CPaaSAA | Research

CPaaS Acceleration Alliance

CPaaSAA Research is produced by, and represents the views of, the CPaaSAA team

AI Voice: Who Will Run The Conversation?

AI Voice is creating fast growing and valuable new business opportunities. How can Intelligent Engagement players maximise their value?

ABSTRACT

AI Voice unlocks the "dark data" in calls and extends the ambition of Intelligent Engagement: coordinated, real-time interaction across customers, partners and workforce. Early deployments already show measurable business impact, including operational time savings of 30% or more and booking or conversion improvements of up to threefold in some AI-assisted interactions.

As conversations become computable, value shifts away from call volume and towards what is understood, decided and acted on, and the profit pool in communications will shift from carrying interactions to orchestrating the intelligence that runs them. Governing how conversations are executed, i.e. how data flows, how models are applied, and how outcomes are delivered, will shape where value accumulates in the next phase of enterprise communications. Connectivity remains essential, but the value of communications is shifting from transport to intelligence.

This report draws on CPaaSAA's case directory, industry interviews, and secondary research, providing an evidence-based view of how enterprises are beginning to deploy AI Voice. It provides an overview of the emerging AI Voice market and examines the strategic choices facing Intelligent Engagement players as conversational AI becomes embedded across communications platforms, enterprise workflows and telecom networks.

Published: March 2026

CONTENTS

CONTENTS **2**
ABOUT THIS REPORT **4**
THE CPAAS ACCELERATION ALLIANCE 5
HOW TO READ THIS REPORT 5
EXECUTIVE SUMMARY **6**
KEY ACTIONS FOR INTELLIGENT ENGAGEMENT PLAYERS 7
CUSTOMER ADOPTION AND RISKS 8
FOR ECOSYSTEM STAKEHOLDERS: SO WHAT? 9
CPAASAA's ACTIONS 11
INTRODUCTION TO AI VOICE **12**
WHAT IS AI VOICE? 13
AI VOICE IS A COMPONENT OF INTELLIGENT ENGAGEMENT 15
AI Voice and Conversational Intelligence 17
Calls and Voice in Context 18
WHY AI VOICE MATTERS TO -aaS PLAYERS & CSPS 19
THE BUSINESS CASE FOR AI VOICE **20**
WHAT CAN AI DO FOR VOICE, AND VOICE FOR AI? 20
AI VOICE IMPLEMENTATION: CASE STUDY EXAMPLES 21
Transcription 22
AI Dialogue: AI Voice becomes a two-way communication 22
Appointments and Scheduling 23
Lead Qualification 23
Streamlining Processes 24
Customer Experience 24
Compliance 25
Case Study Distribution 26
Workforce impact of AI Voice adoption 27
Beyond the call itself 27
BUSINESS BENEFITS OF AI VOICE 28
CUSTOMERS AND SALES CHANNELS: MINDSETS & BUDGETS 30
ETHICS, PRIVACY AND REGULATORY COMPLIANCE 32
WHEN WILL THE BENEFITS MATERIALISE? 33
INDUSTRY ECOSYSTEM & DYNAMICS **34**
AI VOICE ECOSYSTEM 34
RUNTIME: THE CONTROL POINT FOR VALUE 36
Recent Market Signals 37
Compute Location vs Runtime Ownership 37
FOUR RUNTIME SCENARIOS: HOW BUSINESSES DECIDE 38
Intelligent Engagement Led Runtime (CPaaS / CCaaS) 39
Hyperscaler-Led Runtime 40
Telco-Integrated Runtime 41

Enterprise-Controlled Runtime 42
COMPETITIVE RISKS 43
SOVEREIGNTY AND STRUCTURAL TENSION 44
EMERGING REGULATORY QUESTIONS FOR AI VOICE 44
HYPERSCALER STRATEGY: COMPONENT VS. PLATFORM 45
APPLICATIONS: MANY HORSES FOR MANY COURSES 46
AI VOICE ADOPTION & FORECASTS 48
AI VOICE ADOPTION 48
MARKET SIZE, GROWTH AND FORECASTS 51
THE FOUR VALUE LAYERS OF AI VOICE 52
Strategic implications 52
How might it play out, in general? 52
INTELLIGENT ENGAGEMENT: MARKET OPPORTUNITIES 53
WHAT'S DRIVING CUSTOMER ADOPTION TODAY? 53
BUSINESS SEGMENT NEEDS 53
Vertical Specialisation: Creating Defensible Utility 54
INDUSTRY NARRATIVE AND PACKAGING 55
PLATFORM CHOICES & MODEL NEUTRALITY 56
TELCO NETWORK AI VOICE 57
AI in every call 57
Design and pricing considerations 59
Why it matters where the intelligence is orchestrated in Telco Network AI Voice 60
Telco Call Centre AI 60
ANALYTICS 61
VCONS: CONTEXT, COMPLIANCE AND COMPETITIVE ADVANTAGE 61
TRANSLATION 65
BIOMETRICS & SENTIMENT ANALYSIS 66
CONCLUSIONS 68
INTELLIGENT ENGAGEMENT PLAYERS IN AN AI VOICE WORLD 68
STRATEGIC IMPERATIVES FOR AI VOICE SUCCESS 69
THE BOTTOM LINE 70
A SINCERE NOTE OF THANKS 71

ABOUT THIS REPORT

The analysis and the views in this report are the independent work of the CPaaS Acceleration Association (CPaaSAA), and we would like to thank all of our members for their significant input and support.

The report was researched and written in late 2025 and early 2026 by Andrew Collinson, Associate Research Director, CPaaSAA and reviewed by CPaaSAA's Partners.

A Note of Thanks

This report would not be possible without the insight and collaboration of CPaaSAA members and contributors across the ecosystem. We are grateful to the many executives, product leaders and practitioners who shared perspectives, case studies and data. Their input strengthens our shared goal: accelerating Intelligent Engagement and ensuring this industry captures the value it helps create.

We are grateful for the additional support provided by our Accelerate members: Infobip, KPN, Radisys, Sinch and Vonage.

There is a further list of 36 people from these and other organisations here who contributed ideas, insights and information, to whom we extend our sincere thanks.

- The report is also based on our analysis of 46 AI Voice customer case studies developed by numerous members and ecosystem players, which show real-life applications of the technology and list customer accredited benefits. Some of those cases are summarised in this report. They are among the 163 case studies in our Case Directory at March 2026.
- All interactions with CPaaSAA's members were also taken into consideration, including our Network API and AI & Data Working Group meetings, and dedicated briefings and meetings.
- We also carried out extensive secondary research, and cite the findings of a range of analyst companies, press releases and other market articles. We used several LLMs to assist in this, including different instances of ChatGPT, Claude and Gemini, and cross-matched our findings where possible.
- All of the diagrams and most of the copy in this report were created by the author. Where AI-Model based research copy was used it was manually reviewed and re-written by the author and reviewed by at least one other CPaaSAA partner.

While this report represents the independent view of the CPaaSAA team with the goal of uniting and catalysing the industry to faster growth, we hope it does justice to the vision and efforts of all of our members, and we look forward to feedback and further discussions.

THE CPAAS ACCELERATION ALLIANCE

The CPaaS Acceleration Alliance (CPaaSAA) is the global Alliance for intelligent engagement, connecting innovators shaping the future of communications through AI, CPaaS, Network APIs, and real-time platforms.

CPaaSAA's members don't just follow trends – they shape them. From telcos and tech leaders to platform builders and scaleups, we unite people redefining how businesses engage with customers.

We help our 100+ members learn from peers, share what works, and lead the next wave of transformation. Through curated events, practical frameworks, and deep industry partnerships, we turn conversations into co-creation – and visions into ventures. Our members span traditional CPaaS players, CCaaS and UCaaS companies, telcos, techcos, analysts, consultants and investors.

CPaaSAA Research

We publish and provide research services, including:

- The Service Provider Playbook - a 'how to' guide for service providers offering CPaaS
- The CPaaSAA Case Directory is a resource bringing together industry case studies to help accelerate the market, examples of which are shown in this report
- Reports, e.g.
    - Intelligent Engagement: How AI, APIs & CPaaS Change Enterprise Strategy
    - The State of CPaaS 2025
- CPaaSAA Working Group White Papers

HOW TO READ THIS REPORT

- **High level 'skim and dip' readers**: the next section of the report is the Executive Summary, highlighting key actions and recommendations, with references and links to where you can find the supporting details.
- **End-to-End readers:** the report is written so that it can be read from end to end.
- The full **table of contents** is on page 2.

EXECUTIVE SUMMARY

AI VOICE: FAST GROWTH AND STRONG BENEFITS

AI Voice unlocks the "dark data" in calls, and provides new ways to engage with customers in the vision of Intelligent Engagement - the business aim to interact with customers in real-time with intelligence at scale. Early deployments already show measurable business impact, including 30%+ operational time savings, booking and conversion improvements of up to 3×, with further gains expected as AI analytics and automation are applied to conversation data.

- **The market is growing fast.** CPaaSAA members report 4-10X year-on-year growth in 2025-26 in enterprise demand, and analysts forecast five-year CAGRs of 16-38%. Interest is also newly strong from telecoms operators in "AI in every call" - as evidenced at MWC26 by Deutsche Telekom and LG Plus and others (see Telco Network AI Voice section). AI Voice presents new category opportunities for Telecoms and Intelligent Engagement players.

- **Market Potential**. Current total "Conversational AI" market estimates are order of magnitude ~$10bn. AI Voice sits within this, and in value pools exceeding $300bn across customer care, CRM and digital engagement. Even modest participation in these adjacent pools implies revenue potential measured in $ billions / tens of billions. We anticipate new market data and forecasts in 2026 as analysts catch up with the rate of change of the technologies.

- **Recent Market Signals.** Announcements at Enterprise Connect 2026 further illustrate the acceleration of AI Voice within communications platforms. Vendors including RingCentral and Sinch demonstrated AI agents embedded directly into communications infrastructure, enabling enterprises to deploy conversational automation across voice, messaging and digital channels.

- **Early adoption of AI Voice shows clear business benefits.** Transcription and copilots are now mainstream in advanced service organisations, save time (often 30%+) and improve quality. Customer evidenced benefits include facilitating scaling, first-time resolution, improving automation containment, and cost reduction. AI Voice demonstrates lead conversion and improved booking rates uplift up to 3X, and growing ability to address basic transactions. Compliance monitoring can extend from sampling to 100% of interactions. (See business benefits section for details)

- **Analytics are less mature in adoption** but combined with effective governance and automation tools and processes (e.g. vCons) show significant potential to deliver value.

- As voice becomes computable, value shifts from the length of an interaction to what is understood, governed and achieved. In effect, AI Voice shifts communications economics from transport pricing toward the control of data, intelligence and execution. **The strategic question is** no longer who carries the call, but **who controls the runtime, which orchestrates the data flows in the conversation**, and therefore who captures the value.

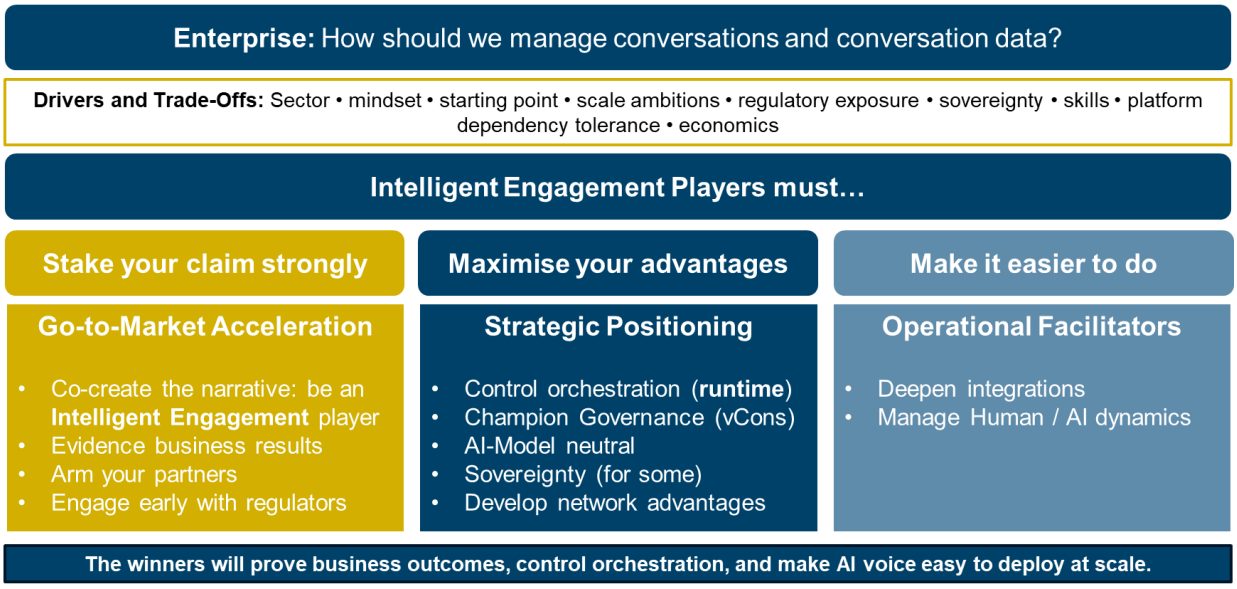
- **Intelligent Engagement players** (CPaaS, CCaaS, UCaaS +) are part of an ecosystem including hyperscalers, telcos and IT players, many of whom are fighting for the prize of a stake in this market.

This report outlines the market, customer and ecosystem drivers, and is intended to help Intelligent Engagement players, enterprises and investors maximise value in AI Voice.

KEY ACTIONS FOR INTELLIGENT ENGAGEMENT PLAYERS

Strategic Priorities for Intelligent Engagement Players in AI Voice

Enterprise choices about governance, sovereignty and platform control shape how the market will evolve.



Source: CPaaSAA

1. **Co-create the narrative: position yourself as an Intelligent Engagement player.** CPaaS/CCaaS/UCaaS tags do not intuitively convey the ability to support customers with AI Voice or any other intelligence based engagement. If you continue to use this language, you will become further marginalised in customers' minds. You need to drive Intelligent Engagement and describe yourself as an Intelligent Engagement player. Position AI Voice within Intelligent Engagement, refine a common narrative structure to describe component propositions, and support with clear outcome focused case studies.

2. **Evidence business results.** Move the conversation from technology features to business results. Focus on benefits of unlocking scale, reducing handling time, improving conversion and scale economics. Product strategy should address budget scrutiny and ROI directly. Shift pricing towards outcomes based when feasible.

3. **Arm your partners** (and educate your/their customers). AI Voice is a new and fast developing field, and sales teams that aren't incentivised and aren't ready to deliver concise, compelling explanations will certainly not be able to sell it. Equally, customers need to know "what can I do with AI?" If you don't help them, someone else will.

4. **Engage early with emerging regulation.** AI Voice may sit at the boundary between AI-generated "content" and electronic communications services. The way regulators classify these interactions could materially affect how services are designed, disclosed and governed. Intelligent Engagement players should actively engage in emerging regulatory processes – including the EU AI Act transparency Code of Practice – to ensure that real-time communications use cases are properly understood and that workable standards for disclosure, marking and governance are developed.

5. **Control orchestration (the runtime layer).** The orchestration layer governs routing, escalation, policy and model execution. Control here determines who captures intelligence margin and who remains a transport provider. Focus your messaging and product design to show customers the strengths that partnering with you as an orchestrator delivers (i.e. points 6-11 below).

6. **Champion governance and trust.** Mission-critical conversations require reliability, auditability and privacy by design. Embed compliance, data sovereignty and policy control into the core platform and your values. Governance can be a source of differentiation, not overhead. vCons (a new data standard) are a potential asset for Intelligent Engagement players in this respect.

7. **Remain AI-Model neutral.** Enterprises need flexibility in a fast-moving model market. Platforms that allow models to be swapped without architectural disruption become the stabilising layer in enterprise AI Voice deployments.

8. **Address Sovereignty requirements.** In some markets, broadly most outside of the US and China, national-level data sovereignty is an increasingly pressing concern. If these are the markets you serve, ensure that your proposition clearly addresses these needs. Even within the US, sovereignty is an issue at a company level. Customers want to be in control of their options, and be able to implement sovereignty if and when required.

9. **Leverage network advantage where relevant.** For CSPs, integrating AI Voice with Network APIs and network-native capabilities can create scalable, high-quality services that are difficult for pure software providers to replicate.

10. **Deepen enterprise integration.** Embed AI Voice into CRM, workforce management, identity and compliance systems. Structural integration increases switching costs and strengthens long-term relevance.

11. **Manage human adoption deliberately.** AI Voice succeeds faster when it helps people rather than threatens them. Prioritise copilots, automation of repetitive work and tools that help teams perform better. Adoption friction is a commercial risk.

Together, these actions determine whether Intelligent Engagement players will shape the next phase of enterprise conversation or operate beneath it.

CUSTOMER ADOPTION AND RISKS

- AI Voice Adoption overall remains uneven but is accelerating. Snapshot indicators suggest meaningful deployment in larger service enterprises, with maturity still early-stage. Forecasts across adjacent segments and recent AI Voice performance announcements show promisingly strong growth.

- The emergence of network-native AI Voice products from operators suggests that it is no longer a purely theoretical debate about where AI Voice value may sit. The market is beginning to test whether AI capabilities can live directly within the communications layer itself, making the question of orchestration and execution control increasingly important.

- Competitive risk. Hyperscalers are expanding from component supply to full-stack platforms. CRM and workflow providers are embedding conversational intelligence directly into systems of record. Contact centre incumbents are strengthening AI control within their environments. Vertical AI specialists are building domain-specific runtime capabilities in regulated industries.

- If Intelligent Engagement players remain focused on transport and usage, margin will migrate to other players. Where they can control orchestration, governance and structured data, they can grow faster. AI Voice will redistribute value across the stack, and runtime is the contested control point.

FOR ECOSYSTEM STAKEHOLDERS: SO WHAT?

- **CPaaS, CCaaS and UCaaS providers: become Intelligent Engagement Players**. AI Voice is a strategic identity test. If you continue to think and describe yourselves in terms of SMS volumes, minutes and API throughput, customers will position you accordingly - as infrastructure suppliers. In that framing, orchestration and runtime arguments will not land. The shift to Intelligent Engagement is not cosmetic but the move from selling transport to shaping outcomes. That requires a new narrative, new commercial language and new proof points built around automation, performance improvement and governed execution. Without that repositioning, value will migrate upward regardless of your technical capabilities. Platform providers must compete to control the execution layer for conversations and integrate deeply with enterprise workflows, as CRM and enterprise software platforms are increasingly positioning themselves as alternative orchestration environments.

- **Telecom Operators: sell AI voice solutions and play to positional strengths**. Telcos can build, resell and integrate enterprise AI Voice solutions, and potentially embed AI directly into everyday calls. Network-level AI can enhance service quality, identity and trust. The opportunity for operators is therefore not simply to add AI features to calls, but to define a trusted execution environment for "network native AI voice". Where identity, quality of service, resilience and sovereignty matter, the communications layer itself may become a natural home for AI orchestration. If orchestration and policy remain outside the network, AI will improve the service but limit operators' control of the economics. Options include embedding runtime capabilities in the network, integrating Network APIs (for example identity validation in sensitive calls), and aligning AI Voice with broader sovereignty and resilience agendas.

- **Hyperscalers: embrace Intelligent Engagement in AI and cloud architectures.** AI Voice can extend cloud platforms from component supply into live operational control. The opportunity is to standardise orchestration within broader AI stacks and deepen enterprise dependency through data and workflow integration. The risk is regulatory, sovereignty and governance pushback in sensitive sectors. Component excellence alone will not secure durable control without operational accountability.

- **CRM and Workflow Platforms: capture the outcomes.** Platforms such as Salesforce and ServiceNow sit where customer interactions become business outcomes. If conversation intelligence flows directly into these systems, they can shape how insights from AI Voice translate into actions such as sales follow-ups, service resolution or operational workflows. Their opportunity is to embed conversational signals into the system of record. Intelligent Engagement players should therefore ensure that conversations executed through their platforms generate structured data and trigger workflows within these enterprise systems.

- **Enterprise Buyers: think through all the outcomes you want.** The decision should not just be about which model performs best, it must include where the controls of execution sit. That choice determines switching costs, compliance exposure, data ownership and long-term flexibility. AI Voice should be evaluated as part of a broader engagement strategy, not as an isolated automation tool.

- **Investors and Ecosystem Partners: check the Intelligence in the Engagement.** The margin pool is moving above transport into intelligence, outcomes and structured data. Players that control runtime and embed deeply in enterprise workflows will capture disproportionate value. Those confined solely to usage-based economics face compression as AI improves efficiency.
- **Regulators and policy makers: clarify the boundary between AI content and communications.** AI Voice increasingly blurs the distinction between synthetic content and real-time communications services. Applications such as live translation, AI call assistants or automated call screening may not fit neatly into regulatory frameworks designed for AI-generated media. Clear guidance will be needed on how transparency, disclosure and governance rules apply to real-time voice interactions. In particular, regulatory treatment may depend on where the intelligence sits in the architecture and who controls the AI model generating the speech.

The Strategic Shift: In AI Voice, the strategic question is no longer who carries the call, but who runs the conversation, and therefore captures the value.

CPAASAA's ACTIONS

To support the industry opportunities and actions, the CPaaS Acceleration Alliance (CPaaSAA) will promote and develop Intelligent Engagement, and continue to prove and develop the value of CPaaS and other applications through:

- Industry events and collaboration. including:
    - CASA26, CPaaSAA Talks, Inner Circle Webinars, and other events and hackathons
    - Working Groups including Telco API and AI & Data
    - CAVE Visionary Exchange Leadership Group
    - Developing and sharing industry research and Insight
    - Intelligent Engagement Case Directory
    - Service Provider Playbook
    - State of CPaaS 2025
    - Plus supporting research reports

We will gather feedback from response to this report and:

- Target an AI Voice opportunity sizing review in the State of CPaaS 2026, to be published at CASA 26 in September
- Revisit this report to create a follow up in late 2026 or early 2027

INTRODUCTION TO AI VOICE

AI Voice is not just a new feature of communications platforms, it is part of the transition toward Intelligent Engagement.

This report is an introduction to the AI Voice market and the possible opportunities it offers Intelligent Engagement players. AI Voice has rapidly moved from theory to practical industry strategy as AI capabilities are now being embedded directly into communications platforms and telecom networks.

In the report we distinguish between three related concepts.

- **CPaaS and CCaaS** describe product categories within the communications software market.
- **Communications platforms** refers more broadly to systems that operate real-time interactions such as calls, messaging and collaboration.
- **Intelligent Engagement** describes the strategic evolution of these and other systems when communications, AI and enterprise workflows are combined to deliver measurable business outcomes.

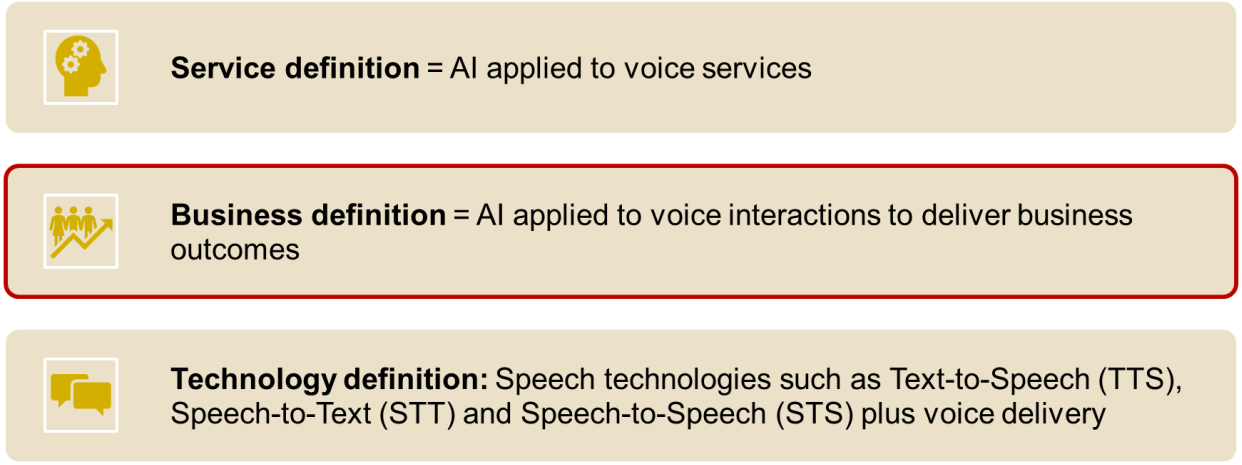
A key thrust of our analysis is that Intelligent Engagement is where the value is moving to, and the market positioning towards which players that wish to compete for that value should also move, supported by the orchestration of conversations, data and workflows rather than communications transport alone.

AI Voice is evolving rapidly. Given this rapid evolution, CPaaSAA plans to update this research within the coming months.

WHAT IS AI VOICE?

Three Ways to Define AI Voice

AI Voice has different meanings depending on whether the focus is technology, services or business outcomes.



Source: CPaaSAA

There are three common definitions of AI Voice, which tend to be used by slightly different audiences. The definition we focus on at CPaaSAA is **the combination of AI and voice in various forms to provide business benefits.**

This is what most of our members and their customers mean when they talk about AI Voice. This is our main focus in this report as we seek to understand how to help customers apply and monetise new technologies, and there is a broad range of applications.

Many, *but not all,* of the solutions we examine are used within the businesses' IT and communications setups. They are designed to be used on calls with specific parts of the organisation (e.g. a specific helpline), and are bought and built in by the company.

A relatively new class of Telco Native AI Voice applications is emerging that is built into the telecoms operators' networks, rather than the businesses' IT setups. An example of this is the recently announced T-Mobile *87 translation service, where customers can initiate a live translation service by keying *87 during a call.

"What's new is the form of access - we can now bring AI into every call on the telecoms network, not just the calls to a specific business number."

Ralph Page, EMEA Markets, Radisys

Such services are not necessarily confined to business contexts. Consumers and employees alike may use capabilities such as translation, transcription or personal assistants during everyday calls. However, these same capabilities can increasingly integrate with enterprise IT systems, creating new opportunities for telecoms companies to support business workflows.

Customer service and contact centre automation remain important entry points for AI Voice. Over time, however, the opportunity extends to AI-native communications across everyday voice interactions, including assistance, translation, coordination and action-taking in live calls. In this model, AI Voice becomes a general interaction layer rather than a specialised service channel.

This report discusses these opportunities in brief in the section on Telco Network AI Voice, and we will extend our coverage in future research, including how the different types and models of AI Voice services may interact. For example, what happens when two transcripts on a key contract negotiation or sensitive customer care call do not align?

Audiences that are more focused on the development of technologies, such as LLM makers and deep "technology component" makers, sometimes use AI Voice more narrowly, to mean the technologies, processes and systems used for:

- Text to Speech (TTS - voice generation)
- Speech to Text (STT - transcription)
- and Speech to Speech (STS - advanced AI systems that convert audible components of speech into meaning without a middle text stage).

NB Voice AI is interchangeable in meaning with AI Voice.

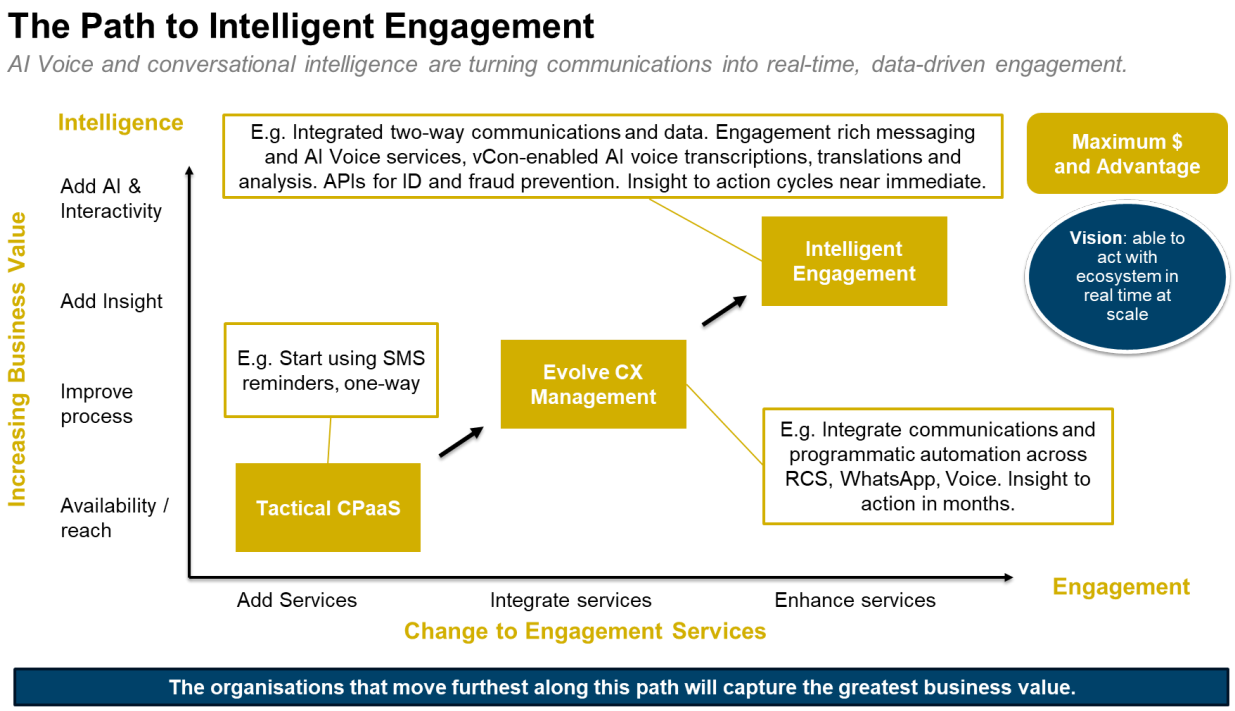
AI VOICE IS A COMPONENT OF INTELLIGENT ENGAGEMENT

AI Voice becomes valuable when integrated with AI, analytics, APIs and enterprise workflows to deliver Intelligent Engagement.

Intelligent Engagement is a new business strategy driving adoption of CPaaS, AI, and communications and network APIs.

Evolving from tactical, piecemeal adoption of engagement services (such as bulk SMS, voice, etc.) and combined with intelligence from AI and network (and other) APIs, businesses are moving towards a goal of acting and responding in near real-time to market changes with their customers, partners, people and things.

Intelligent Engagement was first outlined in the CPaaSAA Research report Intelligent Engagement: How AI, APIs & CPaaS Change Enterprise Strategy.



Source: CPaaSAA

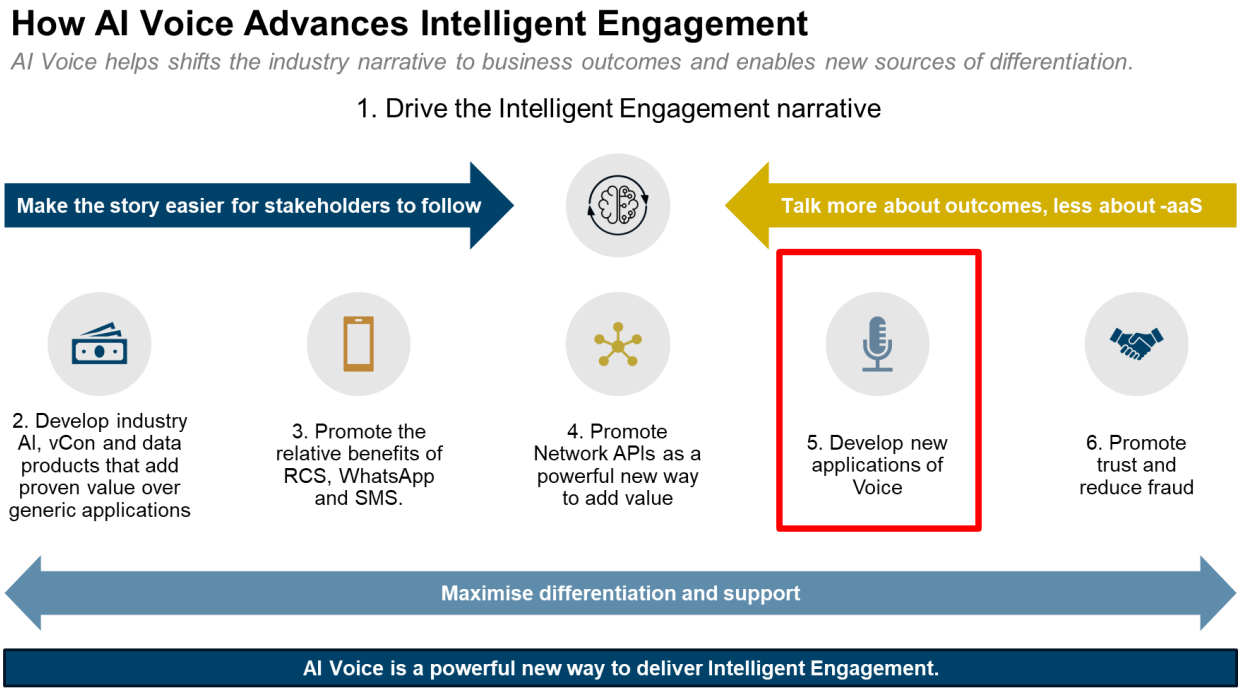
AI Voice is a new component of the mix that is helping to drive and reshape the nature of Intelligent Engagement. AI Voice comprises various applications of AI and Voice to allow:

- Human voice conversations to be turned into data (transcripted to text or other meaningful symbolic tokens)
- Text to be turned into machine speech (voice generation) so that the AI or processes can "speak"
- The data to be used to inform and drive decisions or actions, including translation, biometrics, decision support, automation, etc.

The two primary business drivers of change in AI Voice (as opposed to AI) are:

- AI can now talk with customers in ways and at times that were previously impossible or uneconomic (e.g.s. out of hours calls, with automatic machine-speed access to data)
- It opens up to analytics and machine processing, a vast new trove of information in previously inaccessible voice calls

In our report on the State of CPaaS 2025, we identified Voice, and AI Voice in particular, as a new field of innovation and progress, evidenced by numerous and diverse Case Studies in our case directory of examples of member successes. This report expands on those observations.



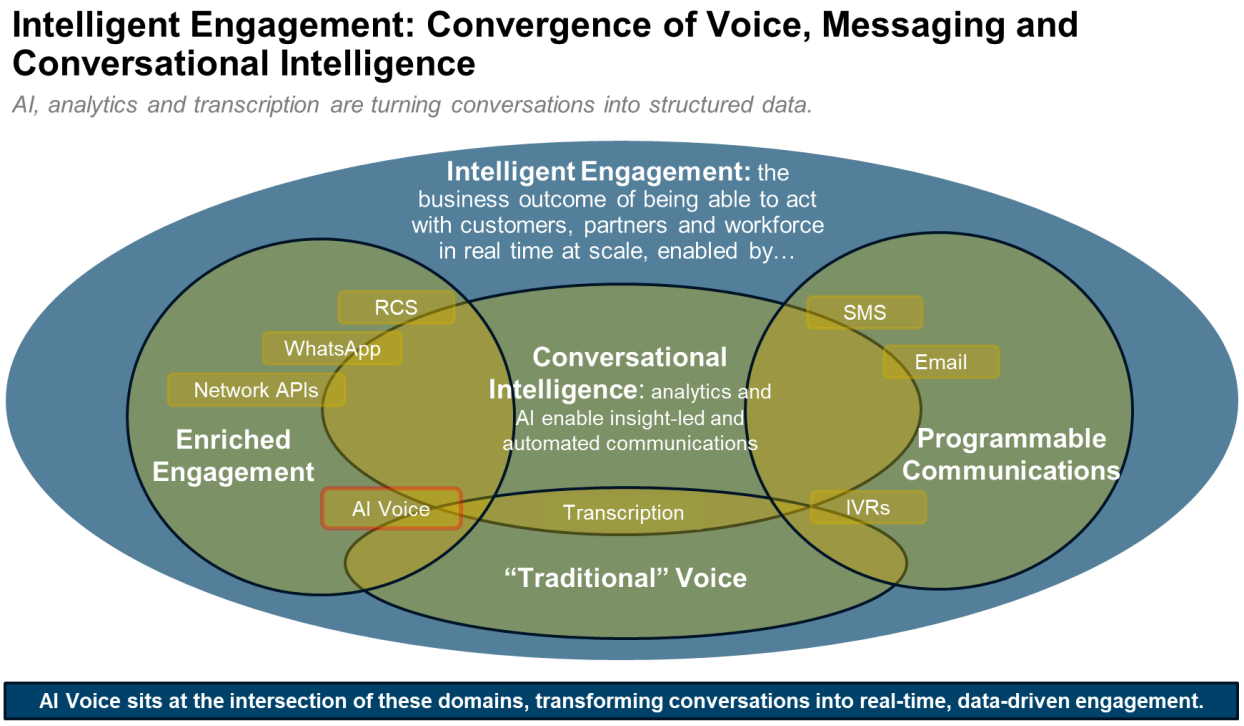
Source: CPaaSAA, State of CPaaS 2025

AI Voice and Conversational Intelligence

It's helpful to put AI Voice in the context of some other terms used in the industry, because some of the words and phrases sound similar but have nuanced meanings:

- **Intelligent Engagement**, discussed above, refers to *customers' desired end-goal outcome* of being able to interact cohesively with their ecosystem in real-time
- **Conversational Intelligence** describes the *process of bringing together different channels of communication* (e.g. calls, emails, texts, etc.) *and applying systematic intelligence* to those communications that treats them as part of an ongoing conversation with the customer
- **Programmatic Communications** is a definition that arose to describe how *a class of earlier stage media* (i.e. such as text, emails and IVRs) could be automated using IT
- **Enriched Engagement Media** describes *the newer class of media or channel* (e.g. RCS, WhatsApp, AI Voice) which are designed to be both *richer in format and more interactive and flexible in function*

AI Voice sits at the intersection of Enriched Engagement, Conversational Intelligence, and Traditional Voice, as AI Voice can also augment person-to-person voice (e.g. through transcription, translation, agent co-pilot, etc.)



Source: CPaaSAA

The vocabulary of Intelligent Engagement continues to evolve:

- **Intelligent Engagement** is the overarching concept, the business outcome of acting with your ecosystem in real time at scale. Everything else sits within it.
- Three overlapping domains sit inside it: **Enriched Engagement** (newer, richer channels like

RCS, WhatsApp, Network APIs), **"Traditional" Programmatic Comms** (established automated channels like SMS, Email, IVRs), and **"Traditional" Voice** (standard calls).

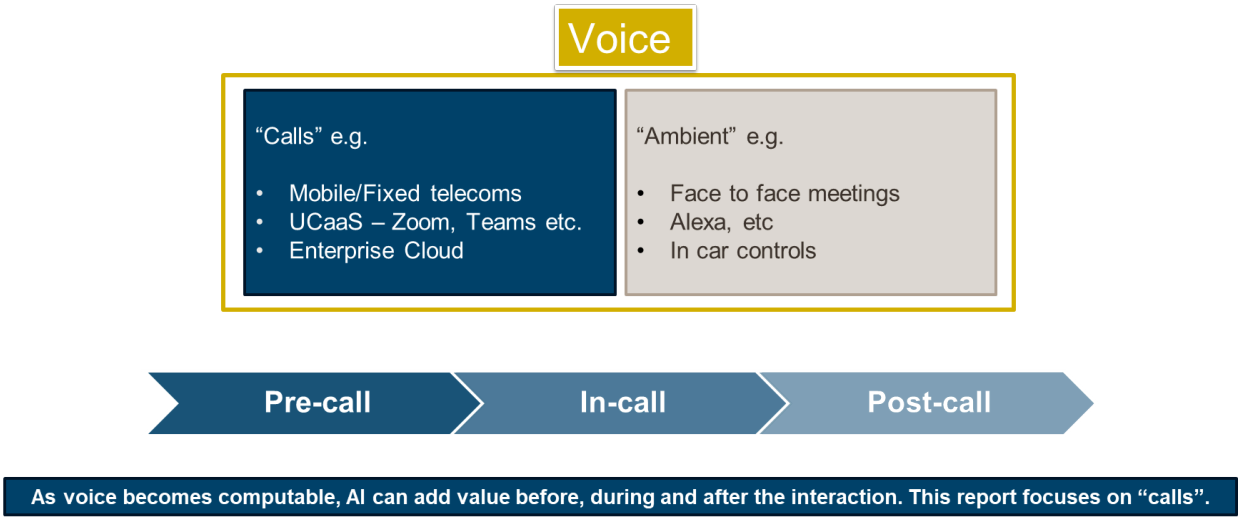
- **Conversational Intelligence**, AI and analytics applied to create insight-led, automated communications, sits at the intersection of all three, representing where the domains converge.
- **AI Voice** (highlighted in red) is shown at the overlap of Enriched Engagement and Traditional Voice, reflecting that it bridges new AI capability with the existing voice infrastructure.
- **Transcription** sits at the intersection of Traditional Voice and Conversational Intelligence, reinforcing the report's point that it is currently the most widely adopted entry point into AI Voice.
- **IVRs** sit at the overlap of Traditional Voice and Traditional Programmatic Comms, older automation that predates the AI era.

Calls and Voice in Context

Our analysis focuses on "calls", although it is notable that the same possibilities are available for any form of voice communication that is recorded - for example in a car, a meeting, or any recorded environment.

AI Voice Is More Than Just Calls

AI Voice spans traditional telecom calls as well as ambient and embedded voice interactions across devices and environments.



Source: CPaaSAA

WHY AI VOICE MATTERS TO -aaS PLAYERS & CSPS

Communications providers face slowing growth in traditional services while enterprises demand AI-enabled engagement.

"AI makes it possible to handle high levels of personalisation and scale… it is also enabling a renaissance in voice."

Robert Gerstmann, Co-Founder, Sinch

Market pressures. CPaaS, UCaaS, CCaaS players face a combination of threats and opportunities, including the plateauing or decline of SMS revenues and creeping and/or disruptive competition from different layers of the AI ecosystem. At the same time, communications platforms are increasingly embedding AI voice capabilities directly into their core infrastructure, accelerating enterprise deployment, although hyperscalers, CRM and enterprise software players are too. Telcos/CSPs also face these pressures and slow growth in many core communications areas.

Customer need: "how can you help me with AI?" In contrast, business customers have a clear need and desire to embrace AI and other technologies to update and improve their businesses (i.e. Intelligent Engagement). We will explore this significant business opportunity in the next section. Many businesses are a long way from being able to harness a powerful and complex development like AI Voice, and will choose different strategies to embrace AI Voice, including working with developers, integrators, and DIY builds from components and platforms.

AI Voice is complex to deliver well at scale with quality, and there is significant scope for -aaS players to integrate AI Voice services within their offerings, reducing complexity and friction for customers by combining pre-packaged solution components and know-how. As AI becomes embedded in communications platforms, the strategic question is shifting from who carries calls to who controls the runtime that executes them.

THE BUSINESS CASE FOR AI VOICE

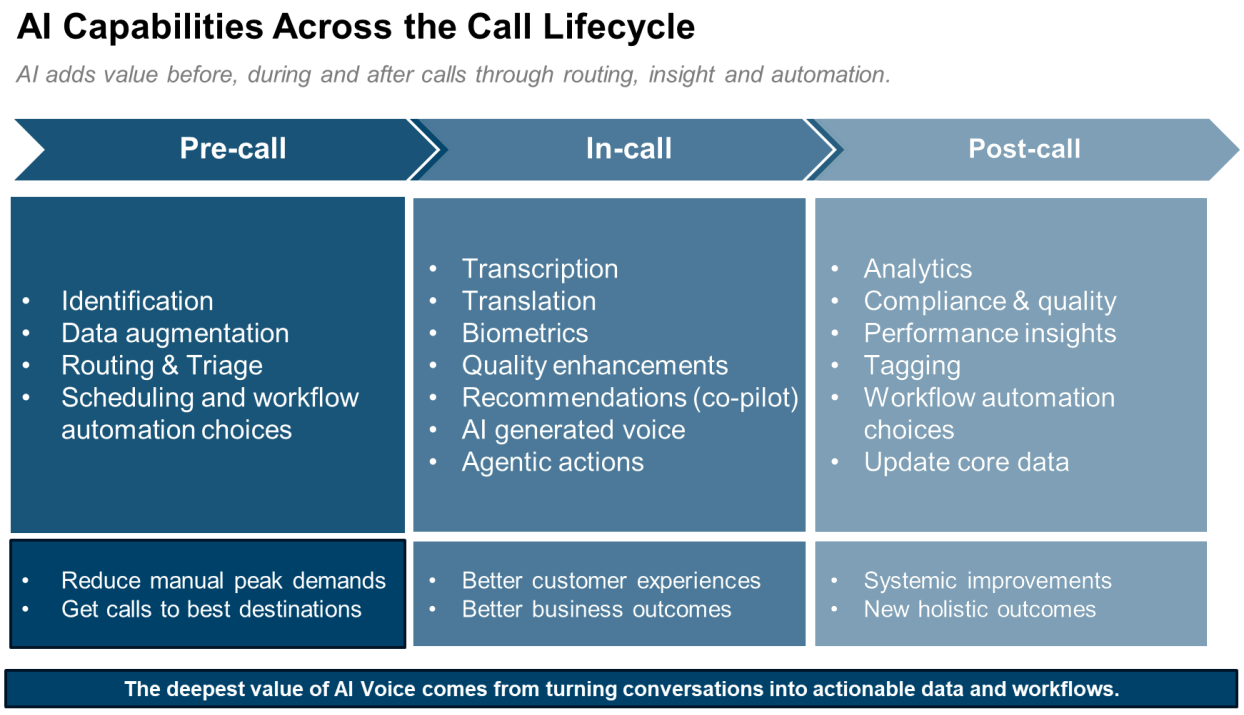
Operational impact: AI Voice improves performance before it transforms business models.

WHAT CAN AI DO FOR VOICE, AND VOICE FOR AI?

As noted previously, there are two fundamental opportunities for the combination of AI and voice:

- Using AI directly to communicate with humans, either by starting conversations in text channels or by implementing AI that listens and talks
- Using AI to turn voice conversations and recordings into data (transcription) - which saves time and effort, and that can be analysed with AI

The diagram below illustrates some of the functions that AI can perform.



Source: CPaaSAA

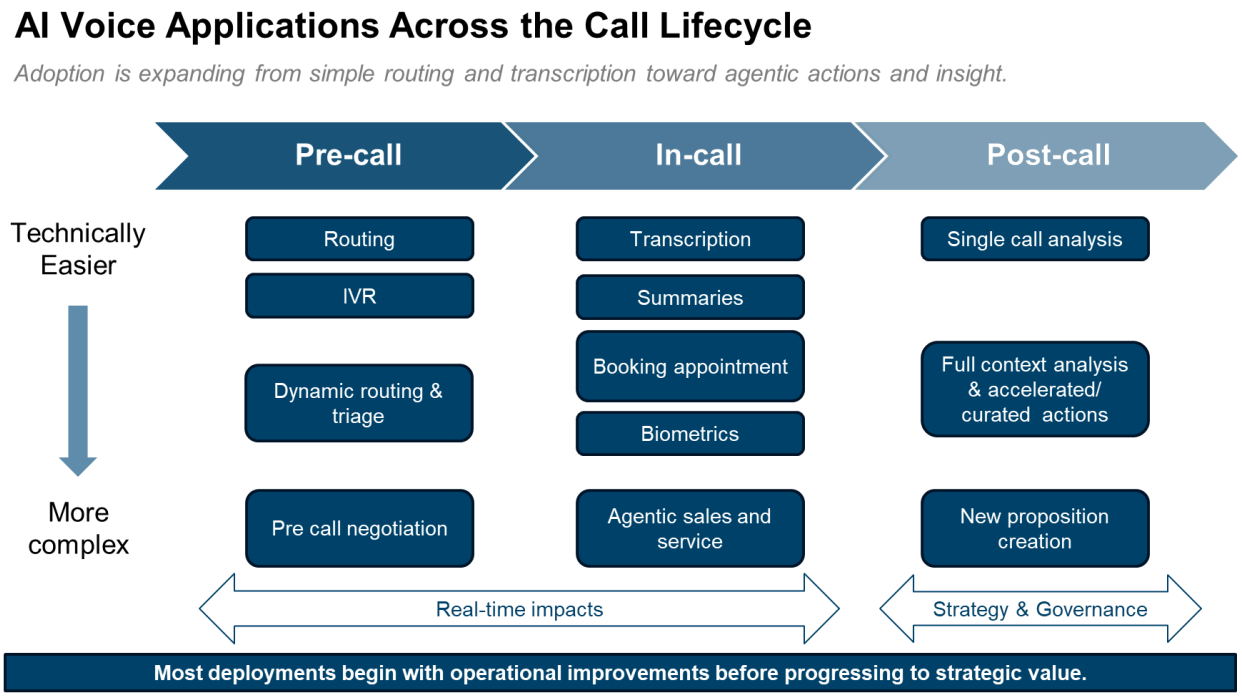
The benefits possible from these applications range from:

- Saving time on calls for customers and employees and resolving queries faster
- Automating actions and decisions, particularly in repetitive, high volume processes
- Managing calls out-of-hours and managing peaks of demand
- Better customer experiences and business outcomes in existing business scenarios
- Developing coaching, compliance and learning delivery
- Creating entirely new solutions, propositions, and outcomes

AI VOICE IMPLEMENTATION: CASE STUDY EXAMPLES

In practice, AI Voice applications come with varying ease of implementation.

AI Voice capabilities are expanding rapidly across the entire call lifecycle. Early deployments tend to focus on technically simpler tasks such as routing, IVR and transcription, which deliver immediate operational improvements. As the technology matures, organisations are beginning to introduce more complex in-call capabilities such as biometrics, booking and agentic service interactions, alongside richer post-call analytics. Over time, the greatest value is expected to come from combining real-time interaction improvements with deeper analysis and governance of conversations to create new services and strategic insight.



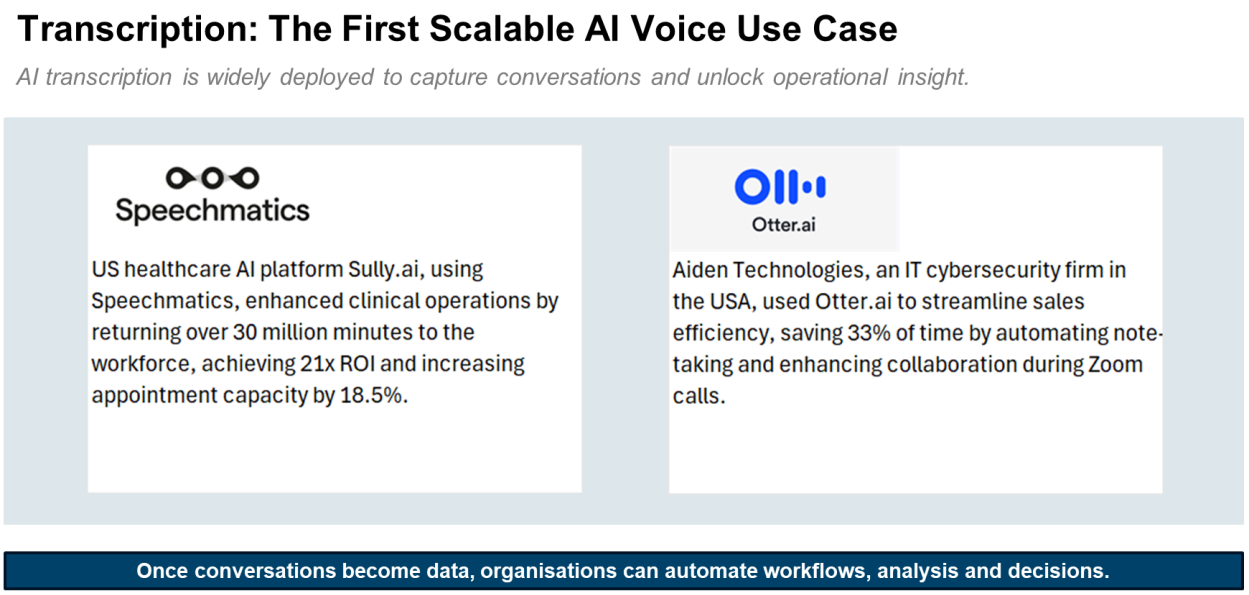
Source: CPaaSAA

This section outlines some of the most common initial application types, illustrated with case studies.

Early AI Voice deployments typically begin with analytics and assistance, before progressing toward automation and autonomous interaction, as illustrated by the following case studies.

Transcription

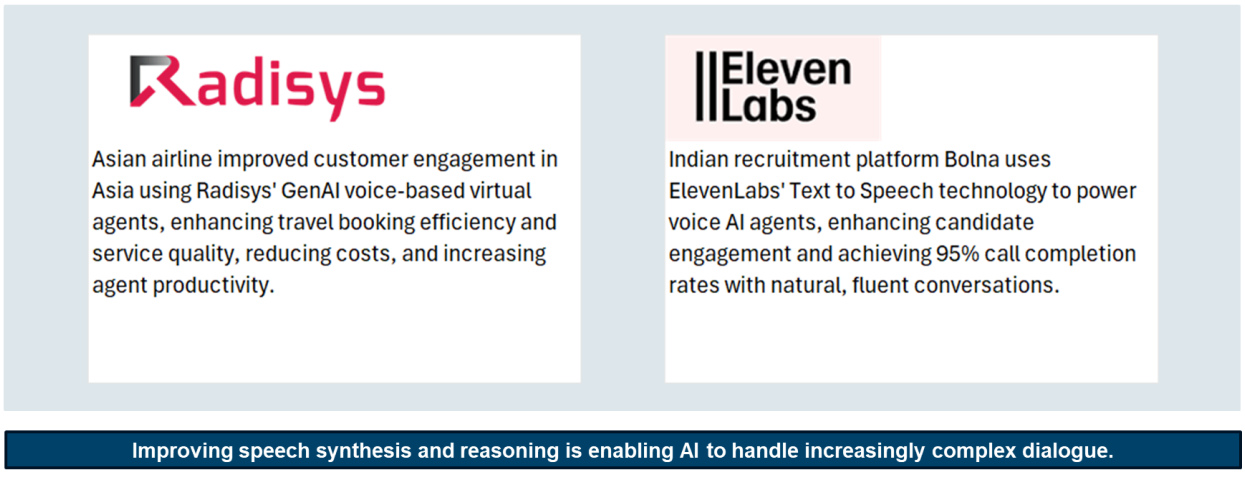
In terms of the technical and operational complexity of implementing AI Voice solutions, the most commonly adopted option is call transcription. Transcription services are now widely available in most online services such as Teams, Zoom and Google, and they are also available from specialised players and integrated into many CPaaS and UCaaS offerings. For example:



Source: CPaaSAA Case Directory

AI Dialogue: AI Voice becomes a two-way communication

Recent advances in speech recognition, synthesis and large language models are rapidly improving the ability of AI systems to conduct natural voice conversations. AI Voice is therefore moving beyond one way capabilities such as transcription or summarisation toward interactive dialogue, where systems can ask questions, interpret responses and complete tasks during a live conversation.



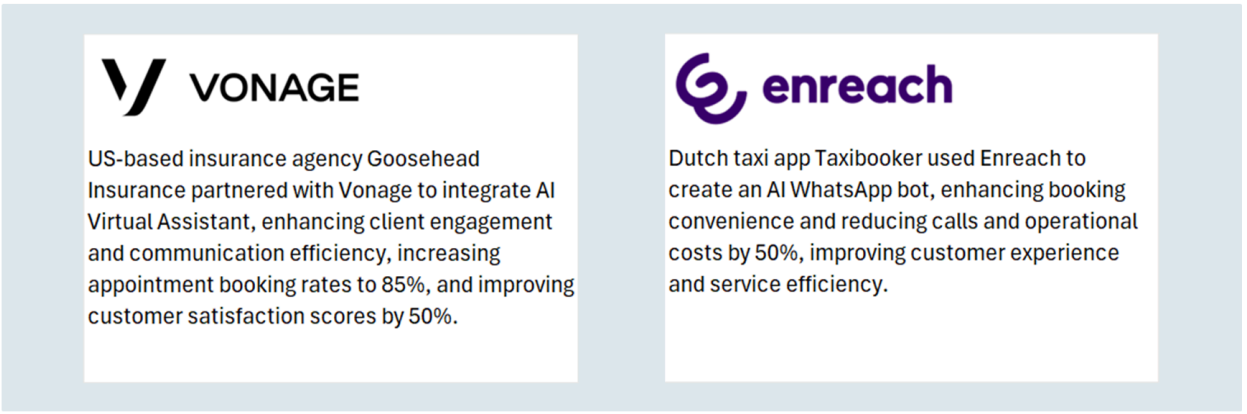
Source: CPaaSAA Case Directory

Appointments and Scheduling

With the addition of two-way dialogue, call triage and concierge services that allow people to book appointments and answer basic questions are the next most popular applications. This can start as a step up from traditional IVRs, but the solutions are significantly more interactive and increasingly capable.

Appointments and Scheduling: High Success Rates for AI Voice

AI assistants and messaging bots automate booking, reducing costs and improving service.



Appointment management is a natural early use case where automation delivers clear ROI.

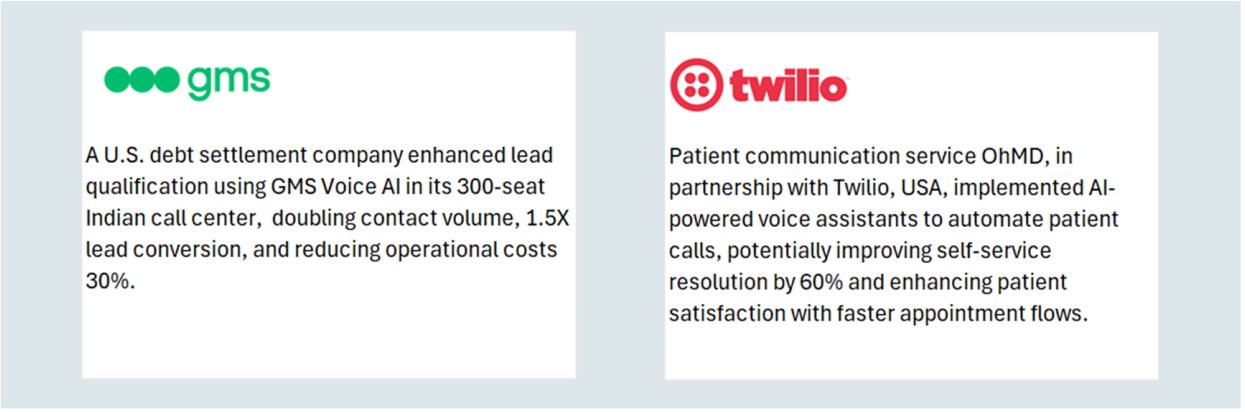
Source: CPaaSAA Case Directory

Lead Qualification

Lead qualification and process acceleration are also popular and practical applications for AI Voice.

AI Voice Can Improve Lead Qualification and Process Acceleration

AI voice agents qualify leads and handle routine interactions to increase conversions and efficiency.

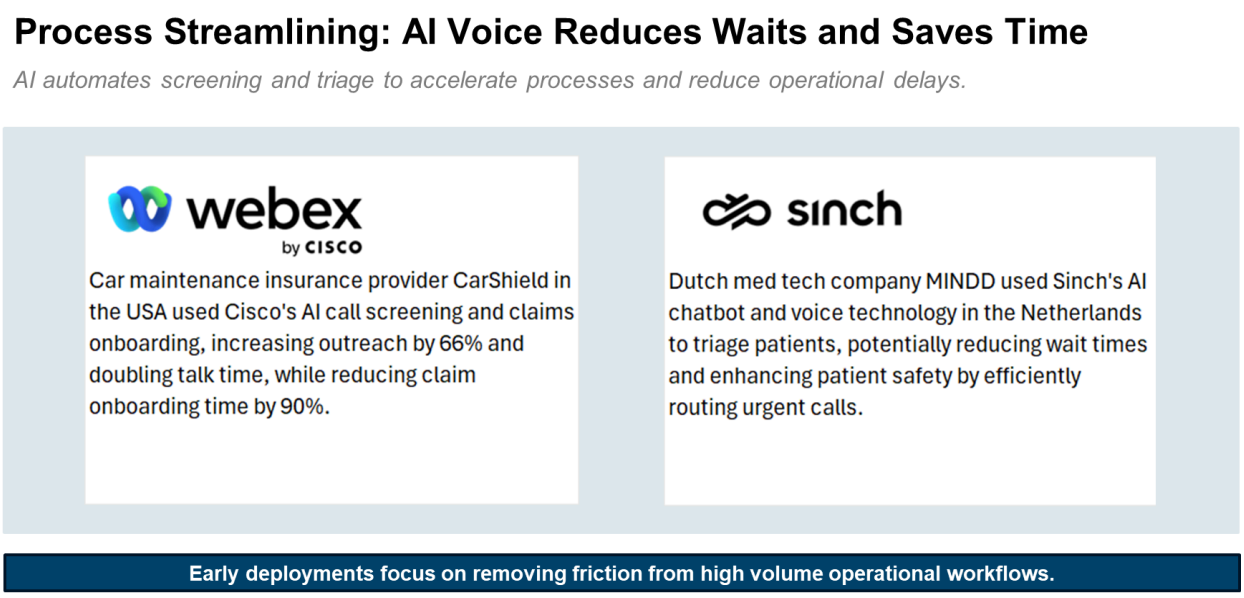


Automating early customer interactions can significantly increase throughput and reduce costs.

Source: CPaaSAA Case Directory

Streamlining Processes

Streamlining existing operations is a common theme in our Case Directory, and AI Voice applications are no exception.



Source: CPaaSAA Case Directory

Customer Experience

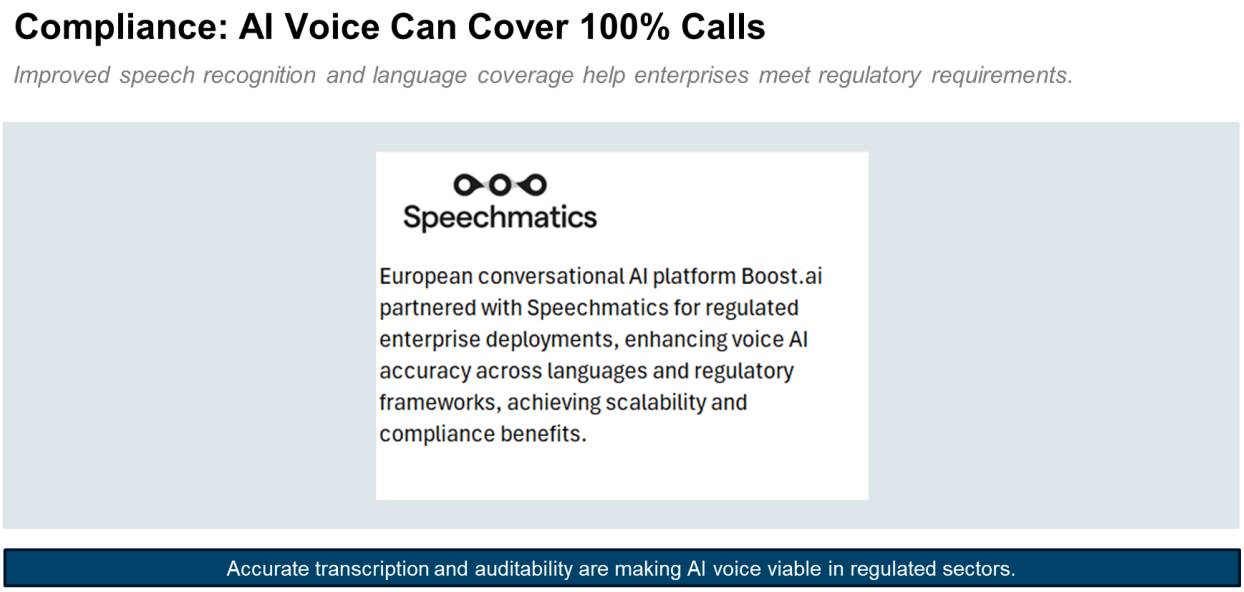
AI Voice is already being deployed to improve customer service operations. By automating elements of call handling, assisting agents and improving responsiveness, organisations can reduce complaints while increasing efficiency. The examples shown illustrate how AI enabled communications platforms are helping enterprises modernise service environments and deliver better experiences for both customers and staff.



Source: CPaaSAA Case Directory

Compliance

With AI Voice, compliance monitoring can be applied to 100% of calls rather than a sample, and this can pave the way to huge risk reduction and performance improvement potential.



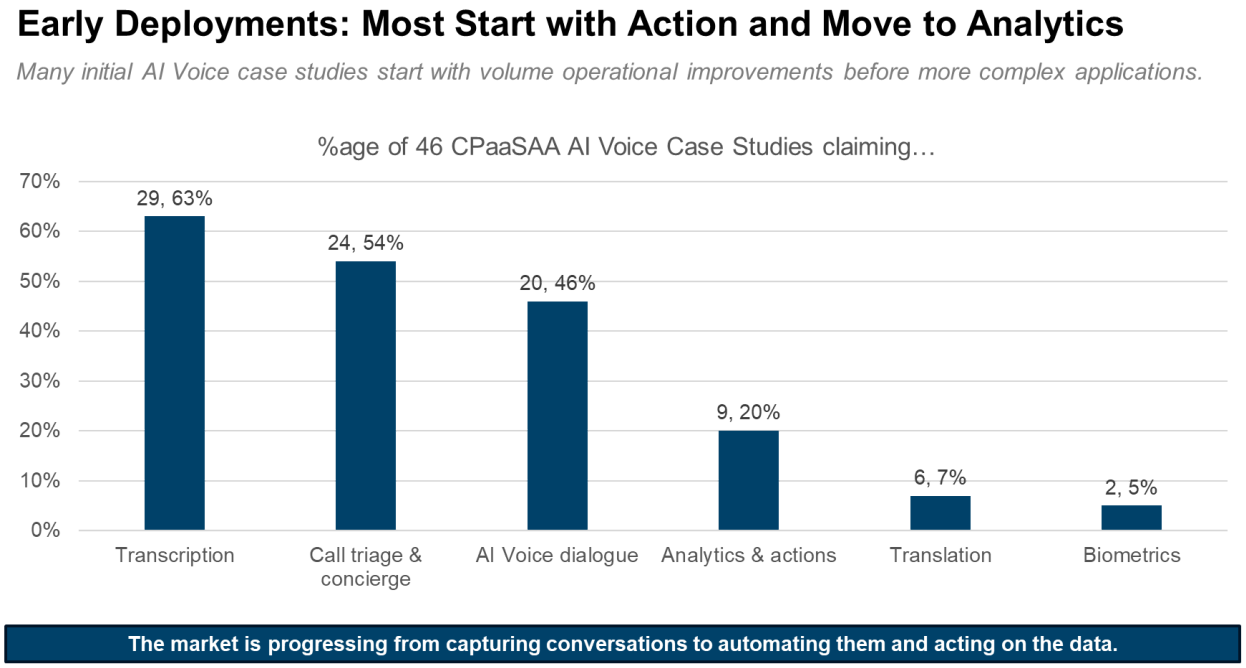
Source: CPaaSAA Case Directory

More advanced services include multi-lingual translation options, and biometrics and sentiment analysis, capable of adding new layers of information on identity and sentiment in a call. Analytics on call information ranges from relatively simple in-call summaries and the analysis of the content of an individual call, through to providing more holistic analysis, management information, recommendations and actions based on the combination of in-call information and a wider set of data. These applications are examined in the section on Market Opportunities.

Robert Arcon, CCO of IoTComms.io, points out that even the simplest implementations may require "plumbing": *"Hundreds of companies are building AI voice services, but most of them, and most enterprises, don't really understand real-time telephony or how to make legacy voice infrastructure work with modern AI at scale. That's the gap we fill: solving the 'boring plumbing' problem so AI platforms and existing telephony systems can reliably work together across many customers, not just in one-off integrations."*

Case Study Distribution

The following chart shows a breakdown of all AI Voice case studies in the CPaaSAA Case Directory (including those shown in this report), and it broadly backs this profile.



Source: CPaaSAA Case Directory

NB This analysis should be seen as illustrative rather than absolute - the Case Directory is a valuable resource that gives substantiated evidence of customer usage and benefits from surveys. However

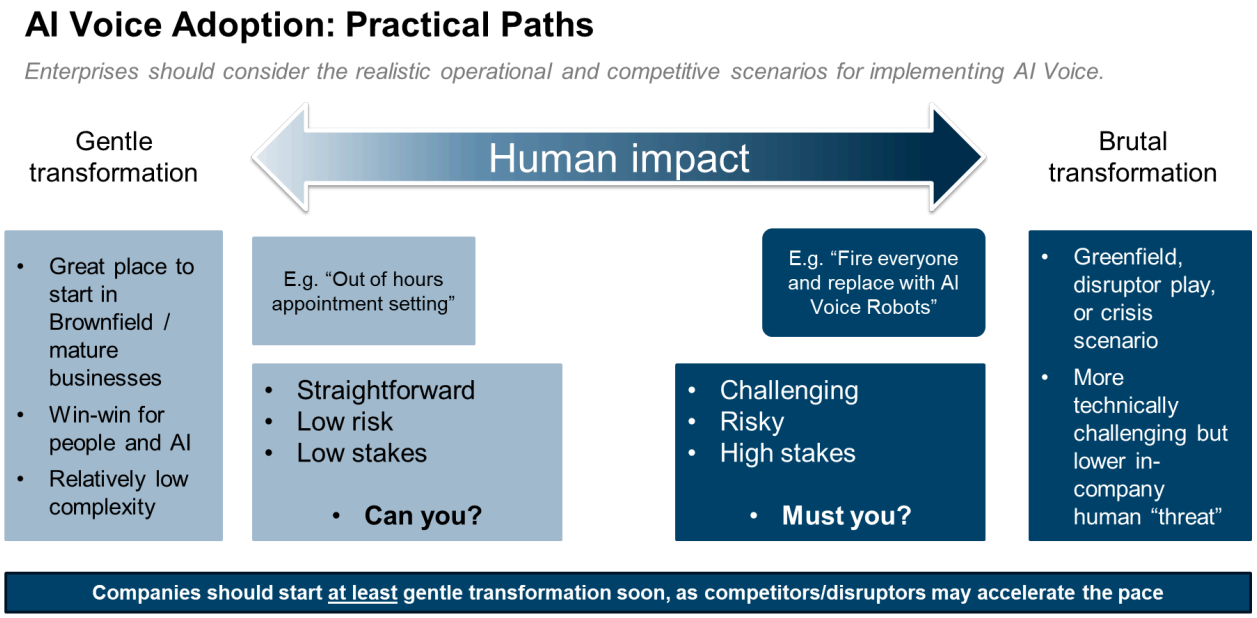
- It is a relatively small sample
- The cases are inevitably selected by their publishers on the grounds of their availability and the willingness of the client to share information
- Case studies take time to compile and publish, so even the most recent studies are not necessarily at the leading edge of application in AI Voice, as the market is moving so fast in some ways.

Nonetheless, it is a useful indicator and is at least based on real-world data rather than theoretical claims.

Data on more nuanced adoption of AI Voice approaches is still scarce, although the section on market adoption provides some snapshots of data available from specific studies in the US and the UK.

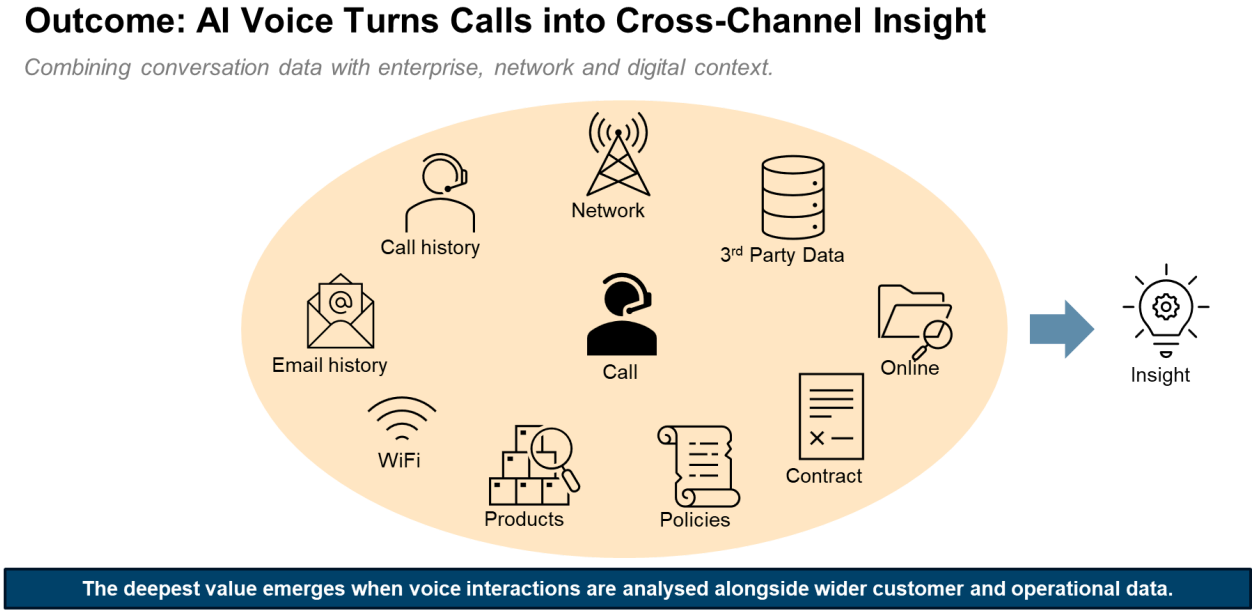
Workforce impact of AI Voice adoption

A further factor that influences adoption is the workforce impact and response to AI Voice. New technologies, especially AI, carry fears of displacement and disruption. In this regard, the simpler applications of AI Voice are much less troublesome - and indeed many initial applications we see of AI Voice are genuinely regarded as displacing work that is tedious, labour-intensive and stressful.



Source: CPaaSAA

Beyond the call itself



Source: CPaaSAA

AI analysis of a single call transcription can yield valuable insights, such as identifying actions needed, missed opportunities, and compliance issues - as well as the time saved by the transcription rather than manual data entry.

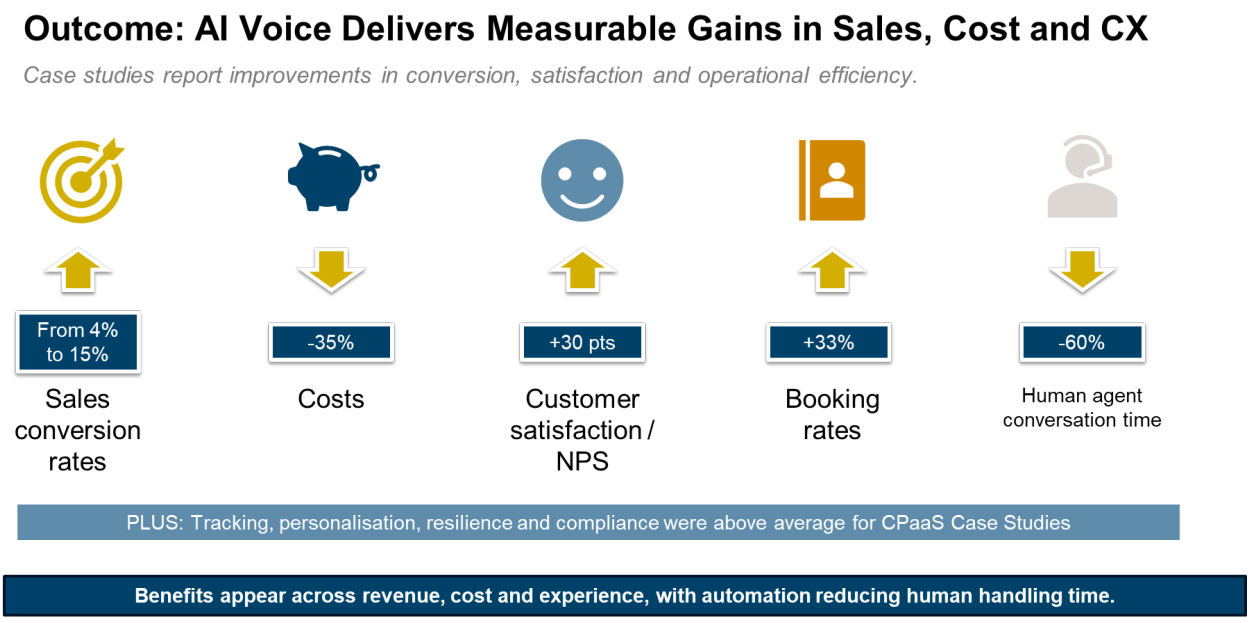
However, when combined with the broader information potentially available to the company from internal and external sources and signals, it is possible to derive even deeper insights and valuable actions.

For example, does this customer have a history of issues that point to a need for a new product, or a better service solution? Are there previously unseen systemic issues in calls over time in a particular call centre, location or for a specific advisor? And for enterprise customers, is there a pattern of issues or needs arising across a range of decision-makers that suggests a new need or opportunity?

BUSINESS BENEFITS OF AI VOICE

From an analysis of the cases in the CPaaSAA Case Directory, which spans a sample of cases from a variety of applications and industry providers, the benefits of AI Voice applications appear promising.

The chart below shows some illustrative examples. However please note: *customers should not expect a point implementation to achieve all of these at once.*



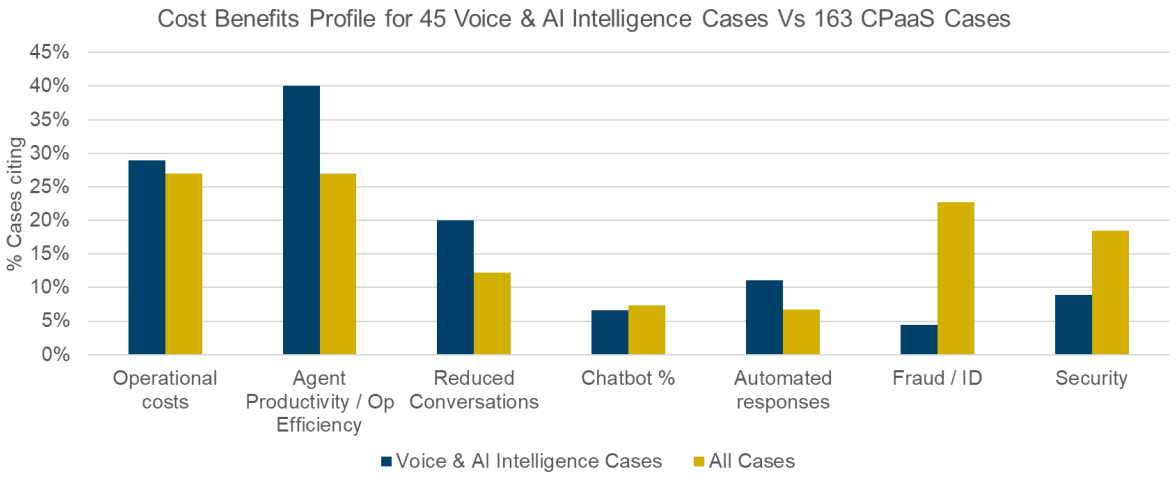
Source: CPaaSAA Case Directory. Note: Benefits represent illustrative examples observed across case studies.

In common with other data-based tools, strategic benefits of better tracking and compliance are rated highly for AI Voice.

There is also consistency between the evidence of benefits and the narrative of adoption. The strongest initial quantified business benefits appear to be in reduced conversation times, increased booking rates for appointments, and better lead conversion rates. This is consistent with the early applications of better call triage using AI Voice, and better availability and first time call resolution.

Operations: AI Voice Over Indexes on Productivity

Case studies emphasise agent productivity and operational efficiency gains.

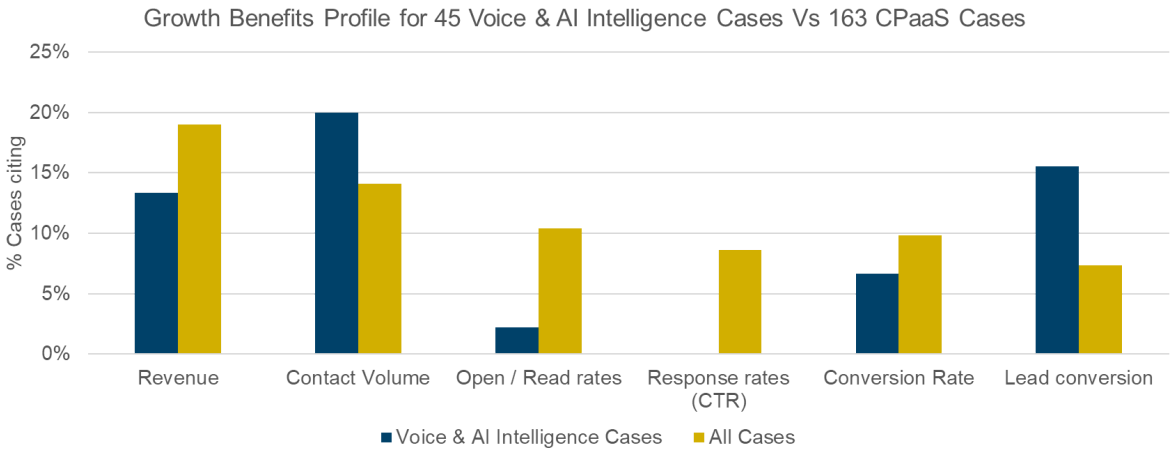


Compared with wider CPaaS cases, AI voice appears to more often improve workforce efficiency.

Source: CPaaSAA Case Directory

Growth: AI Voice Cases Show Volume and Lead Conversion Benefits

Early measures suggest AI Voice cases are still maturing compared to other Intelligent Engagement applications



Voice AI is often used to scale engagement, increasing contact volumes and downstream conversion.

Source: CPaaSAA Case Directory

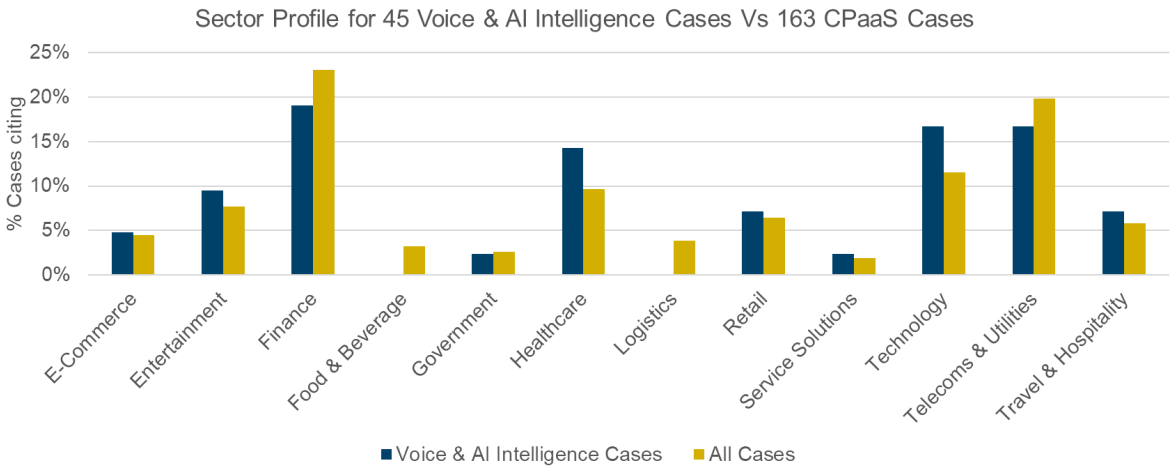
This sequencing of benefits, starting with strategic, response and experience improvements and progressing eventually to growth and revenue benefits, appears to be in line with broader trends in the realisation of benefits from AI and new technologies.

CUSTOMERS AND SALES CHANNELS: MINDSETS & BUDGETS

Data from the CPaaSAA Case Directory provides another snapshot supporting the notion that technology and service sectors lead in adoption of AI Voice at present.

AI Voice Adoption: Strongest in Tech and Service-Led Industries

AI Voice case sector distribution appears broadly similar to other Intelligent Engagement applications.



Use is concentrated where customer interaction volumes are high and service outcomes matter.

Source: CPaaSAA Case Directory

Some customers and channels are well attuned to the possibilities and opportunities afforded by AI Voice. But not all business people have a mindset that values time saving, quality processes and the possibilities of better analysis.

"Some smaller business customers are still using faxes in their daily business, and some telco sales people have spent all their lives selling based on reducing a specific bill by 10%. It can take a long time to change."

Mike Mills, Managing Director, Service Providers, at Gamma Telecoms

Moreover, AI Voice is not free, whereas for example a customer may be used to using an IVR in a legacy telephony system. Most customers need to work through the case to make the expenditure, and budget is rarely abundant in business in 2026.

For example, in research by Contact Center Helper, "What Contact Centers Are Doing Right Now (2025)", respondents from a mix of 10-1000+ seat contact centers said:

- **Budget was the top barrier at 68% to running an "ideal contact centre"**
- 49% had implemented chatbots and 45% AI, with a further 44% saying AI was on their wish list
- 47% had some form of interaction / speech analytics, and 16% had voice biometrics
- AI was on the "wish list" for a further 44%, and Inbound (49%) and outbound (14%) voice amounted for 63% of the contact mix

So overall, while the possible benefits of AI Voice look extremely promising, and there is likely to be fast growth in the field, not every customer and sales channel will immediately be willing and able to

harness the approach, and ecosystem players will need to support both customers and commercial teams with effective information, tools, and training.

"To get real value you need the last mile – the reseller or service provider closest to the end user – properly engaged. They know their customers, but they often don't take that extra consultative step. What they need is a simple recipe: a questionnaire that captures the key pain points and turns them into outcomes. Then you can apply the AI and voice toolbox in a way that fits the use case, instead of pushing generic ideas from a lab. It's really about teaching the channel to fish, and empowering them with tools to easily implement AI solutions."

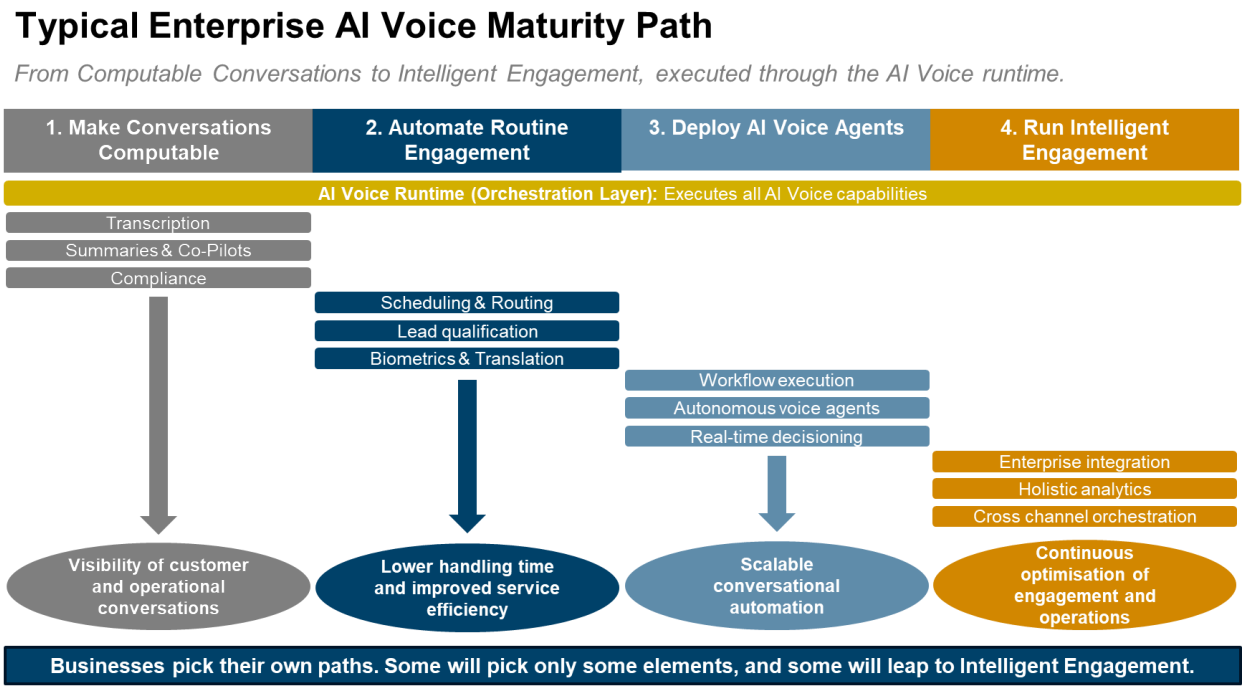
Bertrand Pourcelot, MD, UCaaS for CSPs & MSPs, Enreach

ENTERPRISE MATURITY PATH

Adoption pattern: Most enterprises move from transcription and analytics toward automation and eventually autonomous AI Voice agents.

Most enterprises do not adopt AI Voice as a single large transformation. Instead, adoption typically follows a staged path as organisations first make conversations computable, then automate routine interactions, before eventually deploying autonomous AI agents and integrated engagement systems.

Each stage increases both the value of conversational data and the strategic importance of the runtime that orchestrates it. As AI moves closer to real-time decisioning and enterprise workflows, the question of where the AI Voice runtime sits in the architecture becomes increasingly important.



Source: CPaaSAA Research

ETHICS, PRIVACY AND REGULATORY COMPLIANCE

The rapid evolution of AI Voice capabilities, especially in biometrics, sentiment analysis, and the use of large language models, introduces critical legal and ethical considerations that impact implementation at the enterprise level.

Compliance with evolving data privacy regulations such as GDPR, CCPA, and new legislation governing the use of biometric identifiers and voice recordings is a fundamental prerequisite.

Enterprises must establish robust governance frameworks to ensure transparent data usage, secure customer consent for voice data processing (particularly for biometric authentication), and protect against misuse, such as deepfake vulnerabilities.

Risks from unintended consequences or failures of AI use include:

- **Hallucinations causing compliance breaches.** AI models can generate plausible but factually incorrect outputs. In regulated environments such as financial services or healthcare, an AI agent acting on or communicating a hallucinated response could constitute a compliance violation, create legal liability, or cause direct customer harm.
- **Data retention and consent missteps.** Voice data captured for AI processing (transcription, analysis, biometrics) may be stored, shared or used in ways that inadvertently breach GDPR, CCPA or sector-specific regulations. Consent frameworks designed for traditional call recording may not adequately cover the broader uses enabled by AI.
- **Cost blowouts from transcription and inference.** AI Voice workloads are consumption-based, and costs can scale unexpectedly with call volume, model complexity or real-time processing requirements. Without robust usage monitoring and budget controls, enterprise deployments can exceed projected costs significantly.
- **Vendor lock-in via embedded copilots.** As AI copilots become deeply integrated into agent workflows and CRM systems, switching costs rise sharply. Enterprises that build processes around a specific vendor's AI tooling may find themselves structurally dependent on that vendor's pricing, roadmap and data access terms.
- **Model drift.** AI models can degrade in performance over time as the real-world language, topics and customer behaviours they encounter diverge from their training data. Without ongoing monitoring, a model that performed well at deployment may quietly become less accurate or appropriate, without any obvious failure signal.
- **Evaluation gaps**. Many organisations lack the frameworks, metrics or tooling to assess whether their AI Voice deployments are actually performing as intended. Without rigorous evaluation, underperformance, bias or unexpected behaviour may go undetected until it causes a visible problem.

For -aaS and CSPs, supporting customers with compliant, 'privacy-by-design' solutions is a non-negotiable step to reduce adoption friction and build trust.

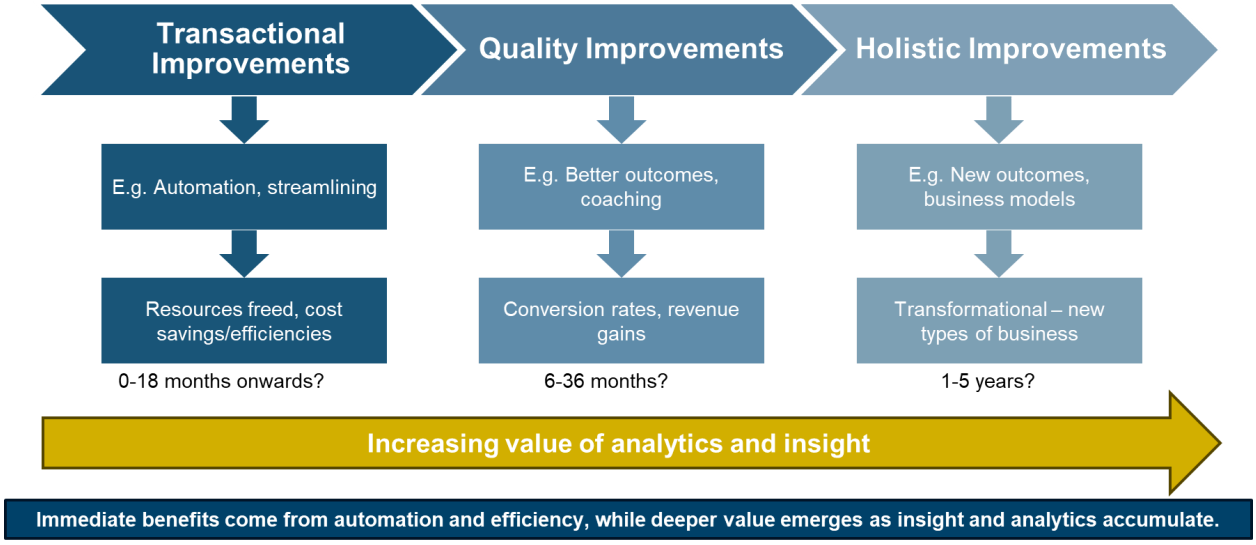
These risks do not prevent adoption, but they mean governance must be designed into AI Voice deployments from the start.

WHEN WILL THE BENEFITS MATERIALISE?

Depending on the complexity, cost and success of implementation, this leads to the following hypothesis that the benefits of AI Voice (and other AI-based technologies) will take time to materialise in industry as a whole, and typically within a given business.

AI Voice Benefits: Typically Operational First, Strategic Later

Businesses typically progress from operational gains to strategic impact over time.



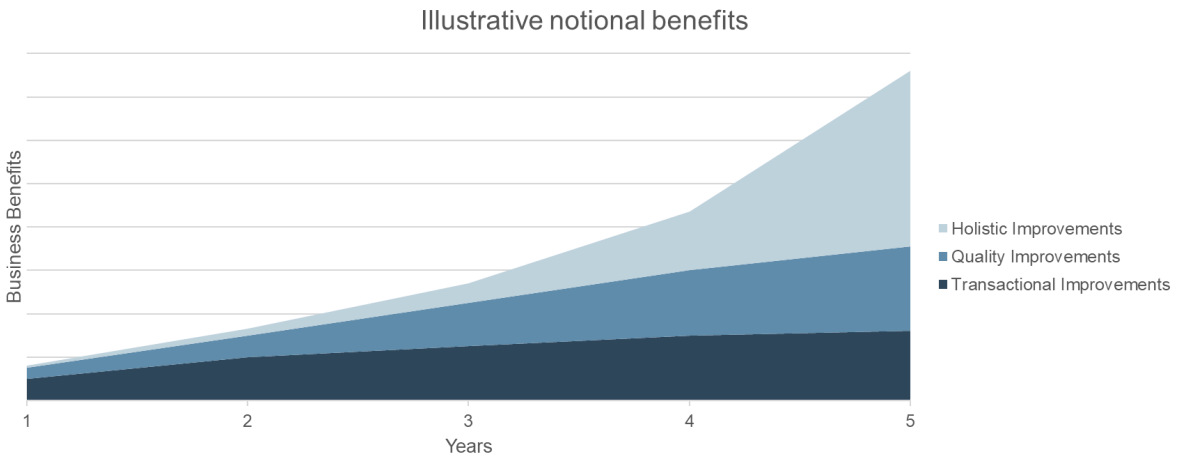
Immediate benefits come from automation and efficiency, while deeper value emerges as insight and analytics accumulate.

Source: CPaaSAA Research

Some businesses - and some industry partners - may be able to accelerate this cycle, and indeed, in the market adoption section, we share some examples of this. However at an overall market level, we expect the benefits will emerge over time, as shown in the following notional illustration,

AI Voice Value is Likely to Compound Over Time

Operational gains typically come first, followed by quality improvements and broader strategic impact.



The largest benefits are likely to emerge later as analytics, learning and new business models accumulate.

Source: CPaaSAA Research

INDUSTRY ECOSYSTEM & DYNAMICS

AI Voice brings telecom networks, hyperscalers and enterprise platforms into direct competition across three functional layers.

THE AI VOICE VALUE STACK

Multiple technology sectors are now converging around AI Voice, bringing telecom operators, hyperscalers, enterprise software platforms and communications providers into direct strategic competition.

Moreover, delivering AI Voice solutions to businesses that are fully operational within required governance, quality and cost standards, and that achieve each business's goals, is driving a complex and growing ecosystem.

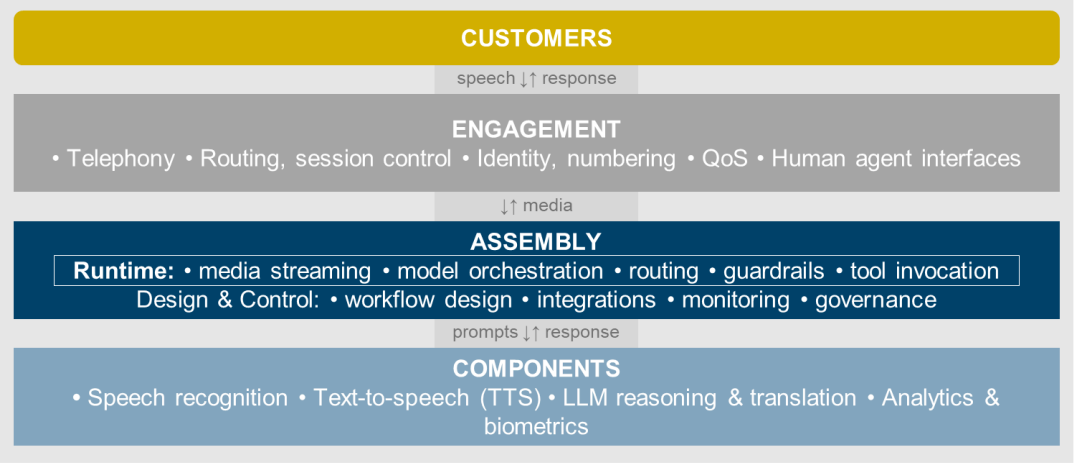
This section gives a high-level overview of the supply-side player structure and player types.

AI VOICE ECOSYSTEM

The AI Voice market is structured around three functional layers: Engagement, Assembly and Components. Understanding who controls the runtime, which sits in the Assembly layer, therefore becomes central to understanding the structure of the AI Voice ecosystem.

AI Voice Ecosystem: Control Sits in the Runtime Layer

How conversations are delivered, orchestrated and powered by AI components.



Control of runtime layer determines how conversations execute, which AI models are used and where value accumulates.

Source: CPaaSAA Research

The three layers are:

- **Engagement** is where live interactions reach customers, via telephony networks, CPaaS APIs, CCaaS platforms and applications. It represents the channels and infrastructure through which customers and enterprises communicate. Players include:
    - Telcos, e.g. Vodafone, Orange, T-Mobile, BT, KPN, etc
    - CPaaS (telephony APIs), e.g. Vonage, Infobip, Sinch, Twilio
    - CCaaS / UCaaS, e.g. Fiveg, RingCentral, 8x8
- **Assembly** combines components into working systems. This is where workflows are defined, tools are connected, policies are enforced and conversations are orchestrated for delivery through engagement channels. It is also where the **runtime** sits: the execution environment that manages conversational timing, routing, model invocation and guardrails during live interactions. Assembly may occur within communications platforms such as CPaaS or CCaaS systems, or within enterprise workflow platforms such as CRM systems where conversational intelligence is integrated directly into business processes. Players include:
    - CPaaS and CCaaS players (see above)
    - Hyperscaler AI platforms, e.g. Google (Vertex AI), Microsoft (Azure AI)
    - CRM / workflow systems (Salesforce, ServiceNow) and enterprise data platforms where conversational intelligence is operationalised. E.g. CRM providers such as Salesforce are embedding AI agents and voice capabilities directly into customer-data platforms, positioning enterprise workflows as an alternative execution environment for conversational AI.
    - Communications platforms, e.g. Webex, CM.com
    - Integrators and developers, e.g. Accenture, Tech Mahindra, Cognigy
- **Components** provide the building blocks: speech recognition, synthesis, LLM reasoning, translation, analytics and biometrics. These capabilities are advancing rapidly and are increasingly modular. Players include:
    - Speech recognition, e.g. Speechmatics, Deepgram
    - Speech synthesis, e.g. ElevenLabs
    - LLMs / reasoning, e.g. Anthropic, Google DeepMind, OpenAI
    - Infrastructure, e.g. NVIDIA

Much current market attention focuses on the release and development of new LLMs and speech models. These improvements expand what AI Voice systems can do. However, the key strategic control point sits in the Assembly layer, and specifically in the runtime that governs how conversations are executed.

The durable advantage in AI Voice is therefore unlikely to come from exclusive dependence on any single model provider. Instead it will come from control of the execution layer, integration into enterprise workflows, and the trust framework surrounding the service.

The following section outlines how control of the runtime in the Assembly layer determines how conversations are executed, how intelligence is applied and where structured data accumulates.

RUNTIME: THE CONTROL POINT FOR VALUE

Runtime determines how conversational data flows, models execute and actions are delivered during live interactions.

In platform markets, the layer that controls runtime execution typically captures the greatest strategic leverage because it determines how components interact and where data and decisions flow.

In AI Voice, the strategic question is therefore no longer who carries the call, but who runs the conversation.

"Runtime" is a term[1] used to describe the execution environment that orchestrates and manages live conversational systems.

Runtime's role in the delivery of AI Voice makes it the key commercial control point. Whoever controls execution shapes data flow, model choice, compliance and switching costs, and therefore captures disproportionate leverage.

From a technology/process viewpoint, Runtime:

- Orchestrates models and ASR, LLM and TTS components
- Controls and manages real-time audio streams and latency
- Enforces policy, and connects to enterprise tools systems such as CRM and payments and maintains operational state
- Enforces guardrails and compliance policies
- Logs and audits interactions, and governs escalation to human agents

While in technical terms, runtime is the operating system of AI Voice, in commercial terms, it is the strategic control point, as whoever owns runtime determines:

- Which models are used and how easily they can be switched
- Where data flows and where it is stored
- How policy and compliance are enforced
- How analytics are captured and monetised
- Where vendor dependency sits

Moreover, runtime is the strategic engine and monetisation layer of AI Voice in Intelligent Engagement. It is where real-time voice data is governed, structured and connected to action.

Where CPaaS/CCaaS players own runtime, it will mean that they can deliver superior customer experience, enforce end-to-end compliance, and most importantly, directly own the value-added service, the intelligence, that customers pay for, allowing the shift from volume-based to intelligence-based monetisation.

[1] The term comes from traditional computing, and refers to the system that loads code, manages memory, handles inputs and outputs, and ensures programs run correctly once deployed. It sits between the raw hardware and the application logic. In modern cloud systems, runtime has evolved to mean the execution layer that governs how services run in production, including scaling, security, logging and performance management.

That makes runtime the key control point in the ecosystem, and the key strategic area of competition between -aaS Player/CSPs, and hyperscale players.

Recent Market Signals

Recent announcements illustrate this architectural shift:

- RingCentral introduced AIR Pro, embedding AI agents directly within its communications platform.
- Sinch announced Voice Relay, enabling developers to connect AI agents into live calls through its CPaaS voice infrastructure.
- Vonage integrated AWS's Nova Sonic speech-to-speech model into its Voice API, allowing conversational AI agents to operate in real-time telephony environments.
- Zoom also expanded conversational AI capabilities across its collaboration and contact centre platforms.
- Other vendors are also emphasising orchestration as the key control point. At Enterprise Connect 2026, Five9 expanded its Fusion partner ecosystem to connect AI agents, enterprise data and enterprise systems across its Intelligent CX platform. This approach positions the contact centre platform as a coordination layer between AI models, communications channels and business workflows.
- Cisco is also embedding AI Voice capabilities directly within communications platforms. Its Webex AI Agent enables enterprises to deploy autonomous voice or digital agents that can interact with customers and complete tasks using enterprise data and workflows. Cisco has also introduced AI Assistant capabilities for Webex Contact Center and an AI Agent Studio environment for building and orchestrating conversational agents. These developments illustrate how communications platforms are increasingly embedding AI runtime capabilities directly into their core architecture rather than relying solely on external AI services.

Together these developments reinforce the importance of the runtime layer that orchestrates live conversations between AI models, communications infrastructure and enterprise workflows.

As Rob Kurver noted following discussions with platform vendors at Enterprise Connect 2026, *"the moat is not just the model. It is the real-time interaction layer around AI."* In practice this interaction layer corresponds closely to the runtime environment that orchestrates conversational data flows, models and workflows during live communications.

These developments illustrate a broader structural trend: the communications platform itself is becoming the execution environment for AI agents. This strengthens the role of CPaaS, CCaaS and UCaaS providers as potential runtime owners rather than simply connectivity providers.

Compute Location vs Runtime Ownership

Compute location does not determine runtime ownership. Most enterprises will consume AI through cloud APIs or SaaS, while some use reserved GPU capacity or on-prem hardware. Enterprises may run models on hyperscaler infrastructure while orchestrating conversations through a CPaaS or CCaaS runtime, while telcos may embed runtime in their networks. Control of runtime ultimately determines where value accumulates in the AI Voice stack.

FOUR RUNTIME SCENARIOS: HOW BUSINESSES DECIDE

When a CIO or COO approves an AI Voice programme, they decide "who runs the conversation":

- Where conversation logic is defined
- Where policy is enforced
- Where structured conversation data sits
- How easily models can be switched
- Who is accountable if something fails

That decision determines who governs live customer interaction, and who controls the intelligence margin that follows. In practice, enterprises gravitate toward four routes.

Who Runs the AI Voice Runtime? Enterprise Deployment Models

Four architectures for operating the AI voice orchestration layer.

Who runs the AI Voice runtime?	Intelligent Engagement Platform	Hyperscaler Led	Telco Integrated	Enterprise-Controlled
Customer characteristics	• Customer engagement led • May already have CPaaS/CCaaS	• AI Voice is just part of AI strategy • Internal teams are cloud-native	• Highly regulated / sensitive • Network trust a priority	• AI Voice is strategically differentiating • Strong engineering capability or enterprise platform control
Choice drivers	• Governance • AI-Model / HSP supplier neutrality	• Speed and architectural simplicity	• Sovereignty • Trust/resilience • AI in every call	• Control of conversation data and models • Minimise vendor lock-in
Trade-offs	• Governance and control over simplicity	• Speed and simplicity over control	• Sovereignty over flexibility	• Maximum control, higher engineering and integration overhead
Estimated. deployments today	~40–60%	~20–35%	<5%	~10–20%

Source: CPaaSAA Research

The following sub-sections outline these scenarios in further detail.

Intelligent Engagement Led Runtime (CPaaS / CCaaS)

Why a customer chooses this:

- AI Voice is framed as customer engagement transformation, not just AI infrastructure
- They already rely on CPaaS or CCaaS for routing, workforce tools and compliance
- They want separation between model supplier and orchestration layer
- Governance, audit and switching flexibility are board-level concerns

What it looks like:

- Telephony routing, session management and escalation logic are managed inside the communications platform
- The platform hosts the orchestration engine that defines conversation flows and policy rules
- External AI models (speech recognition, LLMs, synthesis) are connected via APIs
- The communications platform decides when and how those models are invoked
- CRM, payments, workforce and compliance systems integrate into this platform
- Conversation logs and structured data (e.g. vCons) are captured and governed there, even if inference runs in external cloud infrastructure

In this scenario, models execute in hyperscaler infrastructure, but execution control sits with the communications provider.

A similar architecture is emerging in new CPaaS voice platforms. For example, Sinch recently announced Voice Relay, which sits between the telephone call and the AI agent. The platform converts speech to text, passes it to an AI model, converts responses back to speech, and manages real-time conversational timing, interruption handling and latency optimisation. Developers connect text-based AI agents to the global telephone network through a simple interface while the communications platform manages the real-time conversational loop.

CCaaS platforms are also expanding their role as orchestration layers for AI-driven customer experience. Five9's Fusion ecosystem connects AI agents, enterprise data and partner technologies across its Intelligent CX platform, enabling organisations to deploy AI voice capabilities through their contact centre while integrating with CRM systems and enterprise workflows.

Estimated split:

~40–60% of new enterprise AI Voice runtime deployments, based on:

- High levels of enterprise preference for buying packaged AI solutions rather than building (Menlo Ventures 2025 survey: ~76% buy vs build).
- Significant installed base and growth of CCaaS platforms as the control layer for contact centre environments.

Hyperscaler-Led Runtime

Customers choose on the basis of consolidation and velocity:

- Most AI, analytics and data workloads already sit in one cloud
- Internal teams are cloud-native
- Speed and architectural simplicity matter
- Voice is treated as an extension of a broader AI strategy

What it looks like:

- Conversation workflows built and managed in hyperscaler AI platforms
- Speech, LLM reasoning and synthesis run natively there
- Orchestration logic (branching, escalation, action triggers) defined in hyperscaler tools
- CRM and enterprise systems integrate directly into that cloud stack
- Conversation data, transcripts and analytics are stored inside the hyperscaler environment
- CPaaS or telco providers deliver media streams into this system

Here, orchestration, data and switching costs accumulate in the hyperscaler stack.

Estimated split:

~20–35% of new deployments, based on::

- Strong enterprise GenAI exploration rates (Gartner: high levels of pilot/exploration in customer service)
- Hyperscaler expansion into full-stack conversational platforms
- Cloud standardisation trends in large enterprises

Telco-Integrated Runtime

Customers choose based on infrastructure-level trust:

- Sovereignty, resilience or regulatory exposure is high
- Identity, fraud prevention or trust at the network layer matters
- They operate in finance, healthcare, government or critical infrastructure
- Latency or infrastructure alignment is strategic

What it looks like:

- AI capabilities such as translation, fraud detection or identity verification are embedded into the telecom network or edge layer
- Routing logic may trigger AI services automatically during call setup
- In advanced cases, policy enforcement and escalation logic are embedded closer to the network core
- Enterprise systems integrate with telco-hosted services rather than hyperscaler orchestration

The structural distinction is whether the telco controls orchestration and data governance, or simply injects AI features into the call path. If orchestration remains elsewhere, telco AI is additive, whereas if execution control sits in the network layer, it becomes structural.

Estimated split

<5% of enterprise runtime control today, based on:

- There is limited evidence of widespread telco-owned orchestration layers at enterprise scale. Adoption appears concentrated in specific sovereign or regulated contexts rather than as a dominant global model.

Enterprise-Controlled Runtime

The enterprise controls the runtime, either through internally engineered orchestration frameworks or through enterprise platforms such as CRM systems.

Why a customer chooses this:

- AI-driven interaction is strategically differentiating to the business
- Desire to retain control of conversation data, models and suppliers
- Strong internal engineering capability or digital-native operating model
- Preference to avoid platform lock-in and maintain architectural flexibility
- Need to integrate conversational AI deeply with proprietary systems and workflows

The logic is maximum control.

What it looks like:

- Open orchestration frameworks (e.g. LiveKit, Pipecat) handle media and model pipelines
- Speech and LLM APIs connect directly to enterprise code
- Conversation logic is custom-built internally
- Compliance, audit and data governance are engineered in-house
- Data stored in enterprise-controlled infrastructure

Some enterprises may adopt AI Voice through existing enterprise platforms rather than communications providers. Salesforce recently introduced Agentforce Contact Center, which embeds AI agents and voice capabilities directly within the CRM environment. In this model, conversations are managed alongside customer data and workflows inside the enterprise platform itself.

This approach effectively places orchestration inside the enterprise application layer rather than the communications platform.

Estimated split:

~10–20% of deployments, based on:

- Menlo Ventures data showing majority preference for buying rather than building AI systems
- Open orchestration lowering barriers, but governance and maintenance overhead limiting broad adoption.

"My customers sometimes move from managed platforms to open orchestration frameworks like PipeCat or LiveKit Agents, which let you plug together best-of-breed components for transcription, LLMs, TTS and telephony, tune them for your specific workflows. It gives them a lot of control over their solutions and suppliers."

John George, AI Voice Developer

COMPETITIVE RISKS

The orchestration layer for AI Voice is emerging across multiple domains: communications platforms, telecom networks, hyperscale AI platforms and enterprise workflow systems. Competition for control of conversational execution is therefore expanding beyond traditional communications infrastructure alone.

So while AI Voice expands both the opportunity and the competitive arena for Intelligent Engagement players, the control point in this arena is moving from transport to orchestration, governance and data. If that control belongs to others, margin follows.

The competition for AI Voice is therefore not simply about models or channels, but about which platform becomes the operational runtime for conversations.

There are four areas of competitive exposure.

1. **Hyperscaler Consolidation**. Cloud platforms increasingly offer end-to-end AI Voice stacks: models, orchestration, analytics and CRM integration. If enterprises standardise on hyperscaler runtimes, CPaaS and CCaaS risk being reduced to media providers. The platform that governs execution governs data flows, model choice and monetisation.

2. **CRM and Workflow Platforms**. CRM vendors are embedding voice AI directly into sales, service and marketing workflows. If conversation intelligence is native to the system of record, value migrates into that application layer. Communications platforms risk losing influence over both insight and outcome monetisation.

3. **Contact Centre AI Incumbents**. Established CCaaS and workforce optimisation vendors are strengthening their AI capabilities. Those that successfully integrate copilots, analytics and automation into their core environments can capture runtime control within the contact centre itself, limiting the role of external orchestration layers.

4. **Vertical AI Voice Specialists**. In regulated industries, specialised providers are building domain-specific voice agents and compliance frameworks. Where these players own workflow, regulatory logic and data context, horizontal platforms may struggle to displace them.

If Intelligent Engagement players remain focused on connectivity and usage, value migrates upward. Whereas if they influence or control runtime, governing models, policy, integration and structured data, they retain strategic relevance.

SOVEREIGNTY AND STRUCTURAL TENSION

Geopolitical risk introduces nuance, as sovereignty concerns push towards architectural separation and greater control over execution, whereas operational simplicity pushes towards consolidation around hyperscaler platforms.

Boards increasingly ask where execution control sits and whether models can be switched, while CTOs often seek standardisation and performance.

The likely outcome is a mix of approaches:

- Hyperscalers dominate commercially driven, open segments
- Communications platforms strengthen in regulated enterprise sectors
- Telcos gain selective influence where sovereignty and resilience are strategic priorities

Telcos are not the only players eyeing sovereignty. For example, Remco Magielse, co-founder GLBNXT says it is designed to serve organisations *"who don't want their most sensitive data held by US companies; they want sovereign AI they can control."* He describes the company as *"the Stripe for AI"* – a single platform that bundles hardware, models and governance so customers don't have to build the whole stack themselves, and can scale quickly from experiments to real-world production.

Trust should increasingly be understood as part of the product value proposition, not only as a compliance requirement. Privacy by design, auditability, resilience and sovereign deployment are emerging as differentiators in environments where enterprises and regulators require assurance around how AI systems operate in live communications.

The key for -aaS players and telcos will be how fast and persuasively they act and convey their relative benefits to customers, outlined further in Conclusions.

EMERGING REGULATORY QUESTIONS FOR AI VOICE

Industry analyst Dean Bubley has highlighted an emerging regulatory issue for AI Voice services. Current debates around the regulation of generative AI, particularly in the context of the EU AI Act, focus heavily on AI-generated content such as deepfakes or synthetic media. However, Voice AI systems increasingly blur the boundary between content generation and real-time communications, particularly in scenarios such as AI receptionists, call-screening assistants, voice translation, or AI agents speaking on behalf of users.

This raises important questions about how existing rules should apply to telecom-based AI Voice services. Under draft transparency provisions in the AI Act, systems that generate AI content may be required to implement mechanisms such as watermarking, metadata tagging or disclosure that outputs are AI-generated.

Real-time voice interactions do not fit neatly into these frameworks. For example, it is unclear whether AI-mediated conversations – such as live translation, call summarisation or AI-driven assistants participating in calls – should be classified as "content" or as part of an electronic communications service.

The distinction could have significant implications depending on the architecture. AI embedded within telecom network functions could be treated as a feature of a communications service, whereas AI delivered via cloud services or CPaaS APIs might fall under broader AI-system regulation. Another

possible interpretation is that AI voice generated directly in response to a user request is treated as communications, whereas synthetic content generated without user interaction is treated as AI-generated content.

As Bubley notes, responsibility may ultimately hinge on who controls the AI model generating the speech, or the input data driving it, rather than where the network signalling takes place. As AI Voice adoption accelerates, clearer regulatory guidance will likely be required to distinguish between AI-generated content and AI-mediated communications.

The European Commission consulted stakeholders on the first draft of the AI Act transparency Code of Practice until 23 January 2026. A second draft is expected in mid-March 2026, followed by further stakeholder consultations before the Code is finalised in mid-2026, with transparency obligations applying from August 2026.

We would welcome feedback from Intelligent Engagement players on how they intend to respond to the EU consultation, and whether CPaaSAA should consider submitting input on members' behalf.

HYPERSCALER STRATEGY: COMPONENT VS. PLATFORM

Hyperscaler AI platforms, such as Google (Vertex AI) and Microsoft (Azure AI), play a dual role in the AI Voice ecosystem, acting as both an enabler and a potential competitor to traditional -aaS players. While listed in the Assembly layer, their significant investment in Component technologies (LLMs and speech models) means they wield influence across the entire value chain.

The strategic challenge is whether a Hyperscaler's offering acts as a supplier of Components or as a comprehensive platform.

Component Supply

The most common initial model is as a supplier of best-of-breed components.

- **Speech-to-Text & Text-to-Speech**. Offering highly accurate, scalable, and multi-lingual foundational models for transcription and voice generation.
- **LLM/Reasoning.** Providing access to state-of-the-art Large Language Models (LLMs) which underpin the conversational intelligence and decision-making capabilities of AI Voice applications.
- **Infrastructure.** Supplying the core compute and networking infrastructure (e.g., NVIDIA chipsets) necessary to run resource-intensive AI models at scale and speed.

This role enables -aaS players to focus on their core strength: packaging these components into vertical, industry-specific solutions that reduce complexity and friction for the end customer.

Platform Competition

The Hyperscalers are also developing increasingly comprehensive, end-to-end platforms that move them closer to the Delivery layer. This shift creates strategic friction:

- **Full-Stack Solutions.** Platforms like Google's Contact Center AI or Microsoft's Dynamics 365 Voice offer pre-integrated, rapid-deployment solutions for specific business use cases (e.g., virtual agents, agent assist).

- **Data and Analytics Lock-in.** By hosting the core data and intelligence layers, these platforms can gather deep customer and business data, potentially creating a "data moat" that is difficult for partners to replicate or compete against.
- **"The Path of the Partner".** The Hyperscaler strategy often involves positioning CPaaS/CCaaS providers as channels or integration partners, which can implicitly limit the partner's control over the total customer solution and, critically, the long-term revenue stream.

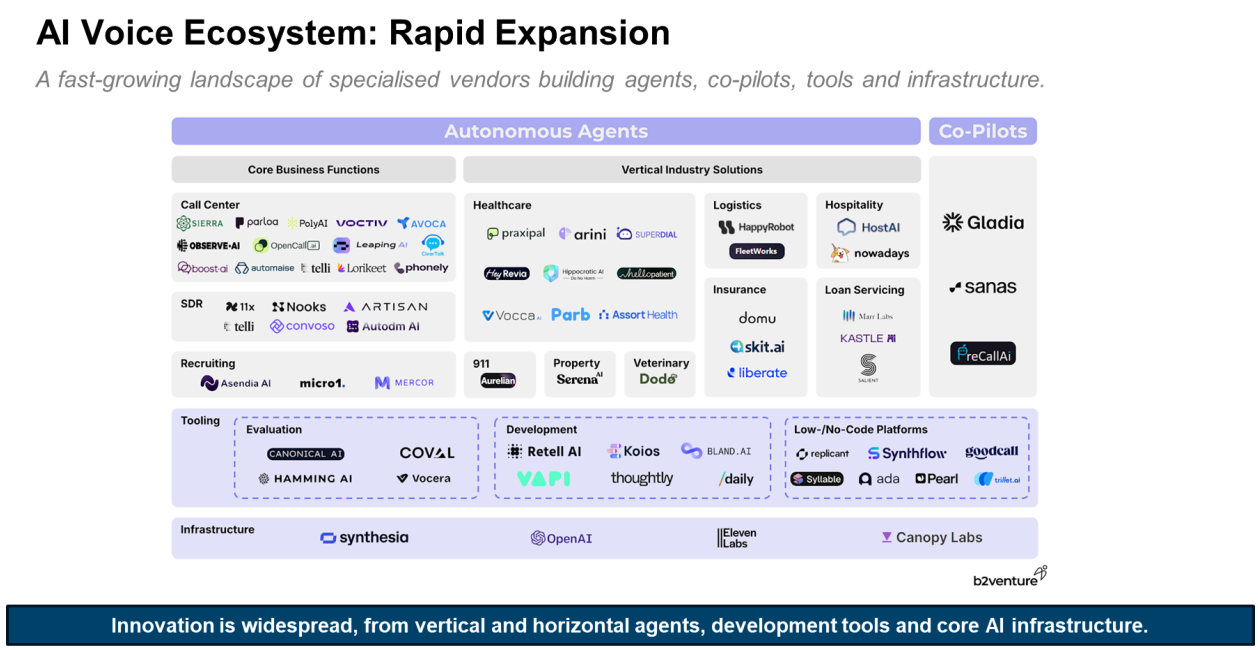
The key strategic control point will be in how the Assembly layer manages the flow and orchestration of value, determining whether Hyperscalers are integrated as foundational component suppliers or as full-solution competitors.

APPLICATIONS: MANY HORSES FOR MANY COURSES

As a final note on the ecosystem, a wide range of players are active in the AI Voice market, offering different tools and specifications:

- Many have a vertical focus - delivering services to e.g. Healthcare, Logistics, Hospitality, etc.
- And others with a horizontal application, e.g. recruitment, call centre, etc

The following chart from B2Venture shows a few examples.



Innovation is widespread, from vertical and horizontal agents, development tools and core AI infrastructure.

Source: https://www.b2venture.vc/stories/the-market-map-of-ai-voice-call-center-agents

NB This market map from Sierra Ventures is even more detailed:
https://www.sierraventures.com/content/voice-ai-market-map

Some of the companies featured in these maps present:

- **Competitive threats**; Agent frameworks and orchestration startups pose competitive risk if they build runtime that bypasses traditional telephony API stacks

- **Partnership opportunities**; component providers (speech, embeddings, synthesis) are strong partners, while orchestration and middleware platforms can integrate with CPaaS/CCaaS to deliver differentiated voice experiences
- **Customer opportunities**: Many of these companies may be users of telephony and runtime infrastructure (CPaaS/CCaaS) rather than direct competitors.

NB Deeper analysis of this ecosystem is outside the scope of this report.

AI VOICE ADOPTION & FORECASTS

AI VOICE ADOPTION

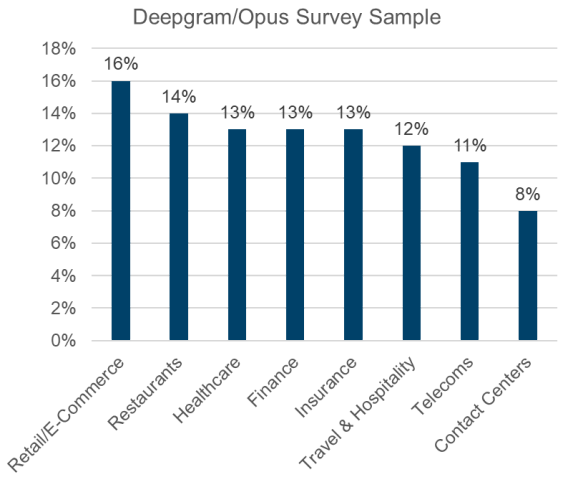
Comprehensive data on the adoption of AI Voice is scarce - it is a very new area. This section looks at some "snapshot" data to gauge order of magnitude indicators. It shows a consistent pattern of demand and fast growth signals across this data, and the combination of scarcity and pace means that there is much still to play for.

US research commissioned by Deepgram in 2025 showed that 92% of 400 US service companies captured speech data, and 56% transcribed more than half of their transactions.

US Service Enterprises Are Expanding AI Voice Use

Most capture speech data today, while investment and agent development are rising.

- Customers: 14% "voice first", 70% "integrated voice", 16% "support specific areas"
- **92% capture speech data, 56% transcribe > ½ transactions**
- 80% use 'traditional' voice agents, e.g. IVRs
- Half of this 80% is service automation
- 15% are actively developing AI Voice agents
- 84% plan to increase budgets in next year
- 72% cite performance quality as critical barrier



Adoption is moving from traditional IVR and transcription toward AI agents, though performance remains a key barrier.

State of Voice AI 2025, Deepgram/Opus Research, 400 US Service Industry Enterprises mix $50m->$5bn

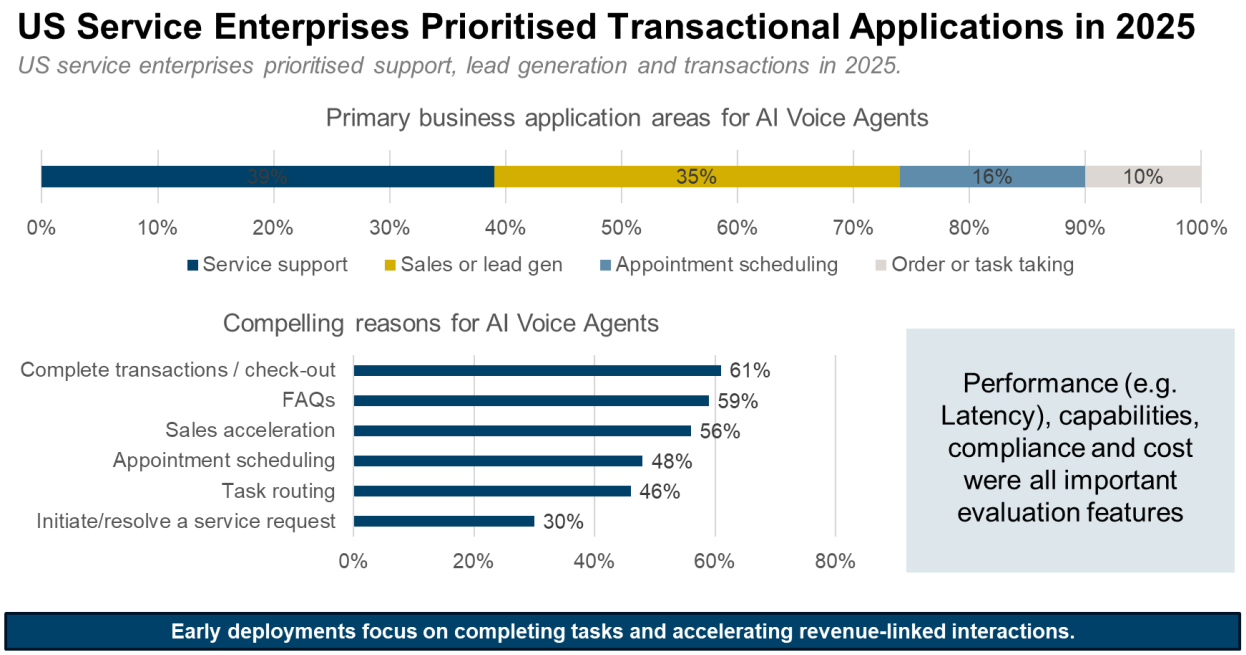
Source: CPaaSAA Research, Deepgram/Opus Research, State of Voice AI 2025

The companies surveyed by Deepgram ranged from $50m to >$5bn by revenue, and were broadly large service businesses.

Deepgram's research supports the notion that AI Voice is mostly being used even today in these relatively advanced businesses in quite transactional areas in service support, sales, lead generation, and appointments.

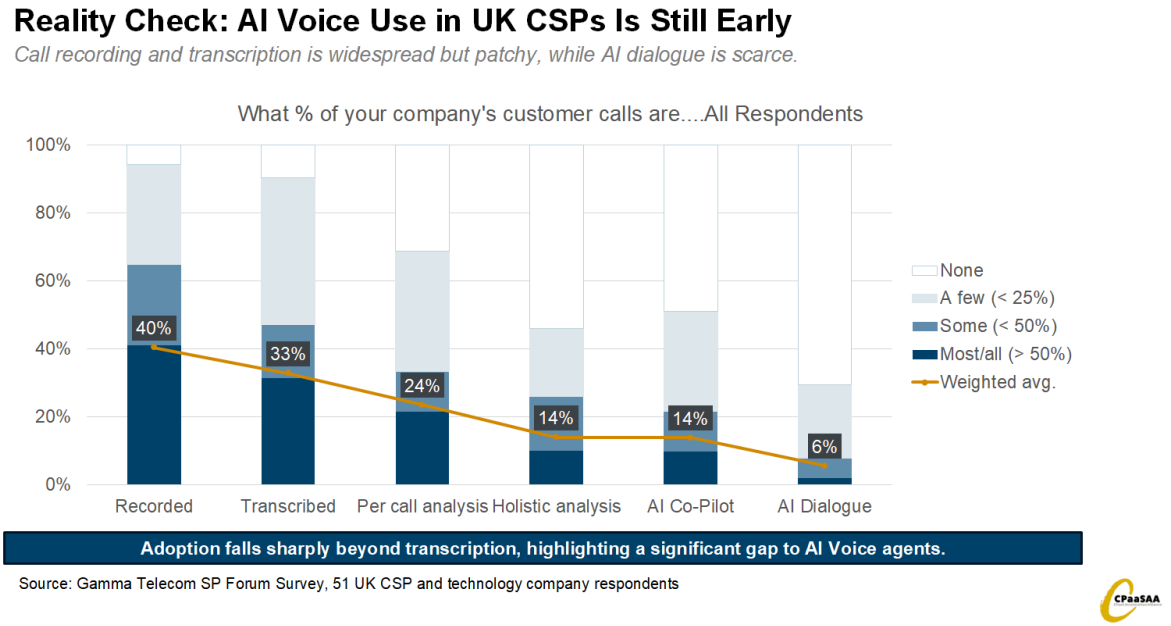
It identified areas such as completing transactions and check-outs, FAQs, sales acceleration and appointment setting as key applications. Additionally, the research described the "most transformational application" as transcription.

While transcription is today considered as "table stakes" *for most AI Voice application providers*, we interpret the finding that transcription is the most transformational application as an indication of the most prevalent current experience *of the majority of customers*. There can be little doubt that the reduced need to write up calls and interviews etc. has proved a blessing to many in terms of saved time and effort.



Source: CPaaSAA Research, Deepgram/Opus Research, State of Voice AI 2025

The following small snapshot survey of UK mid-market CSPs and technology providers in 2026 showed a slightly lower level of usage, with around 30% claiming they transcribed half or more of their calls, This snapshot suggested that, as might be expected, levels of in-depth analysis and AI dialogue (i.e. Voice agents in two-way conversation with customers) were at significantly lower levels than transcription, for example.

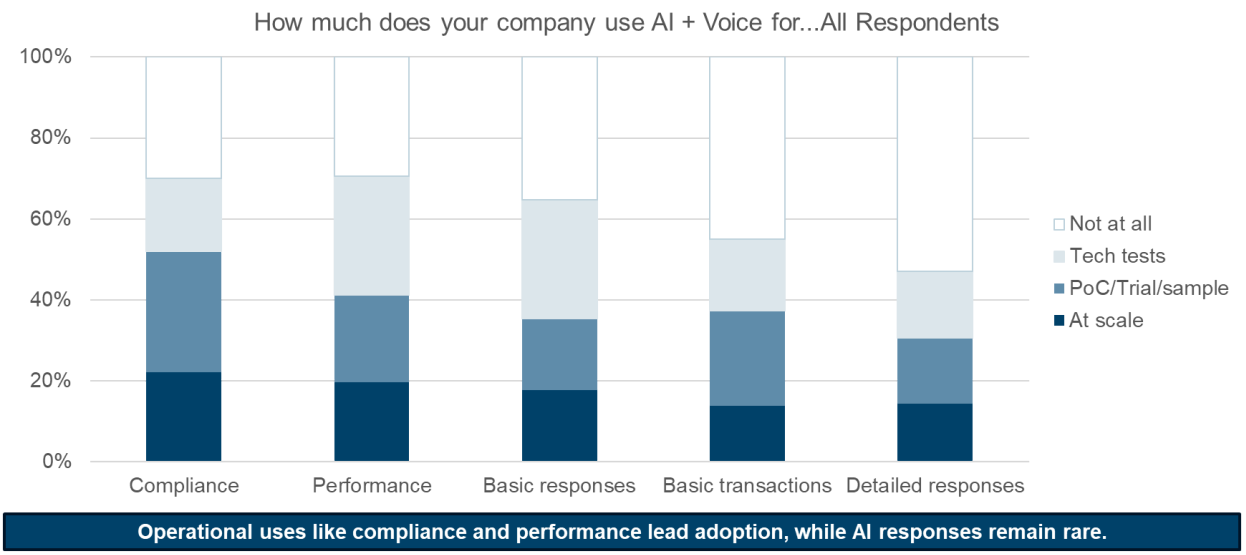


Source: CPaaSAA Research, Gamma Telecom SP Forum Survey / snapshot

In terms of what AI Voice is used for, compliance was marginally the top category, driven by the CSPs in the survey sample, whereas the technology companies tended to use the technology for improving commercial performance areas slightly more than the CSPs. NB This was a small sample, so care should be taken in reading too much into the numbers.

Reality Check: UK CSPs Use AI Voice Mostly for Compliance and Monitoring

Most UK mid-market CSPs remain in trials or limited deployments.



Source: Gamma Telecom SP Forum Survey, 51 UK CSP and technology company respondents

Source: CPaaSAA Research, Gamma Telecom SP Forum Survey / snapshot

Looking at these heuristic measures overall, it looks likely that in large service-focused enterprises, perhaps 1 or 2 in 5 are using AI Voice analytics in markets like the UK and US, but it is unlikely that many are applying it in a truly holistic way across all data available.

The level of application is probably higher in the US than elsewhere, and adoption is likely lower in less service-oriented, smaller and less technical businesses use it less across the board. As one snapshot of wider geographic adoption:

"In the GCC and wider MENA region, organizations are still learning what CPaaS and AI agents are, so their main need is education and clear, real-world use cases rather than just features. When we show how these tools solve concrete problems, like managing hospital appointments, improving banking and government journeys, or turning costly one-way channels into intelligent, two-way engagement, decision-makers are very receptive, especially when solutions integrate with existing systems, meet local regulatory requirements, and deliver measurable gains in efficiency, customer experience and revenue."

Kawthar Awada, Head of Operations and Strategy, CERQLE

Overall, adoption is still early-stage and uneven, and the level of use of AI Voice is currently probably below 10% of its potential, and where it is used, it is mostly in a relatively early stage of maturity.

MARKET SIZE, GROWTH AND FORECASTS

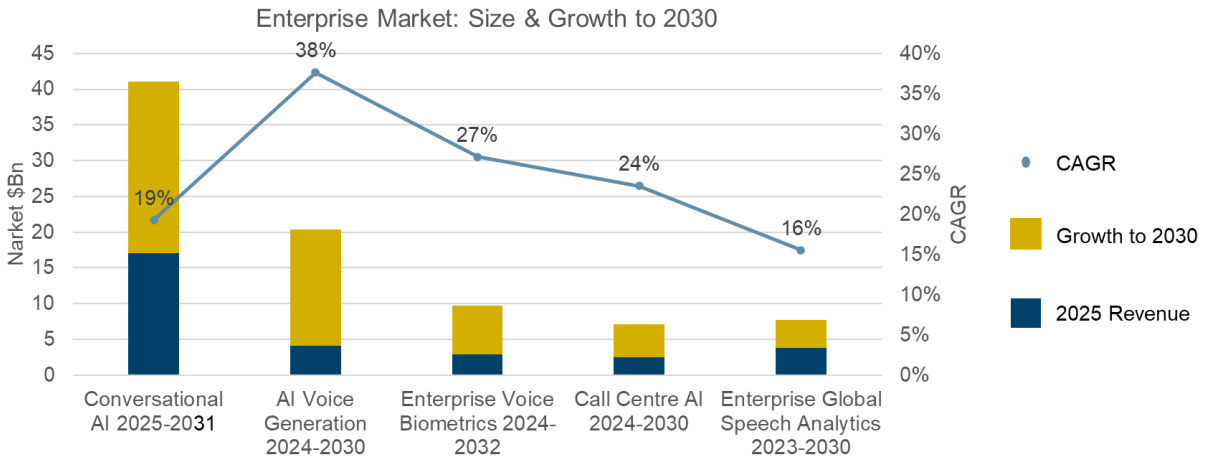
AI Voice is still at an early stage of development and as we have seen, information about the extent of its adoption is still relatively scarce. However there is first-hand evidence from CPaaSAA members that the market is growing very rapidly:

- Data in RingCentral's recent earnings announcement, that its AIR AI Receptionist Q4 2025 sales were up *"44% sequentially"* quarter on quarter, and customer numbers for its ACE (AI Conversation Expert analysis product) were up *"144% year on year"*, according to COO Kira Makagon.
- Speechmatics CMO Lauren King said *"Our real-time usage grew 4x year-on-year...(and) Nordic real-time usage grew 10x, Arabic 6x"*

Market estimates vary by definition. The figures below are indicative third-party market sizing estimates, and illustrate adjacent value pools and growth direction rather than a precise TAM for AI Voice. Perhaps the most striking feature of these numbers is the commonality of strong Compound Annual Growth Rates (CAGRs), ranging from 16% to 38%.

Market Signal: Analysts Forecast 16-38% (CAGR) in AI Voice

Forecasts show strong growth across conversational AI and voice technologies.



AI Voice generation leads with ~38% CAGR, indicating rapid enterprise adoption.

Sources include: GrandView Research, Big Sur, Imarcgroup, Fortune Business Insights, Precedence Research, Markets & Markets

Sources: CPaaSAA Research, Analysts listed

The current market sizes cited range from $2-15bn depending on application definition, with "Conversational AI" forecast to grow to over $40bn in 2031. Some of the numbers we've heard from members are significantly in excess of these forecasts. However, while our members are highly ambitious, they are equally cautious about publicly sharing their views on the market potential.

As further context, in the State of CPaaS 2025 report, we estimated 2025 market sizes as follows:

- $32bn CPaaS market, $7Bn CCaaS market, and $67bn UCaaS market
- $300Bn Customer Care ecosystem
- $600Bn Digital Marketing ecosystem

Another potential yardstick is the global CRM market, worth over $100bn today. There are suggestions that some of the AI Voice and analytics applications could eat into this adjacency - although equally, CRM and data players will also eye these revenue pools.

Even modest participation in these pools implies revenue potential in $ billions / tens of billions.

THE FOUR VALUE LAYERS OF AI VOICE

Where value sits: AI Voice economics are shaped by four layers ranging from infrastructure to outcomes.

AI Voice reorganises where value sits in live enterprise conversations. The economic centre of gravity is moving between the following layers beyond Transport.

1. **Transport & Usage**. Minutes, tokens and media streaming remain the foundation. Telephony APIs, routing and AI processing are typically billed on consumption. This layer is essential, but exposed to price pressure and efficiency gains that can reduce billable volume.

2. **Intelligence**. Analytics, copilots, sentiment detection, biometrics and compliance monitoring make voice computable. Revenue may still be metered per minute or API call, but the value driver shifts to interpretation, risk management and quality improvement. Differentiation expands at this layer.

3. **Outcomes**. AI Voice increasingly links to operational performance: call containment, conversion uplift, reduced handling time, improved compliance. Pricing models extend beyond usage to seats, workflows or enterprise contracts. Budget ownership can move from telecoms to operations, customer experience or revenue functions.

4. **Structured Conversation Data**. When conversations are captured, structured and governed, for example through vCons, they become reusable assets. Structured data strengthens analytics, supports audit, improves model performance and embeds the platform deeper into enterprise systems. This layer increases switching costs and strategic leverage.

Strategic implications

- The largest long-term margin pools sit above transport, and Intelligence revenues are likely to scale more sustainably than transport and usage

- Outcome-based pricing reshapes budget conversations, and ownership of structured conversation data strengthens defensibility. Efficiency may reduce minutes, but intelligence increases value per interaction.

- Control of runtime determines who captures value across these layers

The opportunity for Intelligent Engagement players is to compete in the added value layers.

How might it play out, in general?

- 2026–2027: transcription, copilots, compliance

- 2027–2028: containment & workflow automation

- 2028–2030: outcome-based monetisation & runtime consolidation

INTELLIGENT ENGAGEMENT: MARKET OPPORTUNITIES

WHAT'S DRIVING CUSTOMER ADOPTION TODAY?

Reviewing the AI Voice case studies in the CPaaSAA Case Directory:

- For most clients, the main near-term priority is coping with higher demand and the inefficiencies it exposes.
- Cost savings are a strong, but slightly secondary driver, often a result of addressing the first two.
- Security and compliance are the top priority in regulated sectors (e.g. finance and healthcare), but for others, they are important but not always the immediate trigger for change.

This broadly matches the priorities given in the Contact Center Helper survey on What Contact Centres Want (2025) - "What would be your biggest driver for implementing AI?"

- 33% Contact Reduction (Dealing with High Volumes)
- 32% Customer Satisfaction
- 22% Cost Reduction

"Our clients are no longer asking how to cut costs, they are asking how to scale their brand's reach without the linear expense of headcount."

Chris Angus, VP CPaaS & CX Expansion, 8x8

"What's really changed is what buyers expect from them (AI Voice agents). Not AI that handles everything, but AI that knows which interactions actually need a human. The ROI case is then built on that triage: fewer unnecessary escalations, humans focused where they add real value."

Lauren King, CMO, Speechmatics

However, in the longer-term, value may emerge not from handling conversations, but from analysing and acting on them.

BUSINESS SEGMENT NEEDS

The business sector is extremely diverse, but some identifiable characteristics help to explain some of the differences in needs for AI Voice:

- **Industry Sector**. In general, companies that use voice more than others are likely to be faster adopters of AI Voice. This trend can be seen in the research and uptake metrics, showing that e.g. telecoms, financial services, healthcare, etc., are high among the adopters
- **Strategic, sovereignty, regulatory, and compliance priorities.** Similarly these and other company types that are under higher legal or regulatory scrutiny due to the nature of their business are likely to be early adopters.

- **Innovation/Disruption Mindset**. Lawrence Byrd, Communications APIs, Vonage, described three archetypes of businesses engaged in AI Voice. *"Digital enterprises and disruptors tend to be the most advanced in application, with Industry Platforms a close second, and Traditional Enterprises third. While there are many more businesses in this last category, their rate of adoption is much slower."*

- **Call Centres** are clearly an area in which AI Voice is a priority across these types, and the Call Center Helper survey suggests that automation and assistance are the top near-term needs (with over 60% respondents looking at this application of AI in the next 12 months) with analytics and predictive applications close behind (40-50%).

- For **Small and Medium Businesses (SMB)** there is often a substantial difference in the level of skills and expertise available in IT and communications functions, so many tend to choose more packaged options that are designed to meet their main areas of need, and focused providers are evolving in this space to address these needs (see Industry Ecosystem).

"SMEs are asking, 'what can you do with AI for me?' For small and medium sized enterprises, AI could make them the most successful they could ever be… and is also simultaneously a huge threat."

Kevin Nethercott, CEO, Tresic

Another approach to SMBs is demonstrated by Finland's LastBot, a fully self-service, low-touch, low cost AI customer service bot that onboards itself from an SME's website and data, and that sees telecoms operators as a potential go-to-market partner.

"The quality of AI Voice now available from Google Gemini is amazing, and we are bringing an AI Voice agent to market using that now."

Janne Timonen, CEO, LastBot.

Vertical Specialisation: Creating Defensible Utility

Beyond horizontal platform offerings, there is an opportunity for CPaaS/CCaaS players to rapidly explore and develop **deep vertical specialisation** in regulated industries (e.g., Finance, Healthcare, Government). This involves delivering highly tailored, compliant, and pre-integrated solutions that bake in domain-specific compliance, identity verification, and workflow automation. This strategy transforms a generic platform into an indispensable, high-value vertical utility, dramatically raising customer switching costs and enabling premium, intelligence-based pricing.

As an example of the benefits of vertical specialisation, Speechmatics CMO Lauren King said:

"General-purpose models will keep improving. That's not the threat it might look like. The industries where voice AI is gaining real traction, healthcare, financial services, regulated public sector, have requirements that go beyond what a general model is optimised for. The gap isn't subtle either. In a clinical setting, the difference between "hypertension" and "hypotension" matters. Our medical models, trained on over 16 billion words of clinical conversation, deliver keyword error rates 70% lower than general-purpose alternatives, with 96% medical keyword recall in production. That's the result of focusing on real problems over impressive demos."

INDUSTRY NARRATIVE AND PACKAGING

AI Voice is a complex field when looked at technically or holistically, as this report shows. For customers and the ecosystem, the key success factor will be to make the story simple, business outcomes based, and easy to follow. It's hard to get a lot of people in a market moving the same way fast, but it is a lot easier if they are all hearing clear and consistent messages from every direction.

This is why we are developing and promoting the concept of Intelligent Engagement with all our members, to try to create the overarching benefit led narrative. We see AI Voice as an important new strand in that narrative, and will continue to work with our members to develop it. As one example, we like RingCentral's intuitive 3A's approach "Automate, Assist, and Analyze" as a way to describe the core propositions for enterprise.

Outcome-based pricing is likely to take time and work to evolve. As Bertrand Pourcelot, MD, UCaaS for CSPs & MSPs, Enreach says: *"Everyone likes the idea of paying for outcomes instead of capacity or minutes or tokens. But outcome-based deals are still rare. Service providers want predictable revenue, so discussions often fall back to fixed price or per-minute models, because the criteria and business logic for true outcome pricing just aren't mature yet."*

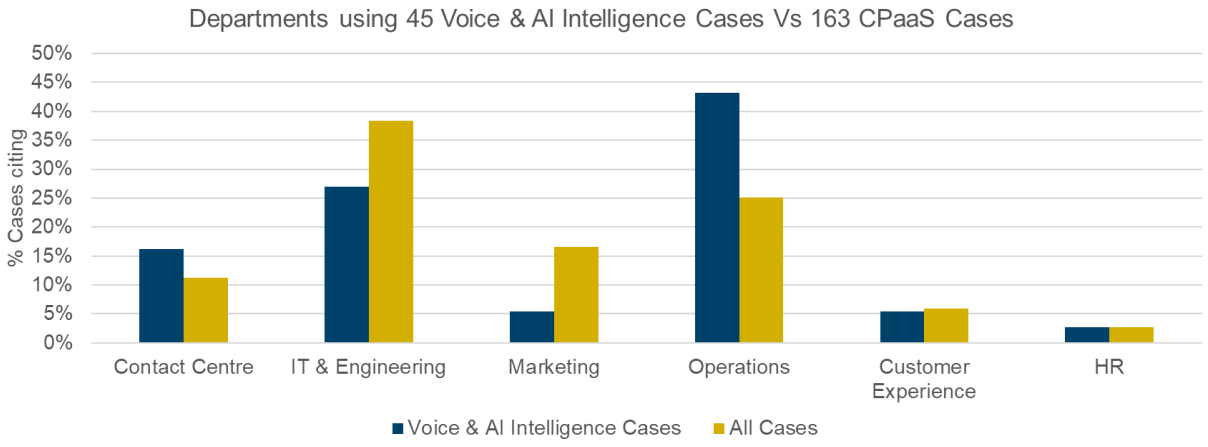
Operations, CX, and revenue leaders are likely to have a stronger preference for outcomes based pricing than IT managers. This may be an opportunity for Intelligent Engagement players.

In the CPaaSAA Case Directory "Operations" features most strongly among the classification of buyers or departments currently using AI Voice. This classification is based on either a direct statement or an interpretation of how each case study is phrased. In the types of business using AI Voice (e.g. technology and service industry organisations) operations covers a broad range of roles responsible for making the core business work, rather than a specific function (e.g. call centre, marketing).

This suggests a) that our members are already finding success with the most attractive buyer types, and that b) with the right formula, there could be scope for outcomes based pricing.

AI Voice Adoption Is Often Driven by Operational Functions

Operations and contact centres lead AI voice deployments in the case data.



Operational teams lead adoption as scaling and efficiency gains are key need drivers.

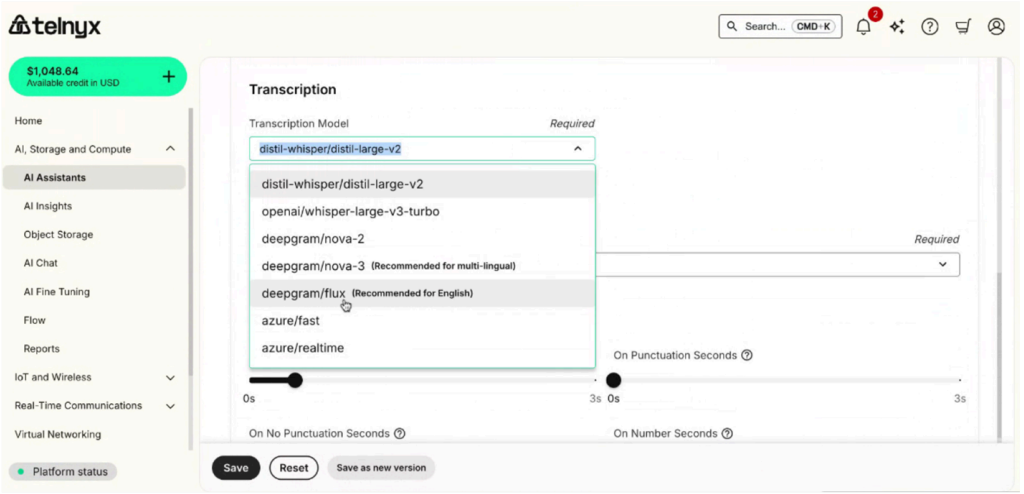
Source: CPaaSAA Case Directory

PLATFORM CHOICES & MODEL NEUTRALITY

One role that Intelligent Engagement players can take is to make it simple to integrate the components of AI Voice, giving choices of models, tools and services in each clients' set up. Ian Reither, COO of Telnyx, demonstrated how a customer could do this.

Example: AI Voice Platforms Orchestrate Multiple AI Models

Intelligent Engagement players enable transcription and AI services from several providers.



Modern Intelligent Engagement platforms enable model-neutral orchestration across AI providers.

Source: Ian Reither, Telnyx demo

As Ian put it, *"We built the real-time voice infrastructure first. Many AI voice platforms stitch together third-party services. Because we own the core layer, the AI models are modular. Customers can swap STT, TTS, or LLMs without rearchitecting their stack or introducing new failure domains."*

Other communications platforms are also adopting this approach. For example, Sinch's Voice Relay infrastructure is designed to operate with multiple AI providers, allowing enterprises to swap models without changing their voice infrastructure.

"We do not believe in locking customers into a single agent model, proprietary data layer, or closed ecosystem."

Daniel Morris, Chief Product Officer, Sinch

Modern communications platforms increasingly expose model-agnostic orchestration layers that allow enterprises to choose between different speech recognition, language and voice synthesis models. Demonstrations at Enterprise Connect 2026 showed platforms enabling developers to dynamically select models during live calls. This approach reinforces the importance of model neutrality in enterprise AI Voice architectures.

TELCO NETWORK AI VOICE

AI in every call

- *"There is a category-level opportunity for operators to define a new execution environment for communications, particularly where trust, identity, quality of service and sovereignty matter. Trust in this context should be seen as part of the product value proposition, not only as a compliance requirement."* Jonathan Abrahamson, Chief Product and Digital Officer, Deutsche Telekom.

Deutsche Telekom's Magenta AI

At MWC26 in Barcelona, Deutsche Telekom announced the launch of its Magenta AI Voice Assistant

- *"With our Magenta AI Call Assistant, we are the first in the world to offer these network-based AI functions. We remove barriers: no apps, no special devices, no technical complexity. AI becomes simple, intuitive and available to everyone."* Abdu Mudesir, Board Member, Product & Technology, Deutsche Telekom

Deutsche Telekom's announcement followed T-Mobile's earlier announcement of its Network-native AI Voice feature,

DTAG's Magenta AI is voice activated by saying "Hey Magenta", and it says no voice content is analysed or stored otherwise. T-Mobile's US service is activated during a normal call by keying *87. It is positioned as "agentic AI in the network" and framed as the first of more network-level AI features. The announcement positioned the service as a beta trial, and says:

- *"By dialing *87*, beta participants will be able to speak to anyone in more than 50 languages, at no cost during the beta and for as long as they need. From family calls to making reservations abroad or handling customer inquiries, even the smallest mom-and-pop shop will soon be able to confidently answer the phone, serve new customers, and avoid missed opportunities caused by language barriers."*

It's notable that the T-Mobile release mentions a "mom-and-pop shop", and also that it is "no cost during the beta". This suggests that it may be a service that T-Mobile envisages marketing to business customers, and that it is exploring whether or not to charge for the service, It is not publicised which, if any partners are involved in the trial, though T-Mobile uses Mavenir in its IMS, and executives have described the app as working with a range of AI partners.

LG Uplus "ixi-O" AI Call Assistant

A critical distinction is emerging between "**AI applied to voice**" and "**network-native AI Voice**." In the first model, AI capabilities such as transcription, translation or summarisation are attached to the call path as external services. In the second, AI orchestration and policy control become part of the communications execution environment itself. This distinction is strategically significant: the former enhances the service, while the latter has the potential to redefine the control point of communications.

LG Uplus' ixi-O assistant illustrates a third emerging architecture between these two models. The service operates as an AI call assistant that can be invoked during live calls (for example by saying "Hey ixi") and provides capabilities such as real-time information retrieval, call transcription and summarisation, voice-phishing detection and AI call answering.

It combines LG Uplus' proprietary AI technology with Google's large language models and voice-to-voice processing so that spoken questions can be answered directly within the call without interrupting the conversation flow. At MWC26 the company's CEO Bumshik Hong showcased ixi-O Pro, an enhanced version designed as a hyper-personalised AI call agent that understands the user's communication context across calls, messages and schedules and proactively provides information during conversations.

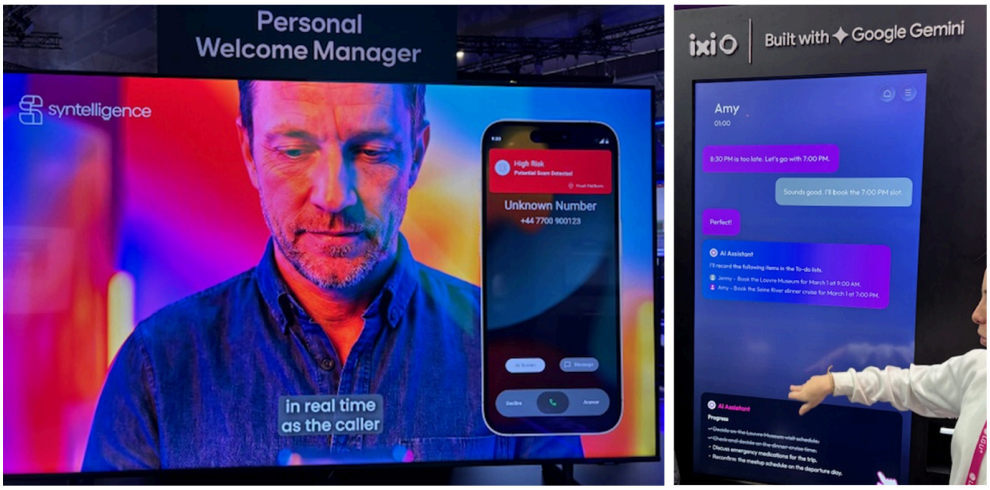
Architecturally, ixi-O differs from many cloud-based AI voice services by running much of the AI processing on the device, with cloud services used selectively when the assistant is invoked. This design reduces latency and keeps most voice data on the device, improving privacy and security while still allowing access to external AI models and online information when needed.

The trade-off is that device-centric architectures may be more constrained in compute capability and orchestration compared with fully cloud-based AI platforms, which can run more complex models and integrate more deeply with enterprise workflows. ixi-O therefore sits somewhere between a device-native AI call assistant and a telecom-integrated AI voice service, rather than a fully network-native AI voice platform. Public information on standalone pricing is limited, although related "AI Phone" packages have been reported at roughly KRW 7,700 (around $6) per month, sometimes waived for premium subscribers.

A similar direction was visible in demonstrations from Syntelligence.ai, a collaborative venture between SK Telecom, Deutsche Telekom, e&, SoftBank and Singtel. At MWC26 the group demonstrated an AI call screening assistant designed to analyse and manage incoming calls in real time. Like ixi-O, these approaches bring AI directly into everyday voice interactions, although the architectural question of where orchestration ultimately resides – device, cloud or network – remains central to determining where long-term control and value will sit.

Some Assistants Use a Mix of Device and Cloud/Network AI

AI Voice assistants such as LG Uplus' ixi-O and the Syntelligence PoC bring AI directly into live calls.



These models blend device, cloud and network intelligence, raising new questions about where control and value will sit.

Source: Images from MWC26 Demos, CPaaSAA Research

Design and pricing considerations

For operators, product simplicity may prove as important as model capability. AI Voice solutions that minimise friction for end users and enterprises requiring no separate app, no device dependency and simple invocation within the call itself, may see faster adoption. Carrier-grade simplicity and trust frameworks can therefore become meaningful differentiators.

How such services will be priced is not yet fully clear. Services that consume external LLM tokens clearly have a cost to the provider - it is not a free lunch. Some services, like translation, may lend themselves to a charge per use for activated users.

Others may suit a per month charge or inclusion in higher value packages. This model may suit assistants, call enhancement services, and possibly anti-scam services.

Anti-scam services, where a listening AI detects signals for malicious or fraudulent intent could be very attractive to many customers. However, there are complexities, for example what happens if the service misses a scam. Is the operator liable? Making expectations clear and setting fair Terms and Conditions will be a critical enabler.

Additionally, many consumers are used to free AI Voice, such as Siri, Alexa and some in-car services. Operators should certainly not overprice AI Voice services as it will restrict take-up and create negative value for money and trust impressions.

The best pricing may therefore follow what is possible in practice, local market regulations and consumer expectations..

Other operator and partner announcements

Embedding real-time translation and other services like transcription and assistants in the mobile network / voice path has ignited interest among other operators in developing similar service. For example, Radisys announced a partnership with Rakuten just before MWC which implies a Voice AI angle to its agentic AI proposition, and we expect to see other announcements soon.

In terms of telco network oriented AI Voice partners:

- Adnan Saleem, CTO, Radisys says of operator demand: *"We're seeing tremendous demand globally with tier-one operators. Operators are pushing through to real deployments with serious resources behind them in multiple markets. They're asking for transcription, summarisation, fraud detection, virtual assistants, and live language translation. These are new, monetisable voice services that go well beyond basic telephony."*

- Alianza envisions a three layer industry model: Infrastructure (components), Orchestration and Experience. It argues that telecoms operators should modernise their orchestration (runtime) control for AI-infused communications services. Chief Product and Technology Officer, Dag Peak says of its Intelligent Communications Fabric: *"we're creating a platform where every call becomes an opportunity for intelligent, revenue-generating experiences."*

- Nokia's Thiery van de Velde, Global Core Network Pre-Sales Lead demonstrated Nokia's IMS based in-call translation and transcription capabilities at MWC26.

Why it matters where the intelligence is orchestrated in Telco Network AI Voice

In Telco Network AI Voice, the operator always runs the call, and determines how AI might be triggered (*87 etc.), the price, terms and conditions etc, and the AI processing, e.g. ASR, translation and TTS may sit on cloud infrastructure in any scenario.

Strategically, it still matters where the intelligence orchestration layer sits - that is, the runtime that decides which models are used, how calls are handled, where data flows, how policies are enforced and when escalation occurs.

If that orchestration sits in e.g. a hyperscalers' domain, AI in the network becomes a feature layered onto connectivity. The telco controls access and billing, but the execution logic depends on that hyperscaler. Costs remain variable. Switching suppliers requires re-integration. Innovation follows the hyperscalers' roadmap.

If orchestration is embedded within IMS, media servers or edge environments, even if external AI models are used as components, the operator controls execution. Models become interchangeable inputs rather than architectural anchors. The telco can manage the cost curve, negotiate from a stronger position with suppliers, and design new services without restructuring the stack each time.

Either way, embedding AI in the call path improves the offer, but controlling the orchestration layer changes the economics and the power balance.

T-Mobile's careful silence on who they've partnered with for their eye-catching beta suggests that they are mindful of this trade-off. It will be interesting to see what approaches other telcos take as the space develops.

We expect to see more announcements and trials in this area, and will update this analysis subsequently.

Telco Call Centre AI

KPN and Deutsche Telekom have also been reported as in partnership with ElevenLabs, although the tenet of these announcements appears to be aligned with in-house customer contact, and possibly also creating enterprise-focused applications.

According to Hans Söhngen, Managing Director, KPN Ventures, *"We believe it is more normal for humans to talk than type into a phone. There are great opportunities in AI Voice, and that's why we are working with ElevenLabs."*

BT has recently launched *"Neura AI …BT's Intelligent Virtual AI Agent, built on top of our award winning Global SIP platform."* Its factsheet says that Neura uses AI to automate interactions and integrates with existing systems seamlessly. It claims *"AI at the Network Layer - Enterprise grade AI voice delivers secure, AI enabled voice and data over BT's global fabric, backed by end to end SLAs that assure performance, reliability, and compliance"* and that it has pre-built LLM and domain AI connectors, with *"usage based, committed or custom pricing models".*

Neura AI targets "repetitious activities", including banking card activations, store information, travel and appointment booking. As it integrates with SIP, it is less like T-Mobile's "AI on any call" and more of a call centre / CPaaS style application.

ANALYTICS

Adding analytics and intelligence to voice with AI is a natural step. The following case study summaries are two examples of where analytics can be used on specific call and engagement instances. The benefits are to increase quality, scalability, and growth, and reduce handling times.



Source: CPaaSAA Case Directory

Many platform vendors are now positioning conversation intelligence as a core capability of communications systems rather than a separate analytics layer. Demonstrations at Enterprise Connect 2026 showed analytics engines monitoring 100% of customer interactions in real time and generating insights during live calls.

There are further examples of analytics and automation and their benefits in the following section on vCons..

VCONS: CONTEXT, COMPLIANCE AND COMPETITIVE ADVANTAGE

vCons are a new IETF (Internet Engineering Task Force) standard for data relating to **v**irtual **con**versations - hence **vCon**. It was pioneered by Thomas McCarthy-Howe, CTO of vConic and CPaaSAA member, and others, as a means to help store data in a way that both humans and AI can intelligently and compliantly access.

vCons present a significant new opportunity to further develop data and AI Voice services by turning everyday conversations, calls, chats, and meetings into structured, trustworthy, portable data that respects privacy and can be put to work by AI.

In Thomas's words *"vCons turn messy, unstructured conversations into clean, structured data. They surface the dark operational data that's hidden in everyday customer interactions and convert it into high-quality 'robot food' for AI, so you can finally use those conversations to run the business better."*

The vCons standard provides two fundamental opportunities for Intelligent Engagement players. It:

- Standardises the process of designing and managing such data processes, saving practitioners the time of resolving this core challenge
- Makes it possible to ensure compliance with global and national legislation such as GDPR despite the complex interconnection of data points involved in a holistic analysis

Standards: vCon Structures Conversations for AI at Scale

The IETF vCon standard converts calls, chats and meetings into structured, portable data.

vCons turn everyday conversations, calls, chats, and meetings into structured, trustworthy, portable data that respects privacy and can be put to work by AI.

They enable conversation intelligence, sales and product optimisation, **compliance** and security

Good standards scale industries

Standardised conversation records enable compliant analytics, AI services and industry scale.

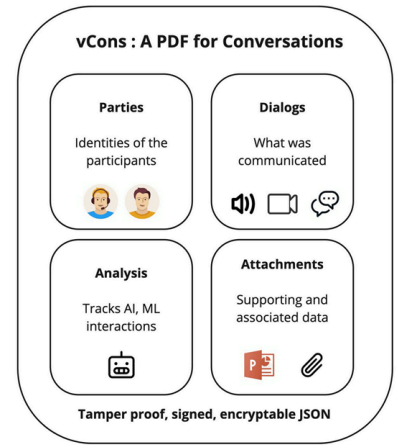
Source: CPaaSAA Research

"A vCon is the container for data and information relating to a real-time, human conversation. It is analogous to a [vCard] which enables the definition, interchange and storage of an individual's various points of contact. The data contained in a vCon may be derived from any multimedia session, traditional phone call, video conference, SMS or MMS message exchange, webchat or email thread."

Source: IETF, The vCon - Conversation Data Container - Overview.

Architecture: vCon Creates a Structured Record of Conversations

The standard captures participants, dialogue, AI analysis and attachments.



A tamper-proof JSON record makes conversations portable, analysable and AI-ready.

Source: CPaaSAA research, Conserver.io

Using vCons within a process called SCITT (Supply Chain Integrity, Transparency, and Trust) enables an organisation to store customer conversation data from all types of media, check it, analyse it, act on it, and if required, remove it in a GDPR-compliant manner.

"Transcripts are great, but few people look at them and even fewer act on them. But imagine if you made sure all the actions you talked about on each call were automatically tracked and acted on. That's what we can enable - for pretty much any business."

Robert Galop, CMO, Tresic

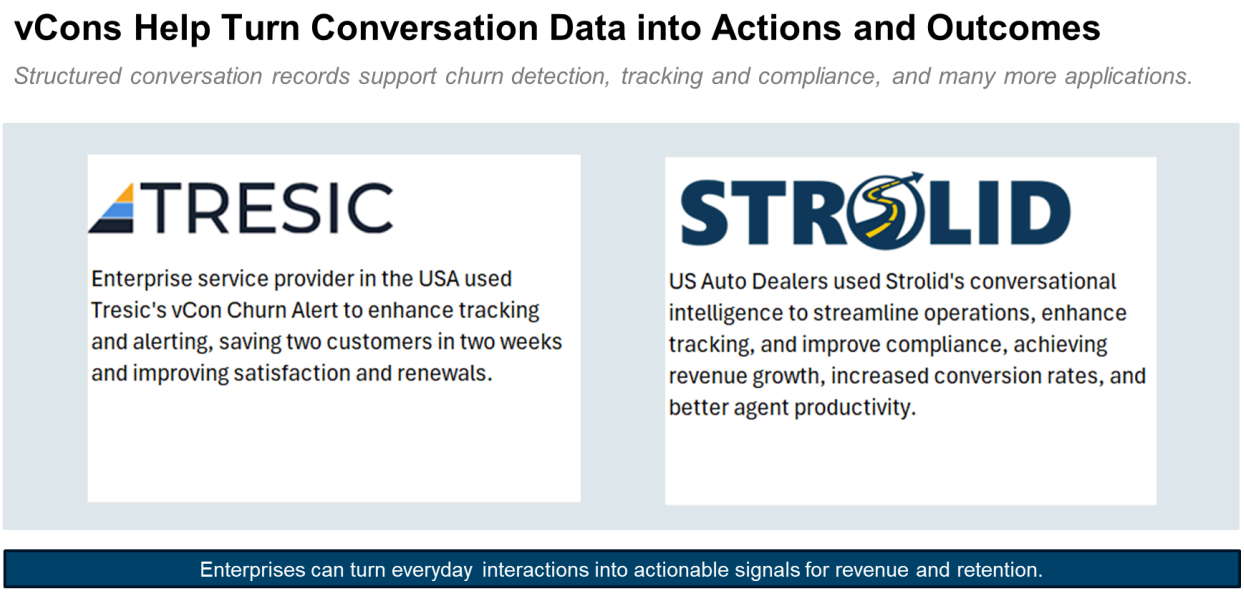
vCons enable:

- **Conversation Intelligence & Insights**
    - **Providing full context of interactions**, capturing structured conversation data, including transcripts, sentiment, and metadata
    - **Customer Journey Mapping** – linking multiple conversations into a unified timeline to understand the lifecycle and needs of each customer.
    - **Agent Performance Insights** – Measuring talk-to-listen ratios, script adherence, and keyword usage to improve training and coaching.
- **New Monetisation Opportunities** for Operators & Providers
    - **Premium Conversation Analytics Services** – Selling packaged insights dashboards or APIs to enterprise customers for sales optimisation, CX measurement, or compliance monitoring.
    - **Targeted Upsell & Cross-Sell Triggers** – Using vCon data to identify buying signals and automatically launch relevant offers.
    - **Integration-Ready Data Products** – Offering conversation intelligence feeds that customers can plug into CRMs, marketing automation, or analytics tools.
    - **Vertical-Specific Solutions** – Creating specialised versions for healthcare, finance, or retail, with domain-specific analytics and reporting.
- **Compliance & Security**
    - **Automated Redaction & PII Handling** – Detecting and masking sensitive information like credit card numbers, health data, or account credentials in real time.
    - **Immutable Conversation Records** – Storing vCons in a tamper-proof format for legal, regulatory, or dispute resolution purposes.
    - **Multi-Jurisdictional Compliance** – Applying region-specific retention policies, consent capture, and language requirements automatically.
    - **Secure Sharing & Access Control** – Allowing only authorised personnel or systems to access vCons through secure APIs or encrypted storage.

For example, CPaaSAA members, including Tresic and others, are using vCons to build:

- **Conversation Co-Pilots** – Actionable Conversation Intelligence in Real Time - using vCon data to summarise interactions and actions across conversations, flag churn risk, coach sales teams and check compliance
- **Insights Hubs** – Centralised Conversation Intelligence & Trend Discovery - turning large volumes of vCons into business intelligence that can be monetised, operationalised, or shared with enterprise customers (e.g. Product Feedback Radar, Market Sentiment Dashboard, Channel Performance Analysis, Upsell Opportunity analysis)
- **Developer Studios** – Making Applications and Experiences Better with Insights from Conversation Intelligence, letting operators, MSPs, Application Developers, and enterprises build their own apps and workflows on top of vCon data - turning conversation intelligence into new products and services (e.g. Vertical Compliance Apps, CRM Auto-Update Connector, Automated Dispute Resolution Tool)

The following vCon case studies demonstrate the additional insight and value unlocked by analysing information across all elements of the conversation, typically identifying issues and actions invisible through other systems (e.g. traditional CRM).



Source: CPaaSAA Case Directory

As well as the new services they enable, vCons can help CPaaS/Intelligent Engagement players to ensure that the Intelligence value-add that they provide is superior to generic AI, because they:

- **Save time and improve quality and security** by providing a clear and standardised template to structure conversational data
- **Build in structured context to the data** - an essential building block for commercially successful applications of AI
- **Improve compliance**, a benefit for customers and key mitigation of enterprise risk

TRANSLATION

Call translation is a natural progression for AI Voice, given the rapid development of translation abilities of many LLMs. (NB note also Telco Network AI Voice section.)



Source: CPaaSAA Case Directory

NB. Call Containment = automated call completion without escalation.

The emergence of real-time voice translation in both enterprise platforms and telecom networks indicates a broader shift toward "AI in every call", where AI capabilities are invoked dynamically during conversations rather than embedded only in contact centre applications.

BIOMETRICS & SENTIMENT ANALYSIS

AI models can be used to identify speakers and analyse the emotional tone of voices etc. The following is an example of biometrics used in identity.



Source: CPaaSAA Case Directory

"The real shift is from scripted responses to systems that can interpret intent, maintain context, and handle the variability of real human speech, including accent, emotion, and ambiguity."

Melvin Maahen, Senior Product Manager, 8X8

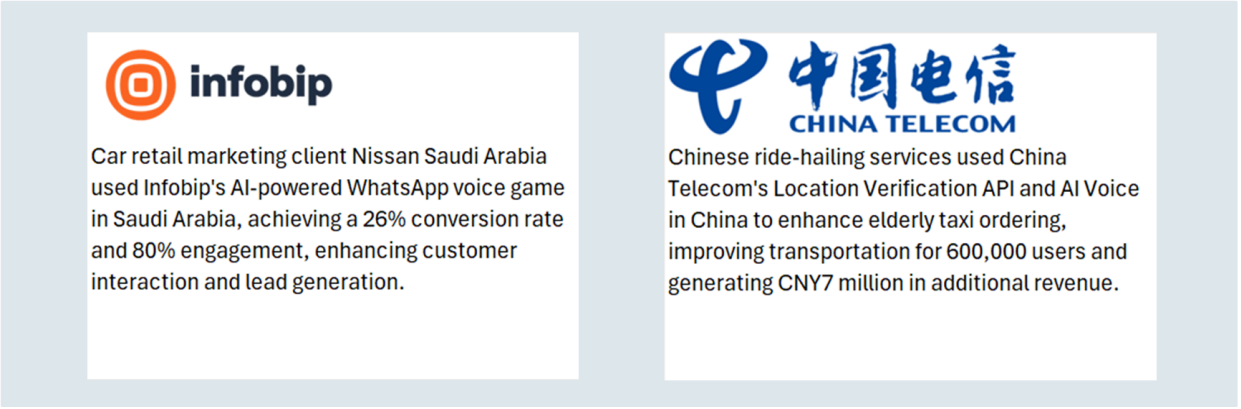
There are many other companies specialising in sentiment analysis including Speechmatics, Deepgram, Verint, Beyond Verbal, Call Miner, AssemblyAI, NICE, Uniphore and Cogito.

INNOVATIVE APPLICATIONS

AI Voice is beginning to support more experimental and differentiated services, beyond efficiency or automation.

AI Voice Is Enabling New Customer Engagement Models

Companies are combining voice AI with messaging and network capabilities.



Source: CPaaSAA Case Directory

For example, AI Voice can be integrated into messaging platforms to create interactive customer engagement experiences, in this instance a WhatsApp voice game that drives high participation and lead generation. Meanwhile China Telecom's example demonstrates how AI voice combined with telecom network capabilities, such as location verification, can enable entirely new services, such as simplified taxi ordering for elderly users.

Together these cases illustrate that AI Voice is not only improving existing processes but also opening new forms of customer interaction and service design, particularly where voice, messaging and telecom capabilities intersect.

CONCLUSIONS

Strategic takeaway: AI Voice will redistribute value in communications toward players who control intelligence, orchestration and enterprise integration.

AI Voice creates a new opportunity within Intelligent Engagement. It turns live conversation into computable, governable and monetisable interaction. Intelligent Engagement is the business framing: the shift from fragmented communications toward coordinated, real-time action across customers, partners and workforce. Runtime is the control point within that shift. It determines how conversations execute, how intelligence is applied, how compliance is enforced and where structured data accumulates. The strategic question is therefore not simply who provides AI, but who governs execution, and who captures the value that follows.

INTELLIGENT ENGAGEMENT PLAYERS IN AN AI VOICE WORLD

AI reduces the economic importance of how long people talk, and increases the economic importance of what is understood and done.

AI Voice opens new fields for Intelligent Engagement players. It can be monetised as a per-minute or per-token service, and/or as a platform offering priced per seat, per workflow, or at enterprise level. The pricing model will depend on customer need, application maturity, and market context.

However, while AI may reduce the number of minutes per call, it does not reduce the value of the interaction. In many cases it increases it. That value now depends less on duration and more on the success of the intelligence deployed. The paradigm is shifting from transport-driven value to intelligence-driven value.

Differentiation therefore lies not in access to models alone, but in bridging the gap between theoretical AI capability and operational business reality. Customers need governed execution, integration with enterprise systems, compliance assurance, quality control, and reliable outcomes. This is where Intelligent Engagement players can create defensible advantage.

How to harness AI effectively is now a board-level question. AI Voice is one of the most immediate and visible arenas in which that question must be answered. The opportunity for Intelligent Engagement players is to provide solutions that are more operationally complete and commercially accountable than hyperscaler models alone.

Enterprise platforms such as Salesforce are also embedding conversational AI into CRM environments, reinforcing that the future control point for AI Voice will sit across communications, cloud and enterprise software platforms.

As conversations become computable, the profit pool shifts toward those who orchestrate the intelligence that runs them.

Because intelligence now drives value, the control point moves from pure connectivity to orchestration. If communications platforms remain at the transport layer, value will migrate upward in the stack. If they own the runtime (the execution environment that governs models, policy, integration and escalation) they remain central to enterprise conversation strategy.

This section outlines what will shape their success and what they must do now.

STRATEGIC IMPERATIVES FOR AI VOICE SUCCESS

Frame the Market Narrative / Value Proposition

1. Co-create the narrative: be an Intelligent Engagement player. Talk less about technology and more about the outcomes (*Intelligent Engagement*). The story requires good structure (e.g., Automate, Assist, and Analyze) and strong case studies to bring value to life.

2. Evidence business results. Move the conversation from technology features to business results. Focus on benefits of unlocking scale, reducing handling time, improving conversion and scale economics. Product strategy should address budget scrutiny and ROI directly. Shift pricing towards outcomes based when feasible.

3. Arm your partners (and educate your/their customers). AI Voice is a new and fast developing field, and sales teams that aren't incentivised and aren't ready to deliver concise, compelling explanations will certainly not be able to sell it. Equally, customers need to know "what can I do with AI?" If you don't help them, someone else will.

4. Engage early with emerging regulation. AI Voice may sit at the boundary between AI-generated "content" and electronic communications services. The way regulators classify these interactions could materially affect how services are designed, disclosed and governed. Intelligent Engagement players should actively engage in emerging regulatory processes – including the EU AI Act transparency Code of Practice – to ensure that real-time communications use cases are properly understood and that workable standards for disclosure, marking and governance are developed.

Secure Platform Control & Champion Governance (Technology & Operations)

5. Control orchestration (runtime). The orchestration layer governs routing, escalation, policy and model execution. Control here determines who captures intelligence margin and who remains a transport provider. Focus your messaging and product design to show customers the strengths that partnering with you as an orchestrator delivers (i.e. points 6-11 below). Communications platforms are embedding AI agents directly into their core products. This strengthens the importance of owning the runtime layer that orchestrates live conversations.

6. Champion governance and trust. Mission-critical conversations require reliability, auditability and privacy by design. Embed compliance, data sovereignty and policy control into the core platform and your values. Governance can be a source of differentiation, not overhead. vCons (a new data standard) are a potential asset for Intelligent Engagement players in this respect.

7. Remain AI-Model neutral. Enterprises need flexibility in a fast-moving model market. Platforms that allow models to be swapped without architectural disruption become the stabilising layer in enterprise AI Voice deployments.

8. Address Sovereignty requirements. In some markets, broadly most outside of the US and China, national-level data sovereignty is an increasingly pressing concern. If these are the markets you serve, ensure that your proposition clearly addresses these needs. Even within the US, sovereignty is an issue at a company level. Customers want to be in control of their options, and be able to implement sovereignty if and when required.

9. Leverage network advantage where relevant. For CSPs, integrating AI Voice with Network APIs and network-native capabilities can create scalable, high-quality services that are difficult for pure software providers to replicate.

Manage the Human Layer (The Integration & Adoption)

10. Deepen Enterprise Integration. Embed AI Voice into CRM, workforce management, identity and compliance systems. Integration increases switching costs and strengthens long-term relevance.

11. Manage Human-AI Adoption. AI Voice succeeds faster in most businesses when it helps people rather than threatens them. Prioritise copilots, automation of repetitive work and tools that help teams perform better. Adoption friction is a commercial risk.

THE BOTTOM LINE

Where Intelligent Engagement players can maximise their value-add by proving the benefits of their focus and know-how, they can become the operating system for live enterprise conversations. Where they can't, they will be part of the integration layers beneath others' control. The next phase of AI Voice is about governing execution. Intelligent Engagement players must decide whether they will lead that layer or operate beneath it.

In AI Voice, the strategic question is no longer who carries the call, but who runs the conversation.

A SINCERE NOTE OF THANKS

We would like to thank all of our members for their support and contribution, and would particularly like to thank the following people for their insightful contributions and quotes:

- Adnan Saleem, CTO, Software and Cloud Solutions, Radisys
- Anastasia Perperidou, Corporate Communications Manager, GMS
- Ben Leaman, Growth Marketing, Speechmatics
- Bertrand Pourcelot, Managing Director - UCaaS for CSP's & MSP's, EnReach
- Chris Angus, VP CPaaS & CX Expansion, 8x8
- Daniel Morris, Chief Product Officer, Sinch
- Dean Bubley, Founder, Disruptive Analysis
- Graham Duffy, Marketing Manager, BT International
- Ian Reither, COO & Growth, Telnyx
- Francesca Rigamonti, Head of Marketing, EnReach
- Jacques Cliques, Director, Market Strategy & Product Marketing, Sinch
- Janne Timonen, CEO, LastBot
- Jo Ann Tizzano, VP, Corporate Communications, Vonage
- Jonathan Abrahamson, Chief Product & Digital Officer, Deutsche Telekom
- John George, Founder, AI Voice developer
- Hakim Annaciri, VP Products & Commercial Management B2B LCE, KPN
- Hans Söhngen, Managing Director, KPN Ventures
- Kawthar Awada, Head of Operations and Strategy, Cerqle
- Kevin Nethercott, CEO, Tresic
- Lawrence Byrd, Communications APIs, Vonage
- Lauren King, CMO, Speechmatics
- Melvin Maahen, Senior Product Manager, 8x8
- Michael Brandenburg, Sr. Product Marketing Manager, RingCentral
- Mike Mills, MD, Service Provider, Gamma
- Nathan Reid, Product Marketing Manager, 8x8
- Neelam Sandhu, CMO, Vonage
- Ralph Page, Snr Sales Director, EMEA, Radisys
- Remco Magielse, Product Leader, GLBNXT
- Rob Kurver, Founder & Partner, CPaaSAA
- Robert Arcon, CCO, IoTcomms.io
- Robert Galop, CMO, Tresic
- Robert Gerstmann, Co-Founder, Sinch
- Sophie Cheng, SVP, Product Marketing, Sinch

- Thomas McCarthy-Howe, CTO, vConic
- Thierry van de Velde, Global Core Network Pre-Sales Leader, Nokia
- Victor Shisterov, Director of Product Management, Vonage