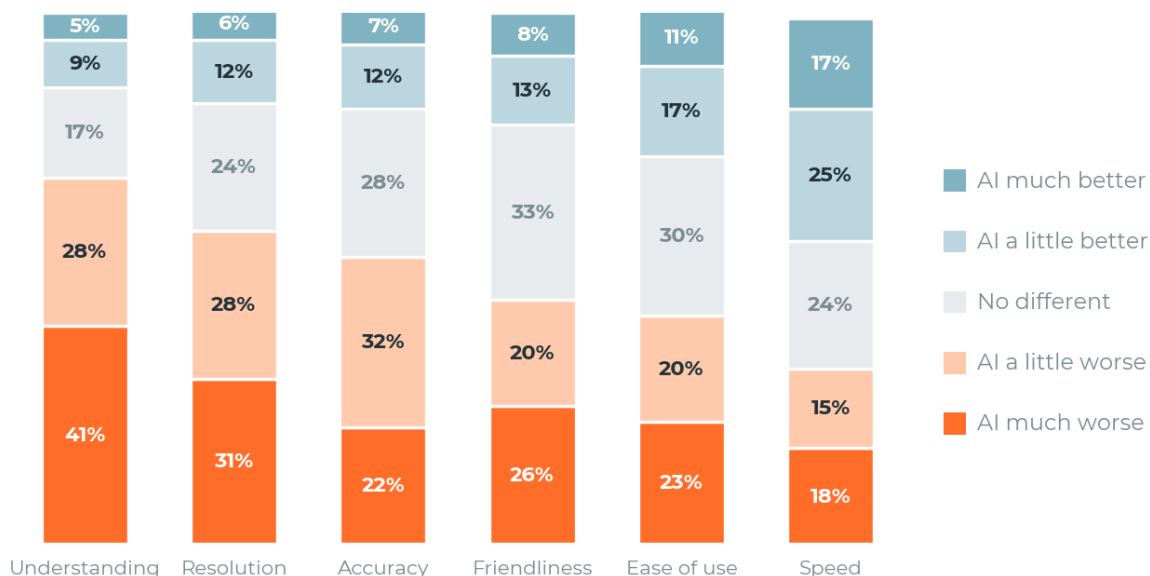
Against that, 38% of customers currently say they would never prefer AI and always want a human, with the figure skewed towards older customers.

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 is unambiguous and points in one direction on five of the six measures.

69% rate AI worse than a human for understanding the issue, against 14% who rate it better. 59% rate it worse for overall resolution and 54% worse for accuracy of information. Even on friendliness and tone, where a machine might be expected to hold up reasonably well, 46% rate it worse against 21% better.

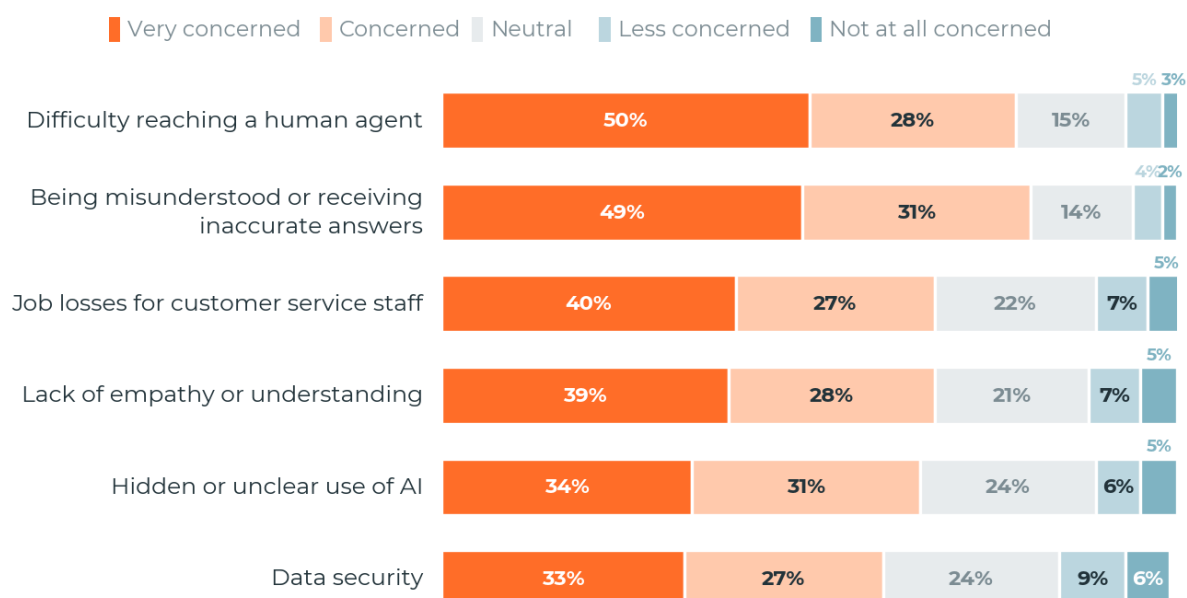
Speed of response is the single exception and the only measure where AI comes out ahead, with 42% rating it better against 33% worse.

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?

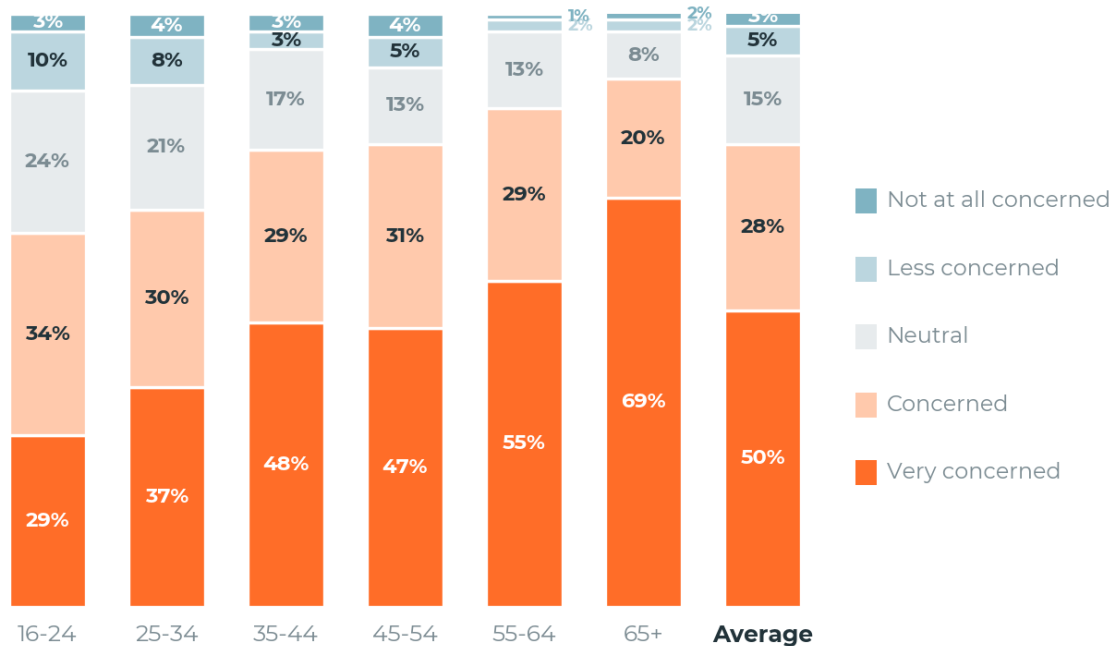


Difficulty reaching a human agent and being misunderstood or given inaccurate answers are the two leading concerns, at 78% and 80% respectively when very concerned and concerned are combined.

Both sit above job losses at 67%, lack of empathy at 67%, hidden or unclear use of AI at 65% and data security at 60%.

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)



69% of customers aged 65 and over are very concerned about being unable to reach a human, and 89% are concerned to some extent. Among 16 to 24 year-olds the very concerned figure is 29%, less than half the level. 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.

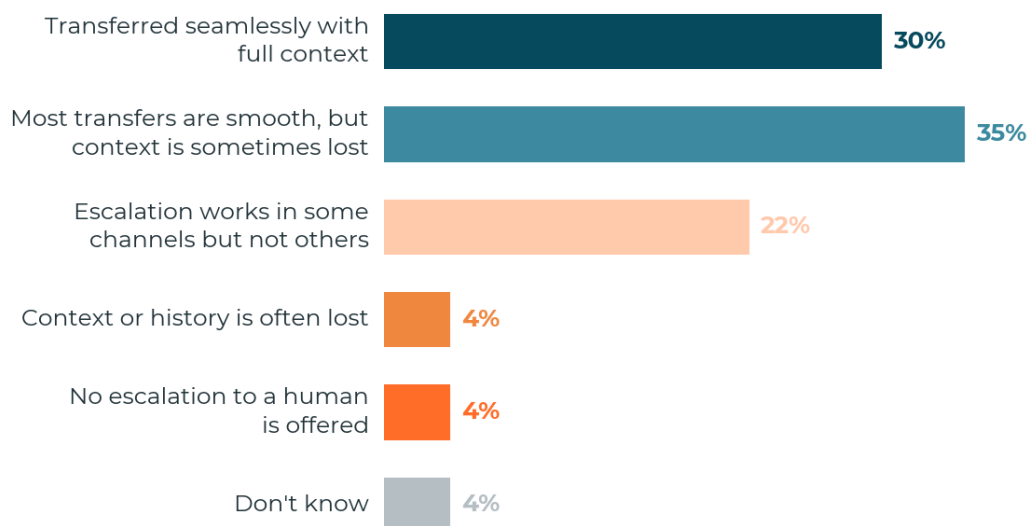
It can identify and authenticate the customer before passing the call on, which saves an average of 48 seconds per authenticated call (across the UK contact centre industry, agent-handled authentication accounts for over £3bn a year).

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



30% report that customers are transferred seamlessly with full context, and a further 35% that most transfers are smooth but context is sometimes lost.

22% say escalation works in some channels but not others, which is a description of a partial integration rather than a design.

The tail is surprisingly small: only 4% say context or history is often lost, and 4% offer no escalation to a human at all.

Read positively, 65% of organisations manage a handover that works most or all of the time, which is better than the industry's reputation would suggest. Read less charitably, only 30% 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 16% of calls handled end to end where telephony self-service is offered (and a median of only 10%) is not particularly impressive. Voicebots are reported to increase this to 19%, and while this is still not spectacular, it will certainly improve as functionality and customer confidence and familiarity increase.

Note that 38% 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: 25-40% would seem a reasonable target.

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 69% of consumers who consider AI worse at understanding the issue, and against the 78% 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

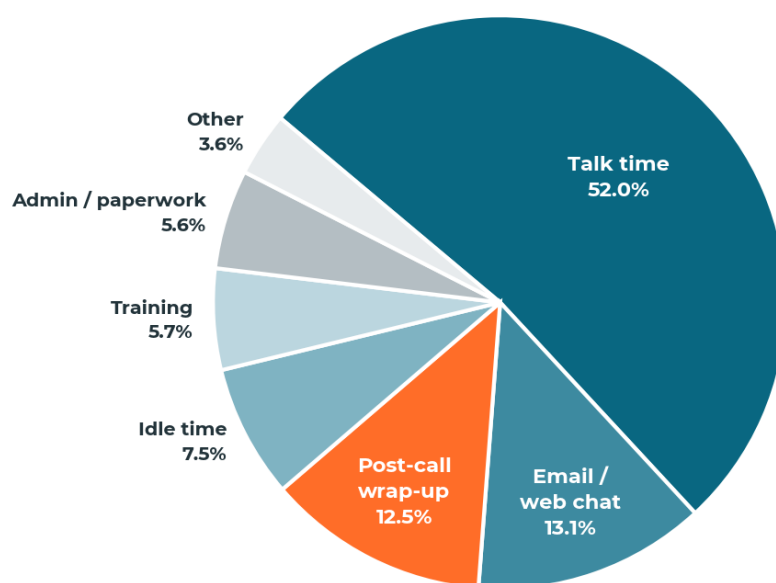
The voice AI applications in this chapter share a characteristic that separates them from everything in the previous one: the customer is still speaking to a person.

AI is listening, retrieving, drafting and summarising 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 centre day is currently spent.

Figure 23: Agent activity

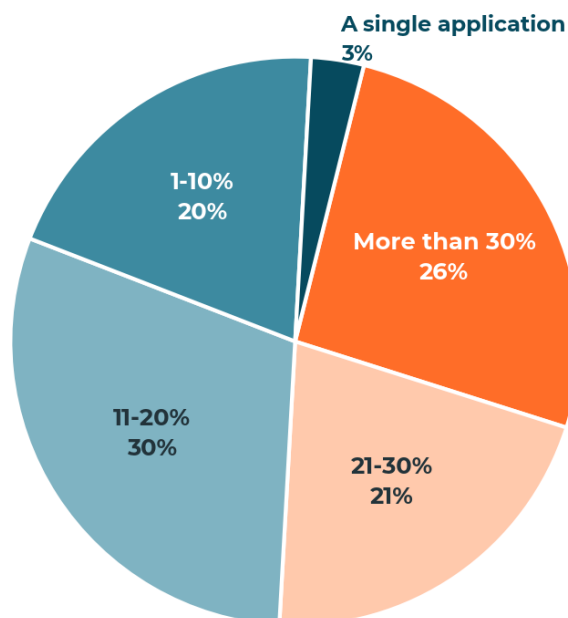


Talk time accounts for 52% of agent activity, with a further 13.1% on email and web chat. The largest remaining block is post-call wrap-up at 12.5%, which is time spent on work the customer never sees: writing notes, coding dispositions and updating systems after the conversation has ended. Idle time takes 7.5%, training 5.7% and administration 5.6%.

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, the conversation itself, has already been captured. We have calculated that at an average length of 101 seconds, the average cost of post-call work is £1.09 per call.

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

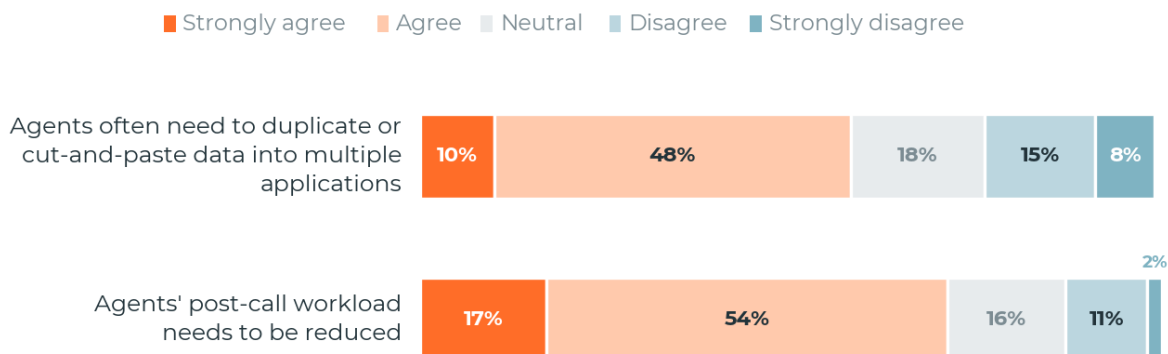
Figure 24: Proportion of inbound call spent navigating between screens / applications



Only 3% of contact centres report that agents handle inbound calls within a single application. At the other extreme, 26% report that more than 30% of the call is spent navigating between screens, and a further 21% put it between 21% and 30%. For roughly half of UK operations, then, something approaching a third of every call is spent on system navigation rather than on the customer.

Respondents recognise the problem and its consequences.

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

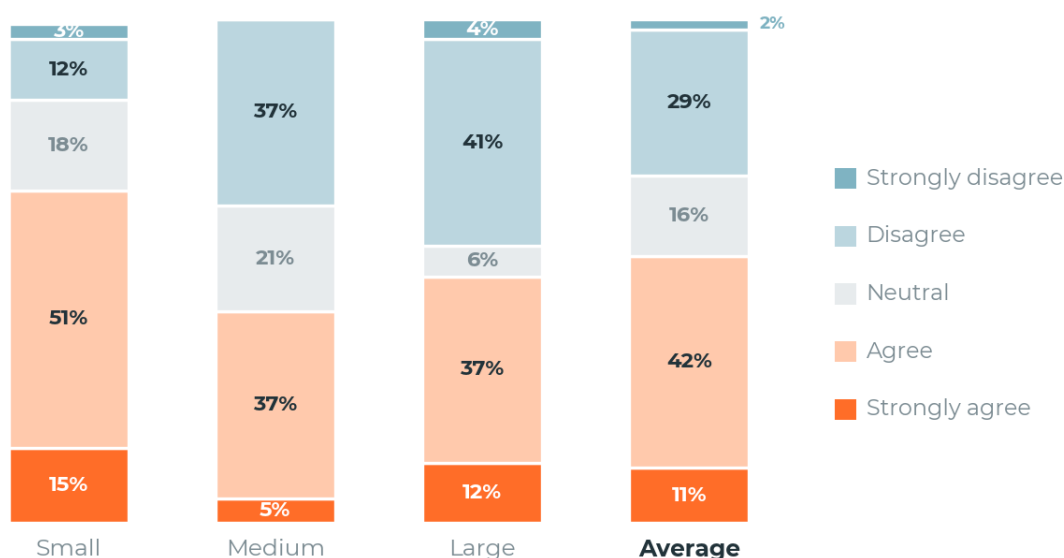


58% agree or strongly agree that agents often need to duplicate or cut and paste data into multiple applications, a pattern found consistently across all contact centre size bands and one that produces transcription errors as well as wasted time. 71% agree that agents' post-call workload needs to be reduced, with only 13% disagreeing.

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 familiarise themselves with our systems", by contact centre size



53% agree or strongly agree that new agents struggle to familiarise themselves with the systems, and the pattern by size is not the one most people would predict.

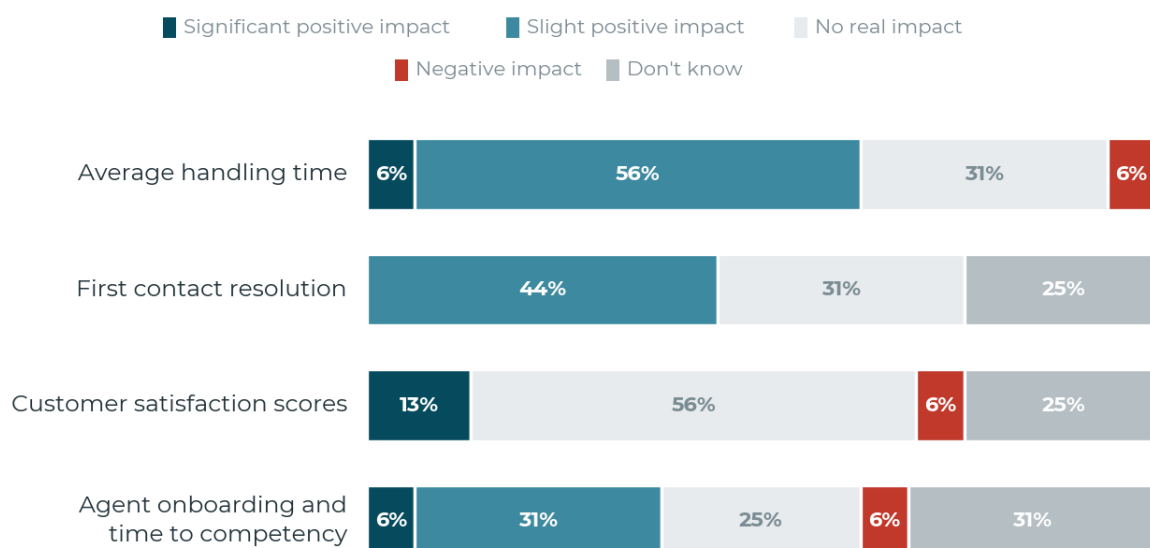
Small operations report the greatest difficulty at 66% agreement, against 42% in medium-sized operations and 49% in large ones. Smaller contact centres tend to have fewer dedicated training resources and agents who handle a wider range of enquiry types, which is a plausible explanation.

This matters for the business case because AI that surfaces the right answer at the right moment reduces what a new agent has to memorise 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. 63% report a positive impact on AHT, of which 56% describe it as slight and 6% as significant, against 31% reporting no real impact and 6% a negative one. It is also the only metric on which every respondent had a view.

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



First contact resolution is weaker. 44% report a slight positive impact and none a significant one, 31% see no real impact, and 25% do not know.

Customer satisfaction is weaker still: 56% report no real impact at all, 13% a significant positive impact, 6% a negative one and 25% do not know.

That pattern deserves attention, because it is not the one the technology is usually sold on. Assistance is making calls shorter without, so far, making them measurably better for the customer. The most plausible explanation is that surfacing the right answer faster removes agent search time without changing the answer itself, which is a real efficiency gain but not the experience improvement often claimed alongside it.

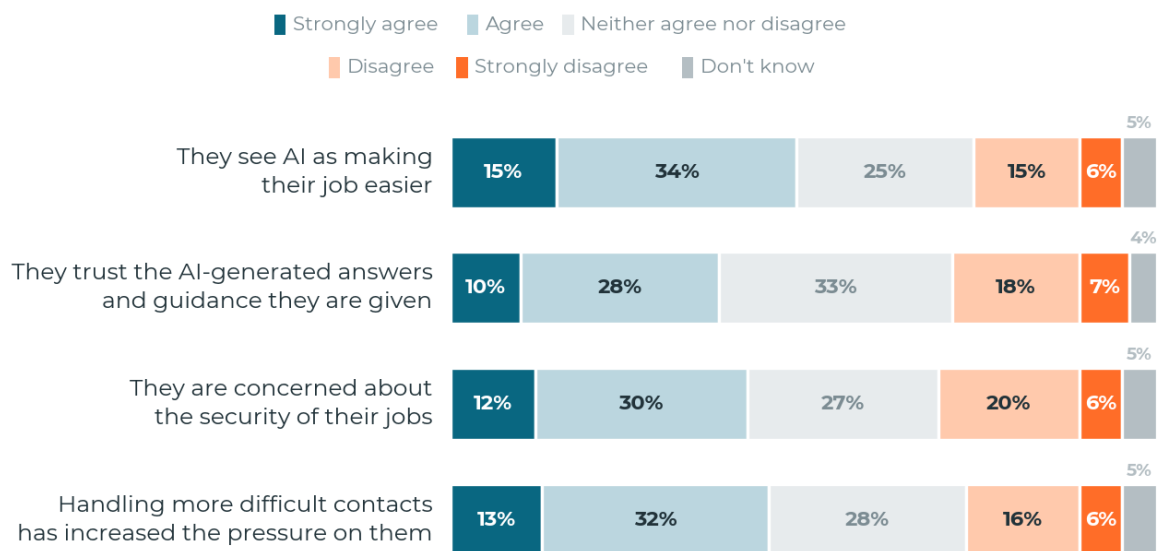
Agent onboarding and time to competency shows 38% reporting a positive impact and 31% not knowing, the highest don't-know proportion of the four. Time to competency is rarely instrumented, so that is likely to reflect measurement rather than absence of effect.

These figures come only from organisations 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



49% agree or strongly agree that their agents see AI as making the job easier, against 21% who disagree. That is a positive balance, though a quarter of respondents place their agents in the middle, which for a question about one's own staff is itself a finding.

38% agree that agents trust the AI-generated answers and guidance they are given, 25% disagree, and 33% sit in the middle. Guidance that agents do not act on delivers nothing regardless of how accurate it is, and a system in which only 38% of agents are believed to trust the output is not yet delivering its full value. It also suggests that the accuracy concerns of customers reported earlier are indeed something agents encounter in daily use.

42% agree that agents are concerned about the security of their jobs and 26% disagree, which sits alongside the headcount data later in this chapter showing that redundancies attributable to AI remain rare.

45% agree that handling more difficult contacts has increased the pressure on agents, which is the predictable consequence of automation removing the simple work first, and connects directly to the doubling of mean call duration shown in Figure 3.

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.

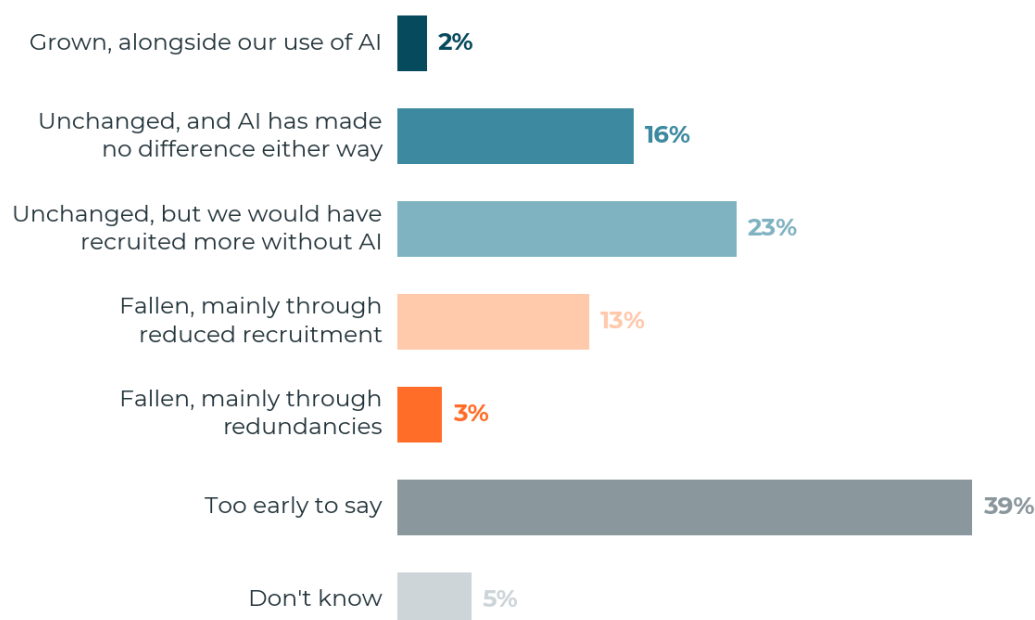
WHAT AI HAS DONE TO HEADCOUNT

Looking at the current effect of AI on agent headcount, the single largest response, at 39%, 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. 23% report headcount unchanged but state that they would have recruited more without AI, and 13% report headcount falling mainly through reduced recruitment.

Only 3% report headcount falling mainly through redundancies, and 2% report growth alongside AI.

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



So 36% describe an effect that operates through hiring rather than firing. This is consistent with the attrition data published elsewhere by ContactBabel, where UK contact centre attrition of 17% means a typical operation replaces roughly one agent in six every year and can therefore reduce headcount substantially without a redundancy programme.

On what organisations did when introducing AI, 64% gave agents formal training on working alongside it and 48% consulted agents or their representatives before deployment.

16% changed agent targets or performance measures to reflect the new way of working, which is a low figure given that 45% believe the residual work has become harder.

23% did none of these things.

ANALYTICS, QUALITY AND ASSURANCE

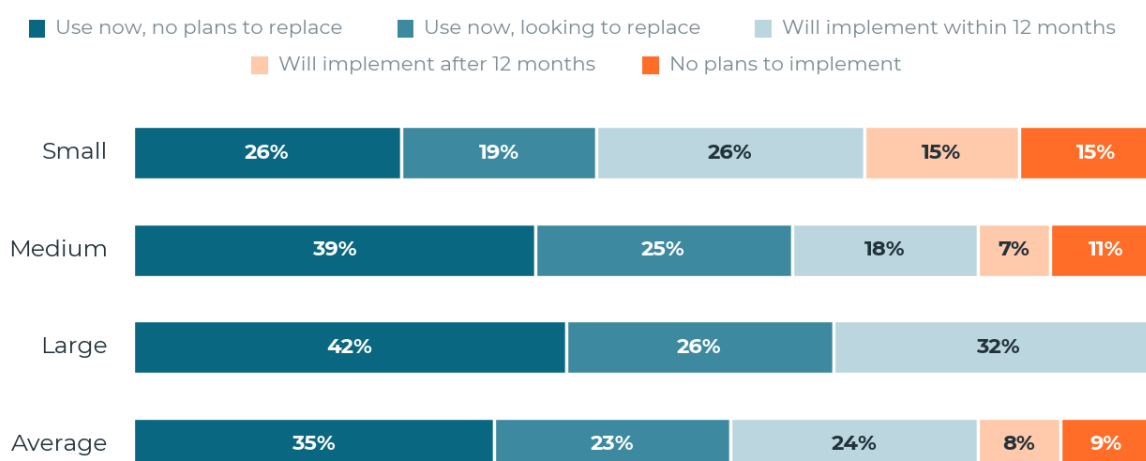
This section is about what is done with the transcripts created through voice AI, how widely interaction analytics is deployed, what organisations use it for, how well quality assurance currently works, and the question that is rarely asked: how AI-handled conversations are themselves quality assured.

HOW WIDELY ANALYTICS IS DEPLOYED

58% of UK contact centres use interaction analytics, comprising 35% with no plans to replace their system and 23% currently using but looking to replace or upgrade.

A further 24% expect to implement within 12 months and 8% after that, leaving only 9% with no plans at all.

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



The size effect is pronounced, with large operations reporting 68% current usage and, notably, none at all with no plans to implement, while small operations report 45% usage with 15% having no plans.

The proportion looking to replace an existing system is high across every size band, which suggests a market in which first-generation deployments are reaching the end of their useful life.

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	74%
Customer journey analytics	57%
Multichannel analytics (email, web chat, social media)	45%
Back office analytics	34%
Desktop analytics	33%
Real-time speech analytics	29%

Post-call analysis remains the dominant use at 74%, while real-time analytics, which is the capability most closely associated with voice AI in supplier marketing, is used by only 29% of analytics users.

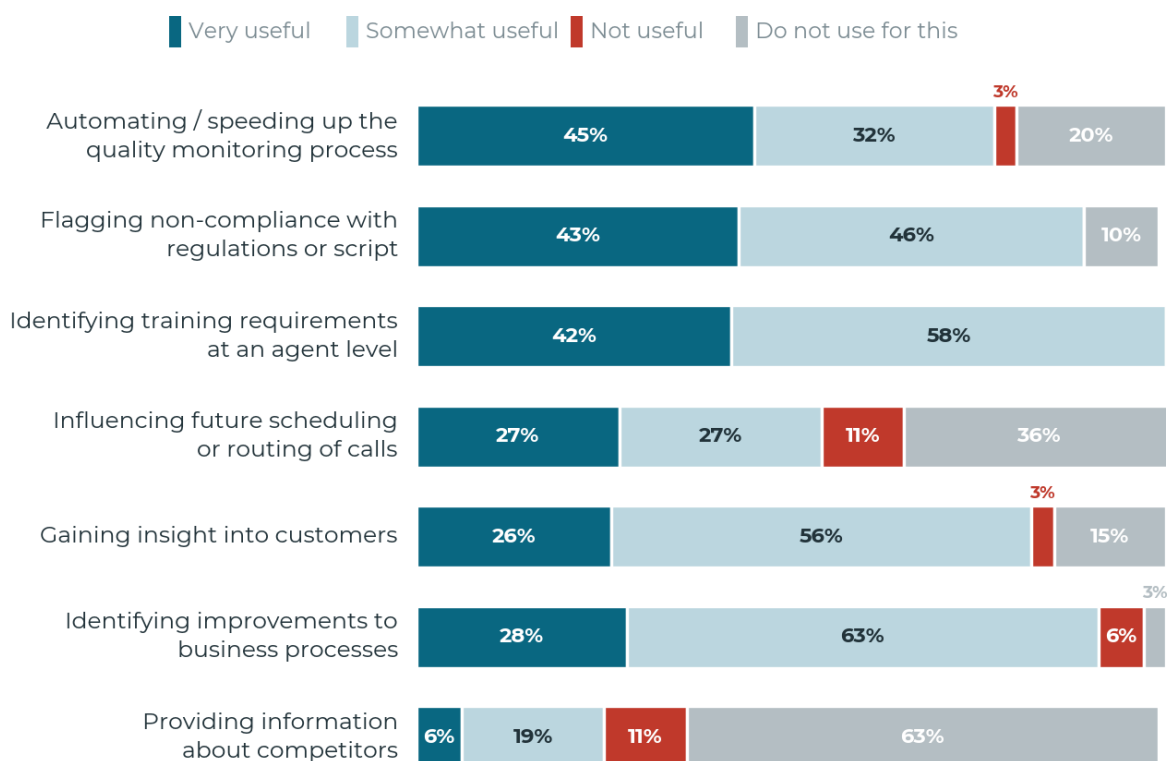
WHAT ANALYTICS IS USED FOR

The three applications rated most useful are all quality and compliance functions rather than customer insight ones.

45% rate automating and speeding up the quality monitoring process as very useful, 43% say the same of flagging non-compliance with regulations or script, and 42% of identifying training requirements at an agent level.

That last application is the only one in the list that no respondent rates as not useful and that every user of analytics reports using.

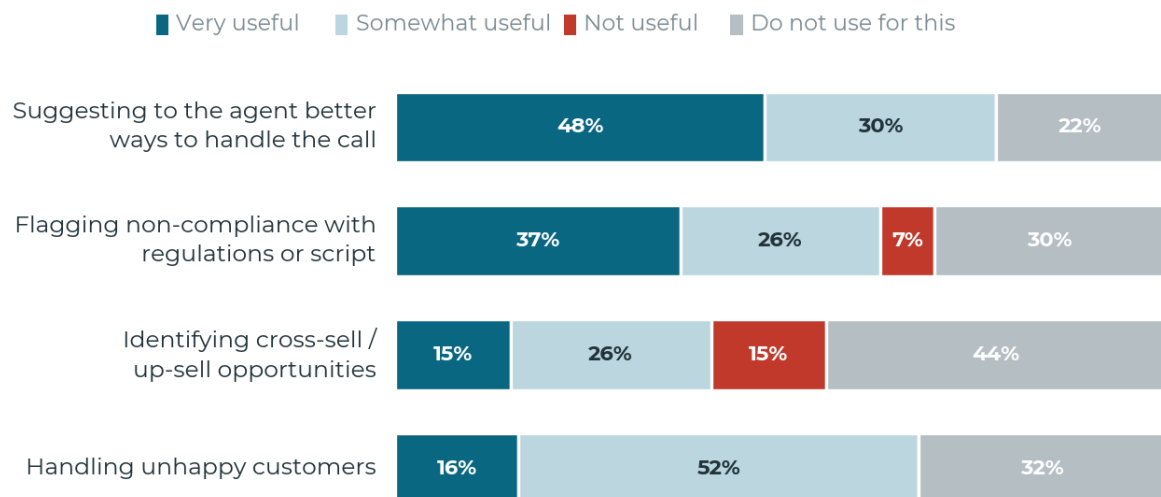
Figure 32: Usefulness of post-call analytics



At the other end, providing information about competitors is not used by 63% of analytics users, and influencing future scheduling or routing of calls is not used by 36%.

Among the minority who use real-time analytics, suggesting to the agent better ways to handle the call is rated very useful by 48%, ahead of flagging non-compliance at 37%. Identifying cross-sell and up-sell opportunities is the weakest application on both measures, rated very useful by 15% and not used at all by 44%. Real-time analytics is being used to help the agent rather than to sell.

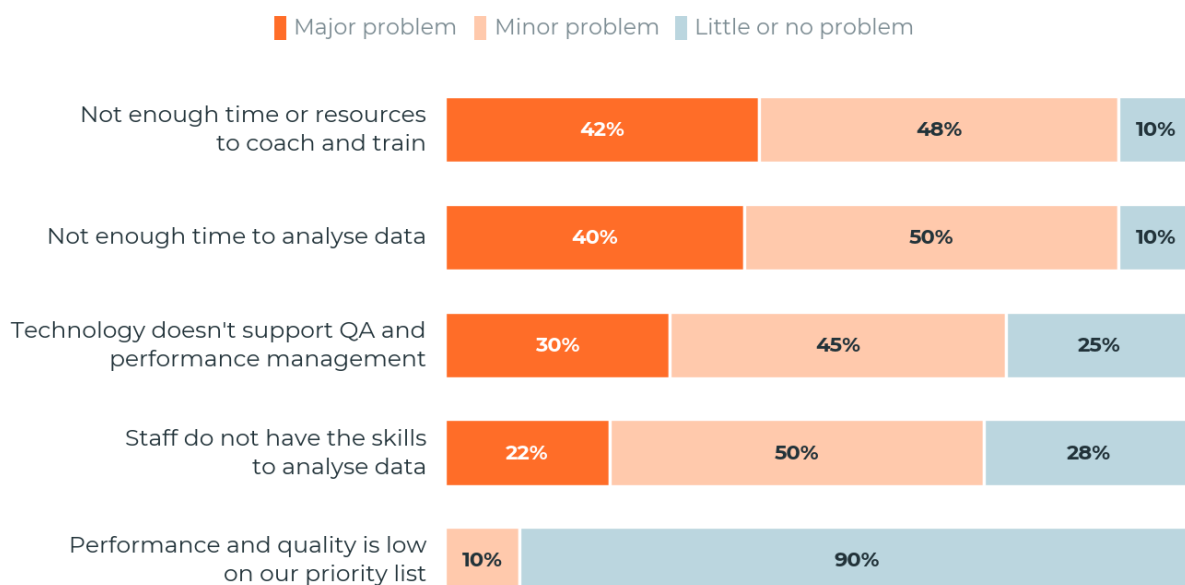
Figure 33: Usefulness of real-time analytics



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

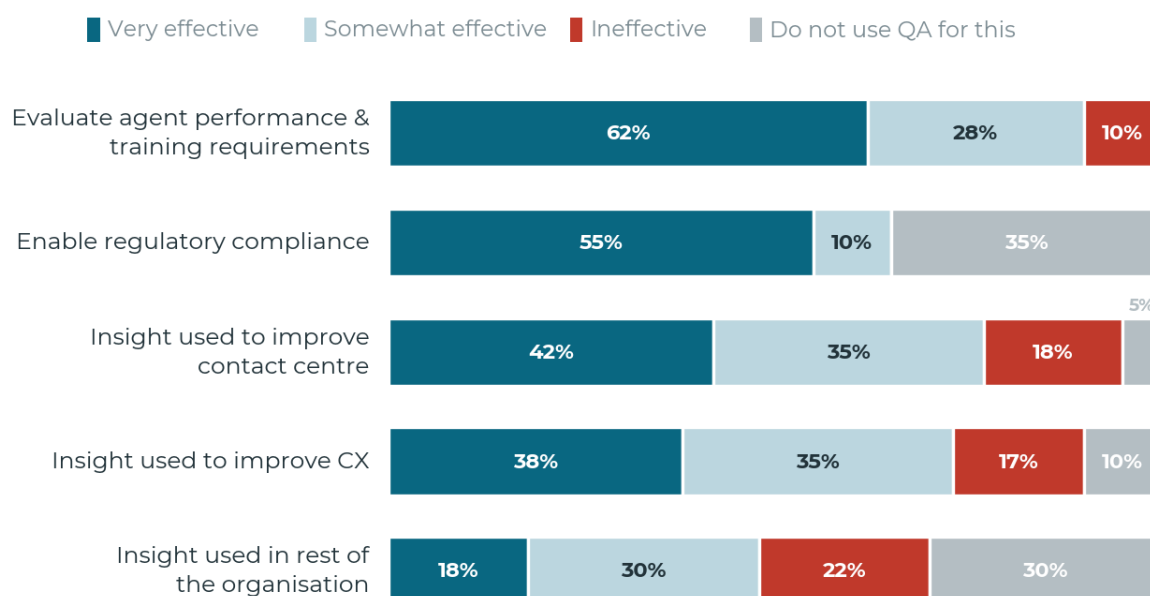


42% describe not having enough time or resources to coach and train as a major problem, and a further 48% as a minor one, meaning 90% experience it to some degree. Not having enough time to analyse the data reaches the same 90% total, with 40% calling it major. Only 10% report little or no problem with either.

The two 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 organisation 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. 90% say that performance and quality being low on the priority list is little or no problem, and only 22% call the lack of data analysis skills a major issue. The constraint is time, not intent or capability.

Figure 35: Effectiveness of quality assurance



Where quality assurance is applied to its traditional purpose it works reasonably well: 62% rate it very effective for evaluating agent performance and training requirements, and 55% for enabling regulatory compliance.

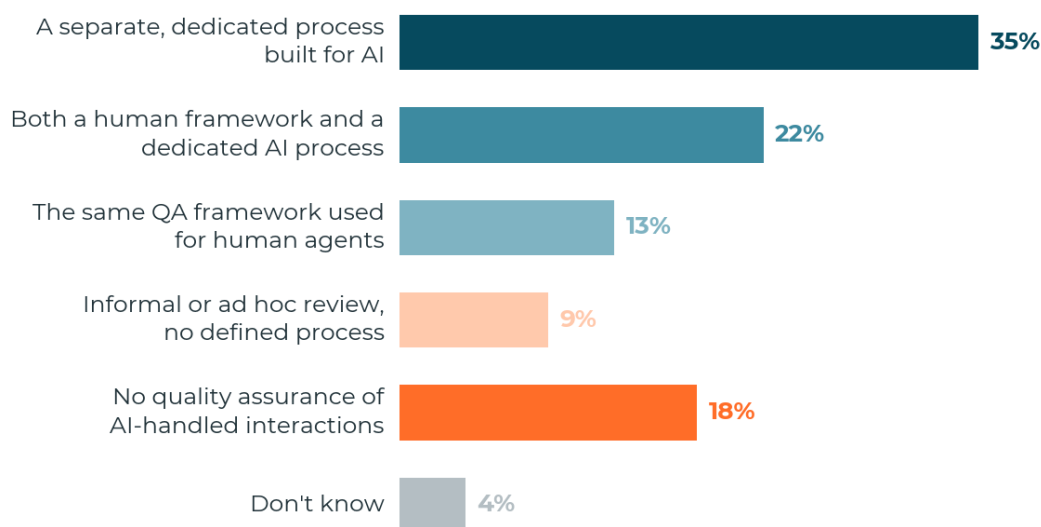
Effectiveness falls away as the insight travels further from the contact centre floor, with 42% rating it very effective for improving the contact centre, 38% for improving CX, and only 18% for insight used in the rest of the organisation, where 30% do not use QA recordings for this purpose at all and 22% find it ineffective.

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 organisations are actually asking those questions of their AI is not something published research has often measured.

Figure 36: How organisations quality assure interactions handled by AI



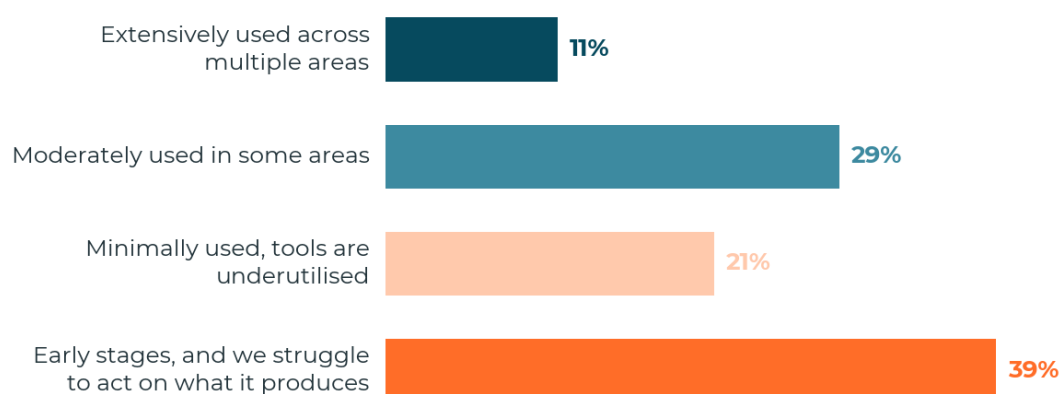
35% have built a separate, dedicated process for assuring AI-handled interactions, and a further 22% run both that and their human framework in parallel: 57% use something purpose-built, which is a higher figure than expected for a recent discipline. 13% apply the same QA framework they use for human agents.

On the other hand, 18% carry out no quality assurance of AI-handled interactions at all, 9% review informally with no defined process, and 4% do not know. Roughly three in ten organisations 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 16% have no formal controls over what bots may do and 25% do not know. An organisation 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 11% report that AI-driven analytics is used extensively across multiple areas to generate actionable operational insight. 29% describe moderate use in some areas, 21% say the tools are underutilised, and 39% say they are at an early stage and struggle to act on what the analytics produce.

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



That is the coaching bottleneck shown earlier reappearing one layer up: 90% of operations report insufficient time to analyse their data, and the technology that solves it produces findings which at least 60% then cannot act on.

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

VOICE, IDENTITY AND FRAUD

Before most UK 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.

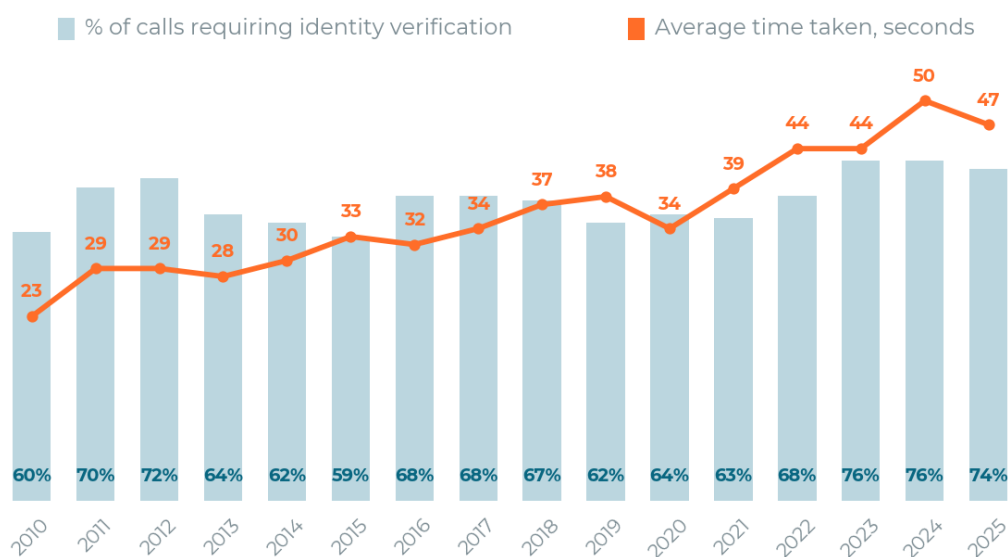
THE SCALE OF THE AUTHENTICATION PROBLEM

74% of UK inbound calls require caller identity verification, close to the highest figure on record. Over the past fifteen years that proportion has moved within a band of roughly 60% to 76%, so the volume of authentication is not what has changed.

What has changed is how long it takes: the average has risen from 23 seconds in 2010 to 47 seconds in 2025, having peaked at 50 seconds in 2024. Verification now takes more than twice as long as it did fifteen years ago, the result of more stringent testing rather than of more calls needing it.

37% of respondents state that all callers go through identity verification, and only 9% that they never do.

Figure 38: Proportion of calls requiring caller identification & average time taken, 2010-25



The distribution of how authentication happens is striking: 91% of authentication is still handled by a person reading questions from a screen, and voice biometrics, a technology designed precisely for this task, accounts for under 1%.

Figure 39: Caller identity authentication methods (only those contact centres which authenticate some or all calls)

Identification method	Proportion of callers identified using this method
Agent	91%
DTMF IVR (touchtone)	7%
Speech recognition	2%
Voice biometrics	<1%

Looking at time taken, the mean of 48 seconds that agents take to authenticate a customer sits well above the median of 30, so a minority of long authentications is pulling the average up.

Authentication through an IVR takes 40 seconds on average, but because the agent is not occupied doing this, the cost to the operation is lower even where the elapsed time is similar.

Figure 40: Time taken to authenticate caller identity using only an agent (seconds)

Seconds to authenticate using only an agent	
1st quartile	20
Median	30
3rd quartile	60
Mean	48

ContactBabel estimates the industry-wide cost of agent-handled authentication at over £3bn a year, derived from the 74% of calls requiring verification, the 91% handled by agents and the 48-second mean.

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 48-second process.

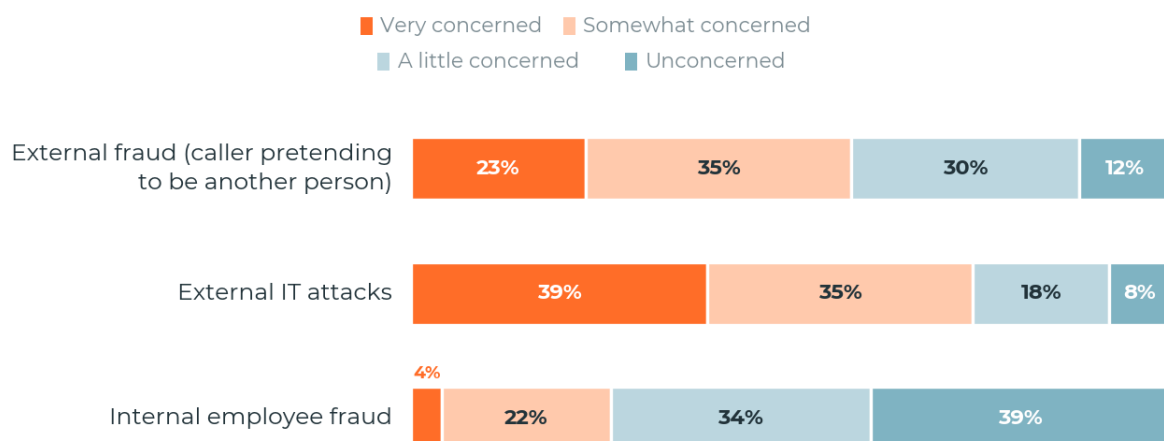
It has nonetheless remained under 1% of authentications for as long as ContactBabel has measured it. Investment intent, however, has moved sharply: voice biometrics appeared in the top five IT expenditure priorities of 11% of UK contact centres in 2025, against 5% in 2024 and a range of 3% to 7% in every year from 2015 onwards. That is the highest figure recorded for the category and more than double the previous year.

There are two things to note: first, that passive authentication placed in front of a voicebot removes the 48 seconds before a human is involved at all, which strengthens the case for voice self-service independently of whether the bot resolves the enquiry.

The second is that 58% of contact centres are very or somewhat concerned about external fraud from callers pretending to be someone else, and 74% about external IT attacks, where 39% are very concerned.

Internal employee fraud attracts markedly less concern, with 39% unconcerned and only 4% very concerned, which is notable given that around 70% of UK agents now work remotely or in a hybrid pattern.

Figure 41: Concerns about security

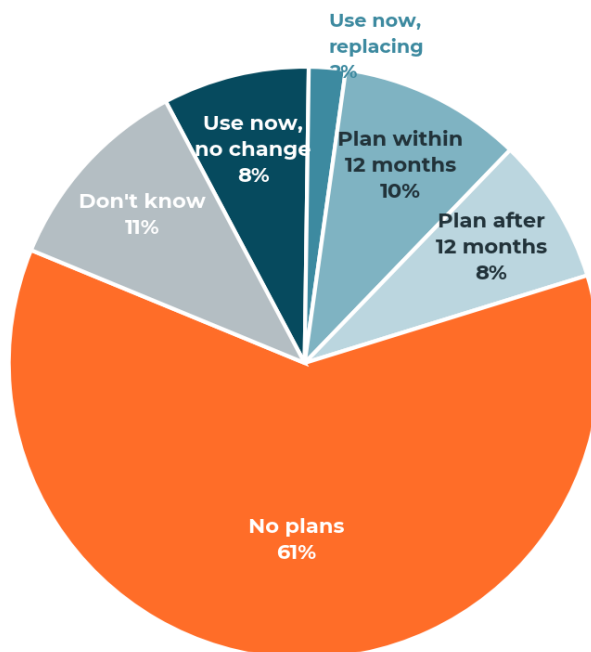


The difficulty for voice biometrics specifically is that generative voice models can now synthesise a convincing imitation of a particular speaker from a short sample. A credential that can be produced from a recording of the customer speaking is a weaker credential than it was when the technology was 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 42: Current and planned use of voice biometrics or AI-based authentication and fraud detection



10% report using voice biometrics or AI-based authentication and fraud detection, with a further 18% planning to implement. 61% have no plans at all, which is one of the highest “no-plans” figure of any application in the survey, and 11% do not know.

While the previous finding that voice biometrics accounts for under 1% of actual authentications might make this 10% 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 organisations running behavioural or device-based risk scoring, transcript-based fraud detection, or anomaly detection on call patterns, none of which identify anyone by voice. The 10% is therefore AI applied to the fraud and identity problem generally, and the under 1% figure is how many callers are actually identified by their voice.

The more useful reading is that 61% of the market has no plans in this area at all. Authentication is simultaneously the largest single mechanical inefficiency in the UK contact centre, at well over £3bn a year, and is also the application organisations are least inclined to address. 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 UK RULES

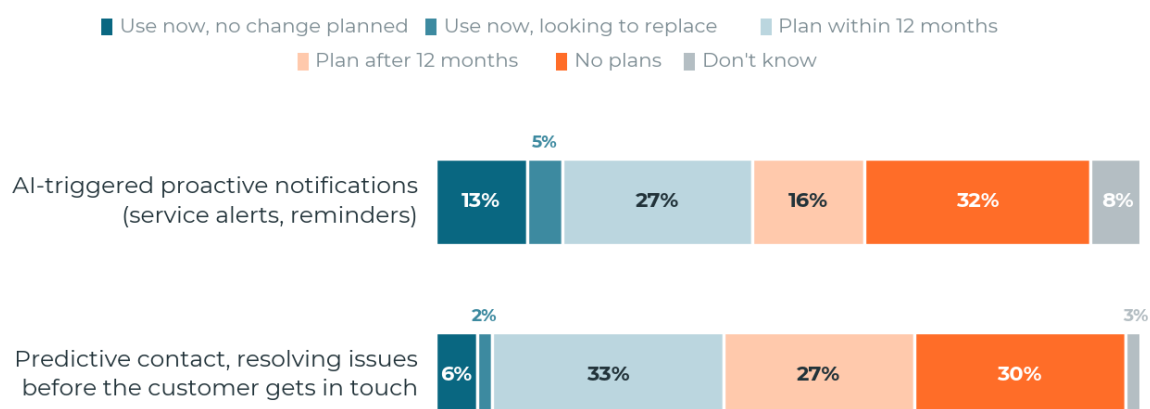
Everything so far in this report concerns calls the customer makes. This section concerns calls the organisation makes, which is where voice AI meets a body of regulation that predates it by two decades and was not written with conversational systems in mind. While the commercial case is quite straightforward, the compliance position is not, and the penalties changed materially in 2025.

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 organisation would have paid £6.17 to handle into a short outbound contact it controls the timing of, and in many cases prevents the enquiry arising at all.

Figure 43: Current and planned use of proactive and predictive outbound AI



18% already use AI-triggered proactive notifications such as service alerts and appointment reminders, and 8% use predictive contact to resolve issues before the customer gets in touch. A further 43% and 60% respectively plan to implement, leaving 32% and 30% with no plans.

Predictive contact shows one of the largest gaps between current and planned use of anything measured in the survey, at 8% now against 67% projected.

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 (even by a bot), and many organisations will use all three depending on urgency and on what the customer has agreed to. The regulatory analysis that follows applies specifically to the voice element, and the equivalent rules for electronic mail and text messages sit in Regulation 22 of PECR rather than Regulations 19 and 21.

A KEY DISTINCTION TO UNDERSTAND

The Privacy and Electronic Communications Regulations (PECR) focus much more on direct marketing, rather than service communications.

An appointment reminder, an outage alert or a delivery update relating to an existing contract falls outside the marketing rules. Marketing dressed as service does not, and the distinction is drawn on the purpose and content of the call rather than on what the organisation chooses to call it.

Most legitimate proactive voice AI in a contact centre appears to us to sit on the service side of that line. The regulatory analysis below matters most for organisations whose outbound has any promotional element at all.

Please note that this information is provided as being our current best understanding and opinion only, and should not be acted upon without first taking legal advice.

WHAT THE RULES REQUIRE

Regulation 19 governs automated calling systems. Regulation 19(4) defines these as systems capable of automatically initiating a sequence of calls to more than one destination in accordance with stored instructions, and of transmitting sounds which are not live speech.

Where such a system is used for direct marketing, Regulation 19(1) requires the prior specific consent of the subscriber. Screening against the Telephone Preference Service (TPS) is not a substitute, because the test is consent rather than absence of objection. Regulation 24 separately requires the caller to be named and, on automated calls, an address or a free-of-charge contact number to be provided.

Regulation 21 covers unsolicited direct marketing calls generally, with the automated ones carved out into Regulation 19. Live calls are permitted subject to TPS and Corporate TPS screening and to respecting any objection the individual has previously made.

The unresolved question is which of the two a conversational agent falls under, and the ambiguity runs through both limbs of Regulation 19. Whether real-time synthesised output is "live speech" within the meaning of Regulation 19(4) has not been tested, and nor has whether it amounts to "recorded matter" for the purposes of the prohibition in Regulation 19(1).

Read narrowly, both terms describe pre-recorded audio and a voicebot escapes each. Read purposively, the mischief the regulation addresses is the automation of marketing calls at volume, which is exactly what the technology delivers.

The Information Commissioner's Office has not yet formally settled the point. The defensible position, and in our view the one a regulator would be most likely to take, is to treat a synthetic voice as an automated calling system and obtain prior specific consent, a materially higher bar than TPS screening.

The cost of guessing wrong has also risen sharply, the maximum PECR penalty having moved from £500,000 to the higher of £17.5m or 4% of global turnover with effect from February 2026, and the previous requirement to show substantial damage and distress having been removed.

Ofcom's policy on persistent misuse applies in parallel, though not in the way most operators assume. It sets no permitted abandoned call rate as the 3% figure was withdrawn in 2017, Ofcom's position being that it had never been a compliance threshold and was only ever a means of prioritising enforcement.

An abandoned call rate is still calculated, per calendar day and on a per campaign or per call centre basis, but solely as one input into how cases are prioritised. Misuse becomes actionable where it is repeated enough to show a pattern of behaviour, or recklessness as to consumer harm.

For voice AI the exposure is not dialler pacing but latency and handover. Silent calls, Ofcom's stated top enforcement priority, are defined to include hearing nothing on answering, or hearing anything that falls short of a message directed at the recipient.

Dead air while a model generates its opening turn is capable of meeting that description. Ofcom sets no hard limit, but in assessing harm it considers whether the recipient is connected to a message as quickly as possible and no longer than two seconds after they begin to speak, where the technology relies on recognising a voice to classify the call.

That test applies to automated and interactive messages, not only to information messages on abandoned calls, which is what brings a conversational agent within it. Ofcom also treats an unfairly long wait as misuse where a recipient has asked to be put through to an agent, which is the standard voicebot escalation path.

Answering machine detection is not restricted, but false positives that disconnect on a live person become silent calls, and organisations using AMD are expected to evidence its accuracy on request. Liability extends to organisations that engage third parties to call on their behalf.

General Condition C6 requires communications providers to supply valid, dialable calling line identification that uniquely identifies the caller and, where technically feasible, to block calls failing those tests. General Condition B1 governs the allocation and use of telephone numbers. Both bind the network operator rather than the calling organisation, but the practical effect on a non-compliant outbound operation is the same, in that the traffic is filtered before it reaches the called party.

UK GDPR sits over all of it: a lawful basis for the processing, transparency about what the call is doing, a data protection impact assessment where the processing is high risk, and, where the call produces a decision by solely automated means, the safeguards regime in Articles 22A to 22D. These replaced Article 22 in February 2026 and permit such decisions for most personal data provided the required safeguards are in place, with special category data treated more restrictively.

THE ENFORCEMENT PRECEDENT

In September 2025 the ICO fined two energy sector companies a total of £550,000 over automated marketing calls made using avatar software. The software presented the calls as coming from named UK-based advisors, when they were in fact pre-recorded scripts voiced by actors and triggered by call agents overseas.

The calls in that case were pre-recorded scripts, which is recorded matter on any reading of Regulation 19, so the case does not resolve whether real-time synthesis falls within the same provision.

Two findings do transfer to voice AI. The first is that human involvement in initiating the call is no answer: the scripts were triggered call by call by agents overseas, and the ICO still treated the calls as automated. The argument that a voicebot campaign is a live calling operation because a person starts or supervises each call is weakened accordingly.

The second concerns Regulation 24 and the concealment of the instigator's identity. That limb turns on how the call presents itself, not on how the audio is produced, so a synthetic voice introducing itself under a human name is exposed on identical terms.

It seems clear that an organisation running automated outbound marketing at scale is now exposed at a materially different order of magnitude from the one that applied when most outbound compliance policies were written: both penalties here were imposed under the previous £500,000 cap but the same conduct today would sit against a ceiling 35 times higher.

DISCLOSURE

There is no UK AI Act, and we are not aware of any general statutory requirement to disclose that a caller is speaking to a machine. We believe there are three reasons to do it anyway.

The first is that the ICO's avatar enforcement turned partly on the false impression of a human being given.

The second is that customers say they want it: as shown earlier, 65% of UK 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 organisations serving EU customers are subject to Article 50 of the EU AI Act, which requires people to be informed that they are interacting with AI and which applied from August 2026¹.

For firms regulated by the Financial Conduct Authority, the Consumer Duty applies to outbound as it does to everything else. An automated outbound call has to evidence good outcomes and appropriate treatment of vulnerable customers, and the evidencing burden does not reduce because the call was made by a machine.

¹ Source: Regulation (EU) 2024/1689, Article 50, <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>

A PRACTICAL CHECKLIST

- Establish whether each campaign is service or marketing, on content rather than label, and document the reasoning.
- For anything with a promotional element, treat conversational AI as an automated calling system and obtain prior specific consent.
- Record consent provenance against each number, including when and how it was obtained and what it covered.
- Screen against TPS and CTPS regardless, and maintain suppression lists that operate across campaigns rather than within them.
- Set a latency budget for the opening turn and test it under load. Dead air while the model generates its first response is capable of meeting Ofcom's definition of a silent call (its stated top enforcement priority), and the two-second reference point runs from the moment the recipient begins speaking.
- Disclose that the call is automated, early and in plain terms, and avoid giving the system a human name.
- Build opt-out handling into the conversational flow itself, not only into a menu at the end.
- Monitor latency at handover as well as at answer. Ofcom treats an unfairly long wait as misuse where the recipient has asked to be put through to a human, so an agent pool that is sized for pre-AI volumes becomes an exposure as soon as the bot increases the number of conversations reaching it.
- Present accurate and reachable calling line identification.
- Decide and document how long call audio and transcripts are retained, and on what lawful basis.

None of this is unique to AI, but what is new is the volume such systems make possible, and the fact that the penalty for getting it wrong is now considerably higher than it was.

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 organisations 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.

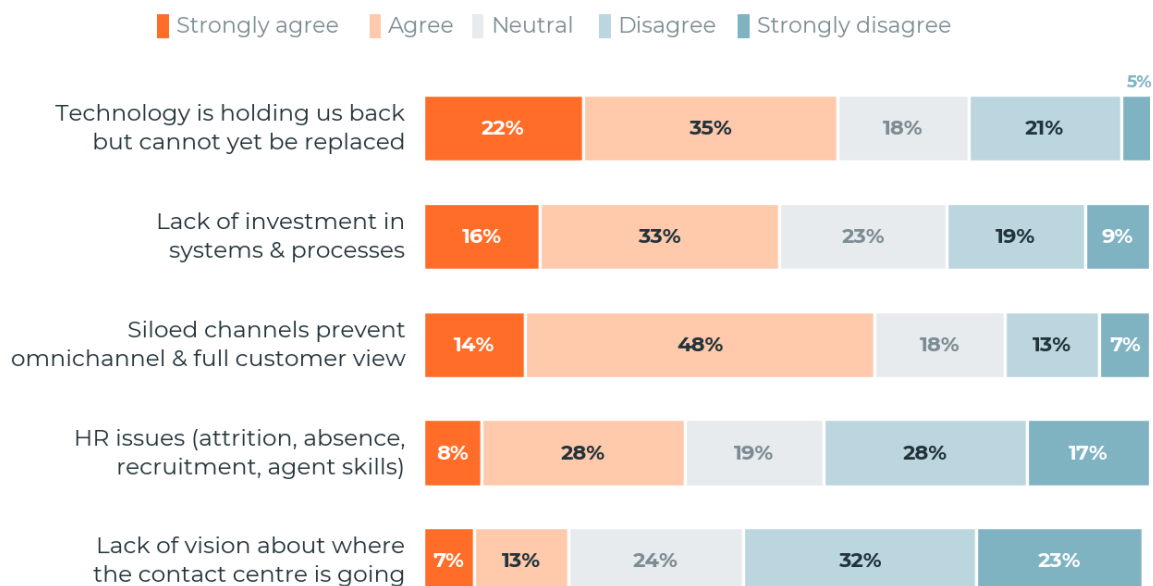
- **Non-linear speech.** Conversations do not follow linear paths. People repeat themselves, rephrase halfway through a sentence, and interrupt. A text interface never has to cope with any of that.
- **Latency.** Users of automated speech recognition have long been accustomed to pauses of several seconds 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 travellers, 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 organisations 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 CENTRES BACK GENERALLY

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

Figure 44: What is preventing the contact centre from achieving its aims?



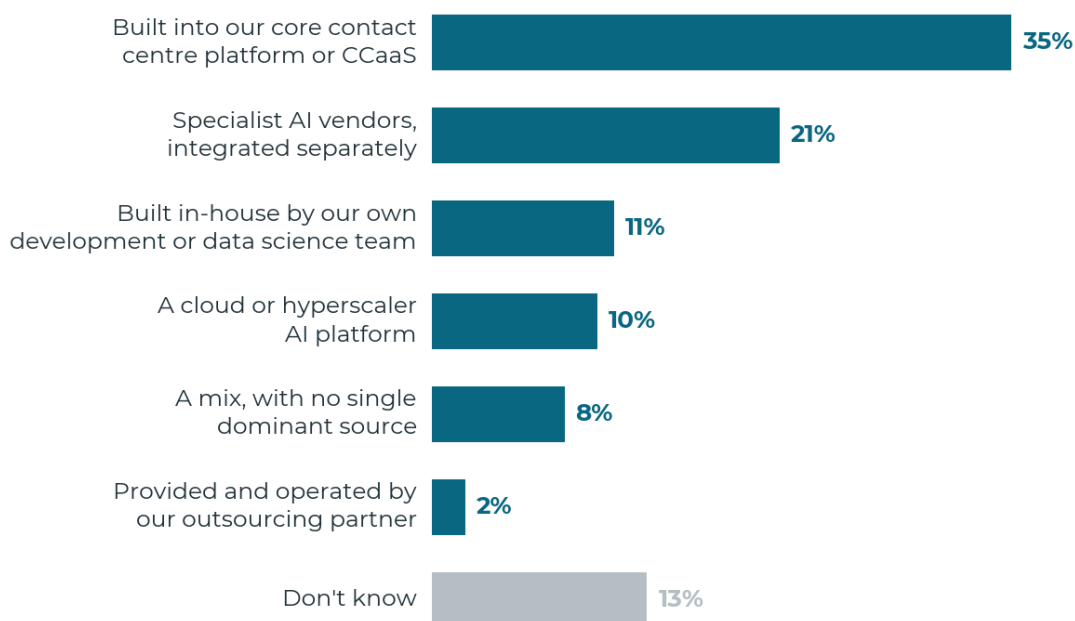
Siloed channels preventing omnichannel working and a full view of the customer is the most widely agreed constraint, at 62% agreement, followed by technology that is holding the operation back but cannot yet be replaced at 57% and lack of investment in systems and processes at 49%. Lack of vision about where the contact centre is going attracts the least agreement, at 20%, with 55% disagreeing.

The first two of those bear directly on voice AI. An organisation 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 45: Where the majority of AI capability in the contact centre comes from



35% source the majority of their AI from the core contact centre platform or CCaaS provider, which is the single largest category but well short of a majority. 21% use specialist AI vendors selected separately and integrated, 11% build in-house, 10% use a hyperscaler platform and 8% describe a mix with no dominant source. 13% do not know, which for a question about where an organisation's own technology comes from is itself informative.

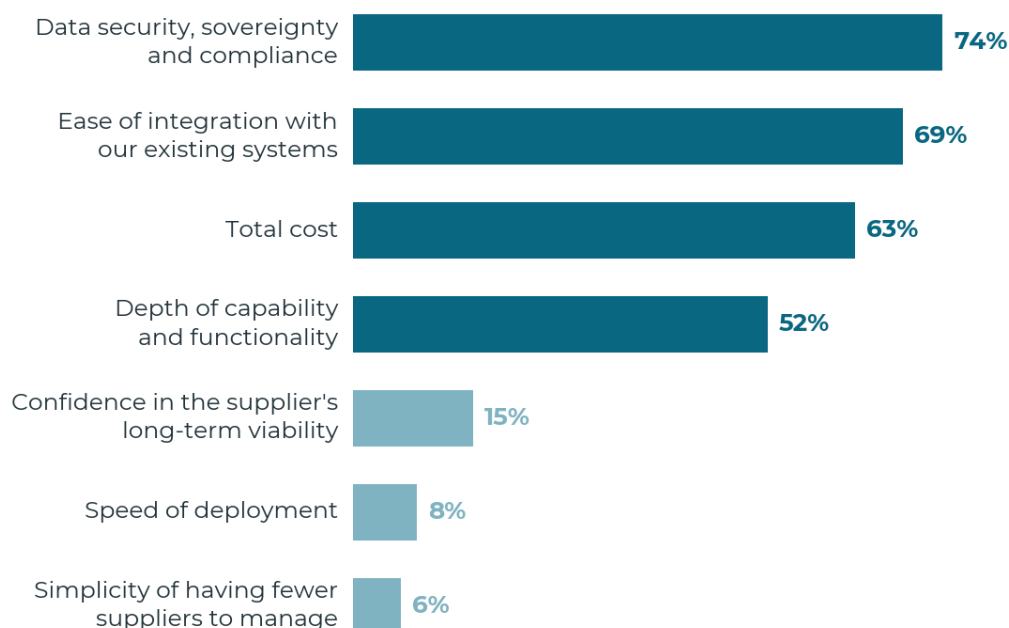
The forward view divides almost evenly three ways and does not resolve the question. 27% expect to move more towards specialist AI vendors, 27% expect to build more in-house, and 25% expect to consolidate onto their core platform provider. 8% expect no significant change and 13% do not know.

There is no consensus about where this market is going, and the three directions are mutually incompatible. A platform vendor reading these figures should note that only a quarter of the market intends to consolidate onto it; a specialist vendor should note that as many organisations intend to build the capability themselves as intend 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.

Data security, sovereignty and compliance leads at 74%, followed by ease of integration with existing systems at 69%, total cost at 63% and depth of capability at 52%. Those four account for almost the whole distribution.

Figure 46: What matters most when deciding where to source AI (top 3 answers)



Speed of deployment matters most to only 8% of buyers, and simplicity of having fewer suppliers to manage to 6%. “Deploy in days” is among the most common claims made for voice AI products, and it addresses a criterion that only one buyer in twelve treats as a priority. Confidence in the supplier’s long-term viability, at 15%, is similarly low given how young many of the vendors in this category are.

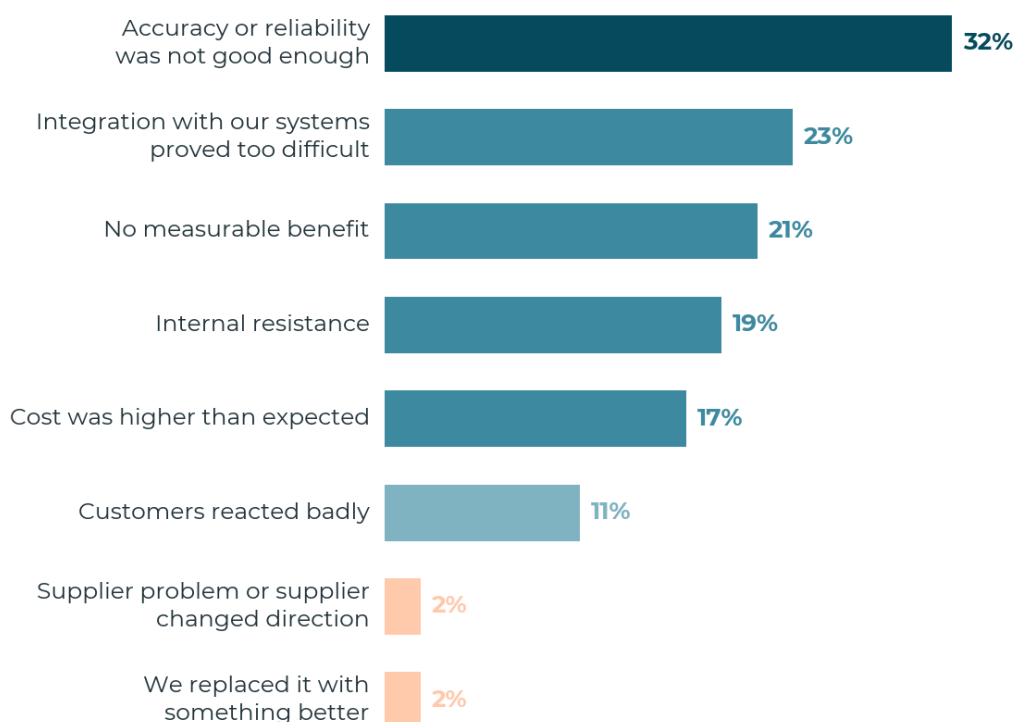
The practical implication for anyone building a business case is that a proposal which leads on speed and convenience is answering questions the buyer is not asking, while the questions they are asking, about data handling and integration, are the ones where evidence is hardest to produce.

WHAT GOT STOPPED

36% of respondents have stopped using, paused or rolled back an AI solution after deploying it. 53% have not and 11% do not know.

The stage at which those projects stopped matters more than the raw proportion. 59% stopped at a limited pilot and 27% at proof of concept or demonstration, which together describe evaluation working as intended: an organisation tried something, it did not work, and it was discontinued before significant commitment. 14% stopped in full production, which is a different matter.

Figure 47: Main reasons for stopping, pausing or rolling back an AI solution



Accuracy or reliability not being good enough is the leading reason at 32%, followed by integration proving too difficult at 23% and no measurable benefit at 21%. Internal resistance accounts for 19% and cost being higher than expected for 17%.

Two figures are worth drawing out. Customers reacting badly accounts for only 11%, which suggests that most failures happen before customers are exposed to them in any volume. Supplier problems, together with replacement by something better, account for 2% each, so almost none of these projects stopped because the market moved on. They stopped because the thing did not work well enough, could not be connected to anything, or could not be shown to have helped.

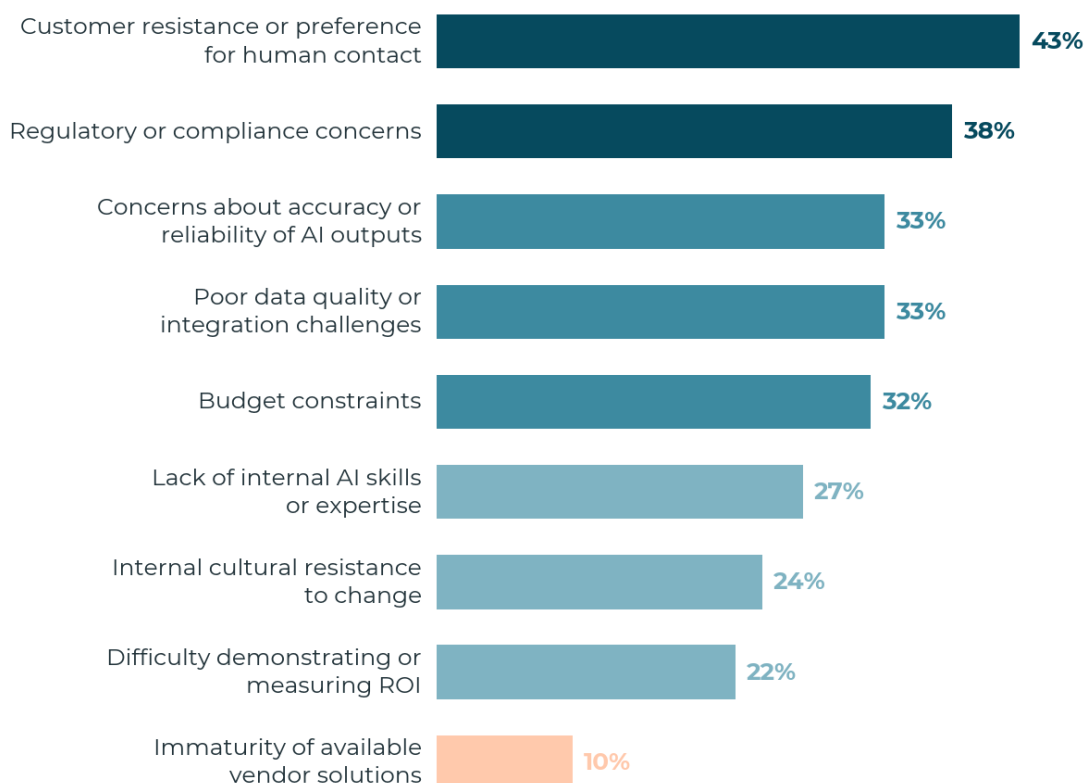
This is valuable data precisely because nobody has an incentive to publish it. Suppliers do not, and buyers rarely volunteer it. A 36% rate of abandonment, with accuracy the leading cause, is the appropriate context for any proposal that quotes a deployment timeline in days.

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 43%, ahead of regulatory or compliance concerns at 38%, and concerns about accuracy and poor data quality at 33% each, with budget constraints a similar level.

Figure 48: Most significant barriers to further AI adoption over the next two to three years



Customer resistance is a barrier on which this report has independent corroboration. 38% of UK consumers say they would never prefer AI to a human under any circumstances, 78% 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.

It is also the barrier least susceptible to the remedies usually proposed. Skills gaps at 27% and data quality at 33% 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 organisations 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 10%, 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	UK figure
Mean cost per inbound call	£6.17
Mean call duration	437 seconds
Calls requiring identity verification	74%
Mean agent authentication time	48 seconds
Calls affected by repetition (modelled)	30%
Time added per repetition event	15 seconds
Agent time on post-call wrap-up	12.5%
Average time spent on screen navigation in call	20%
Self-service containment where offered (mean)	16%
Self-service containment where offered (median)	10%

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 £0.0141, derived from £6.17 over a mean duration of 437 seconds.

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

Figure 49: Modelled saving per 1,000 inbound calls, by intervention (as examples only)

Intervention	Calls affected	Saving each	Per 1,000 calls
Double self-service containment, 16% to 32%	160	£6.17	£987
Automate 70% of post-call wrap-up (average cost of wrap-up is £1.09 per call)	all agent-handled	£0.76 (70% of £1.09)	up to £760
Automated authentication replaces agent ID&V	673	£0.68	up to £457
Remove repetition caused by poor audio	300	£0.21	£63
NB: 70% of post-call wrap-up and doubling containment to 32% are worked examples to demonstrate potential savings			

Containment is the largest prize, as a contained (and fully resolved) call avoids the whole £6.17 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. Doubling containment from 16% to 32% is worth roughly twice what removing agent authentication entirely is worth, and fifteen times what eliminating repetition is worth.

Two qualifications apply: automation is not free, so the £6.17 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. The second is that 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 up to £457 from authentication is more likely: the task is mechanical and the customer has no strong preference on who performs it. The cost of removing repetition caused by poor audio should also be viewed in the context of the much wider ongoing benefits accruing across the board that having a clean and accurate transcription available for other AI applications to use will bring.

The £987 from doubling containment is less certain: the median operation currently contains 10% of voice calls (the 16% figure is the mean), and the current top quartile achieves only 23%.

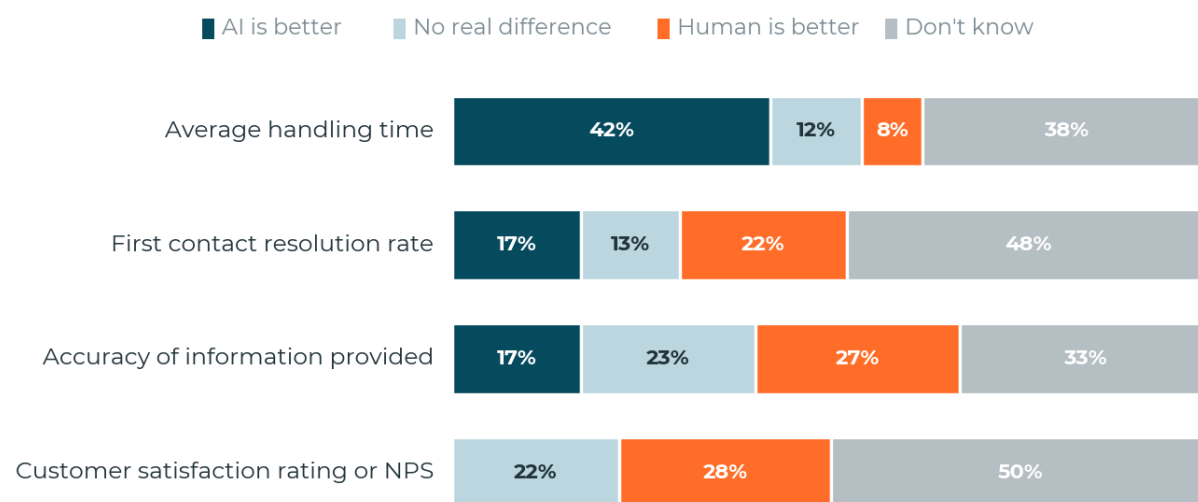
38% of consumers say they would never prefer AI under any circumstances, and the leading reasons for abandonment are trust at 66% and functionality at 64% rather than anything a better voice interface fixes. A target of 32% would put an organisation above the present top quartile: perhaps achievable, but by no means guaranteed.

PERFORMANCE AGAINST EXPECTATIONS

Looking at the outcomes of AI, average handling time is the one measure on which AI is judged clearly superior. 42% say AI is better, 8% say the human is better and 12% see no real difference.

On every other measure the balance runs the other way. For first contact resolution 22% say the human is better against 17% for AI; for accuracy of information 27% against 17%; and for customer satisfaction or NPS 28% say the human is better while not one respondent said AI was.

Figure 50: Differences in performance between AI-handled and human-handled interactions



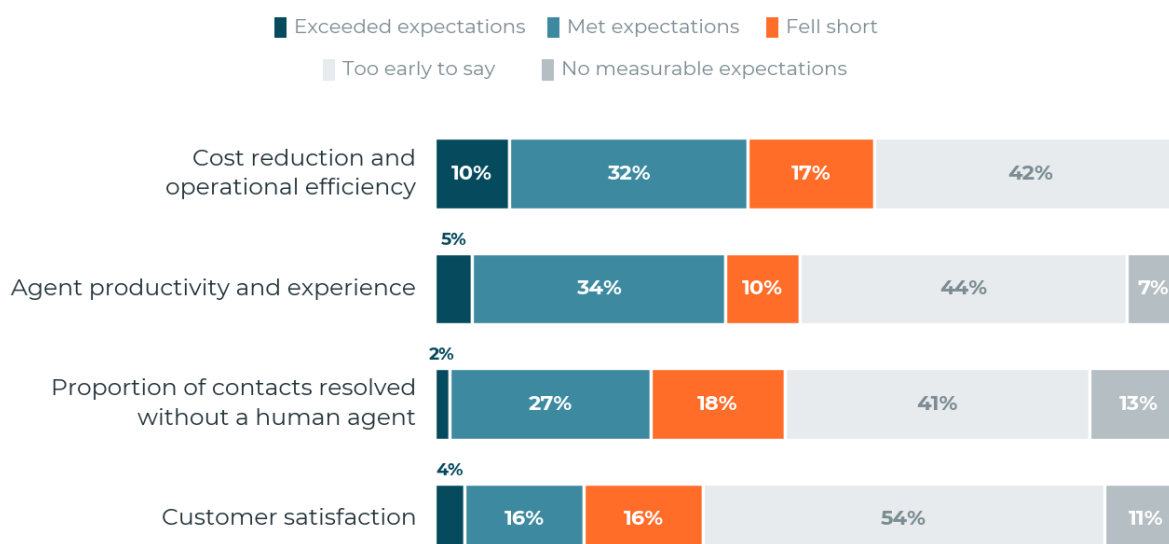
That is the same shape as the consumer data seen earlier, where customers rate AI worse than a human on five of six dimensions and better only on speed. Businesses and their customers are describing the same thing from opposite sides of the call, and the agreement between two independent samples is a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

The don't-know column is the other key finding. It runs from 33% on accuracy to 50% on customer satisfaction. Half the market cannot say whether its AI produces better or worse satisfaction than its people for comparable contacts, which is a glaring gap.

When considering performance against expectations, cost reduction and operational efficiency is where AI most often meets what was promised: 42% report it met or exceeded expectations, against 17% falling short.

Agent productivity and experience is close behind at 39% met or exceeded. Customer satisfaction is the weakest, at 19% met or exceeded and 16% falling short.

Figure 51: Performance against the expectations set in the original business case



“Too early to say” dominates every row, from 42% to 54%, which is consistent with a market in which most deployments are recent.

Between 7% and 13% of organisations set no measurable expectations at all for three of the four outcomes, and the highest of those is for the proportion of contacts resolved without a human, at 13%. That is the outcome most often used to justify the investment in the first place, and one organisation in eight committed to it without defining what success would look like.

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

No respondent expects their contact centre AI budget to fall over the next two years. 34% expect an increase of up to 25%, 20% an increase of 26% to 50%, and 17% an increase of more than 50%.

14% expect it to stay roughly the same and 14% do not know. A market in which 71% of buyers expect to spend more and none expects to spend less is not one that is reconsidering the technology, whatever the rollback figures in the previous section suggest about individual projects.

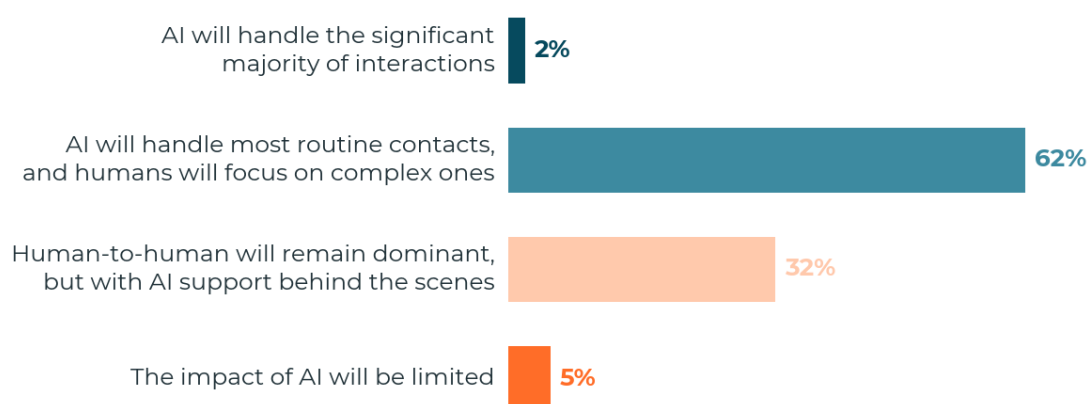
Our survey respondents were asked to consider what the future held for AI in contact centres.

Only 2% expect AI to handle the significant majority of interactions including complex ones, with 62% believing that AI will handle most routine contacts with humans focusing on complex ones.

32% expect human-to-human interaction to remain dominant with AI supporting behind the scenes.

5% expect the impact of AI to be limited on their operations.

Figure 52: Which statement is closest to the expected position in five years



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

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

The 62% and the 32% are separated less by disagreement about capability than by where each organisation 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 maximising.

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 organisation 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 £63 per 1,000 calls, but that figure massively understates its importance because speech-to-text accuracy is what every layer above it consumes.

57% of speech recognition users already say accuracy puts customers off, and 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 summarisation is the second largest prize at up to £760 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 £457 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 74% of calls require that. Organisations already confident in their identity verification will find personalised self-service easier to build than those that are not, though the two can perfectly well be developed in parallel.

Analytics and automated scoring address the 90% of operations reporting insufficient time to analyse their data, and they are also how an organisation 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 organisations, 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 69% of consumers rate AI worse than a human at understanding the issue.

To summarise: 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 organisation which has attended to them will get more from the same investment than one which has not.

ABOUT CONTACTBABEL

ContactBabel is the contact centre 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 centres 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 analysing the contact centre 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 Centre Solutions
 - Customer Engagement & Personalisation
 - Customer Interaction Analytics
 - First-Contact Resolution
 - Fraud Reduction & PCI Compliance
 - Omnichannel Workforce Optimisation
 - Remote & Hybrid Working Contact Centre 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 centres' 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.