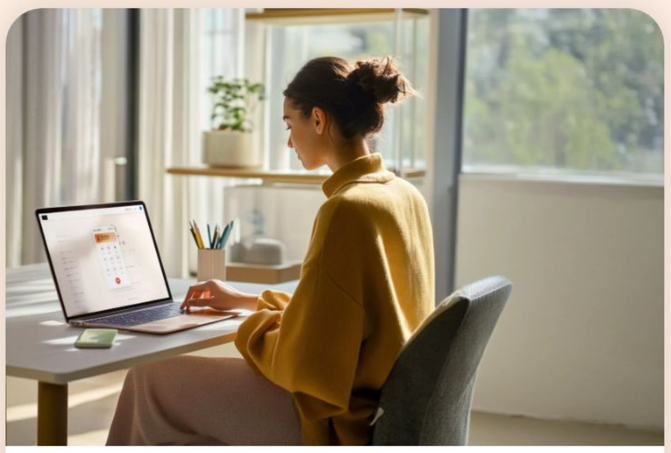
RingCentral

The Agentic Voice AI Model

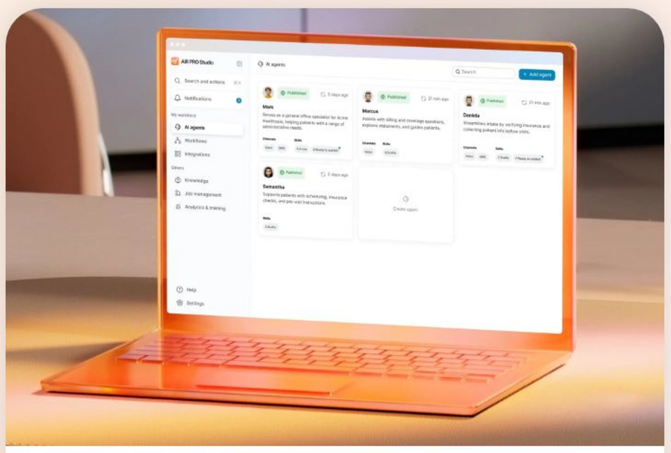
From conversations to action

Table of contents



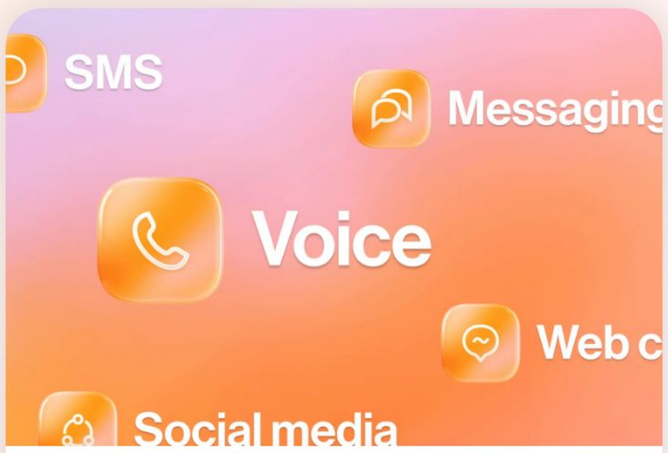
Conversations are becoming operational interfaces

Introduction



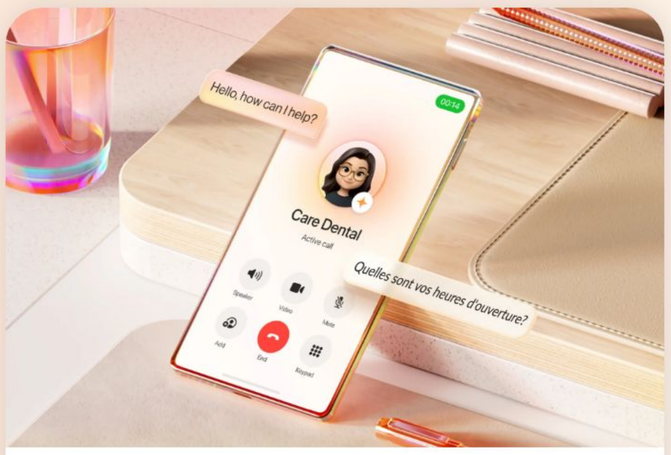
The gap between AI engagement and AI execution

Chapter 1



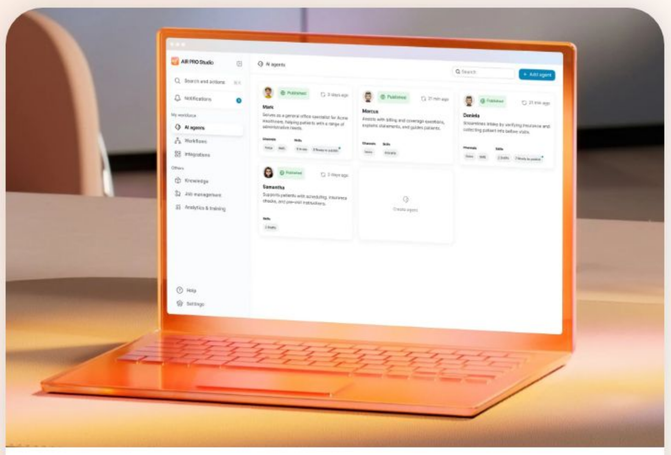
What is the Agentic Voice AI Model?

Chapter 2



Why voice AI agents are different

Chapter 3



Mapping the Agentic Voice AI framework

Chapter 4

What does the Agentic Voice AI Model look like in action?

Chapter 5

Industry applications: Where Agentic Voice AI delivers immediate value

Chapter 6

The future of Agentic Voice AI

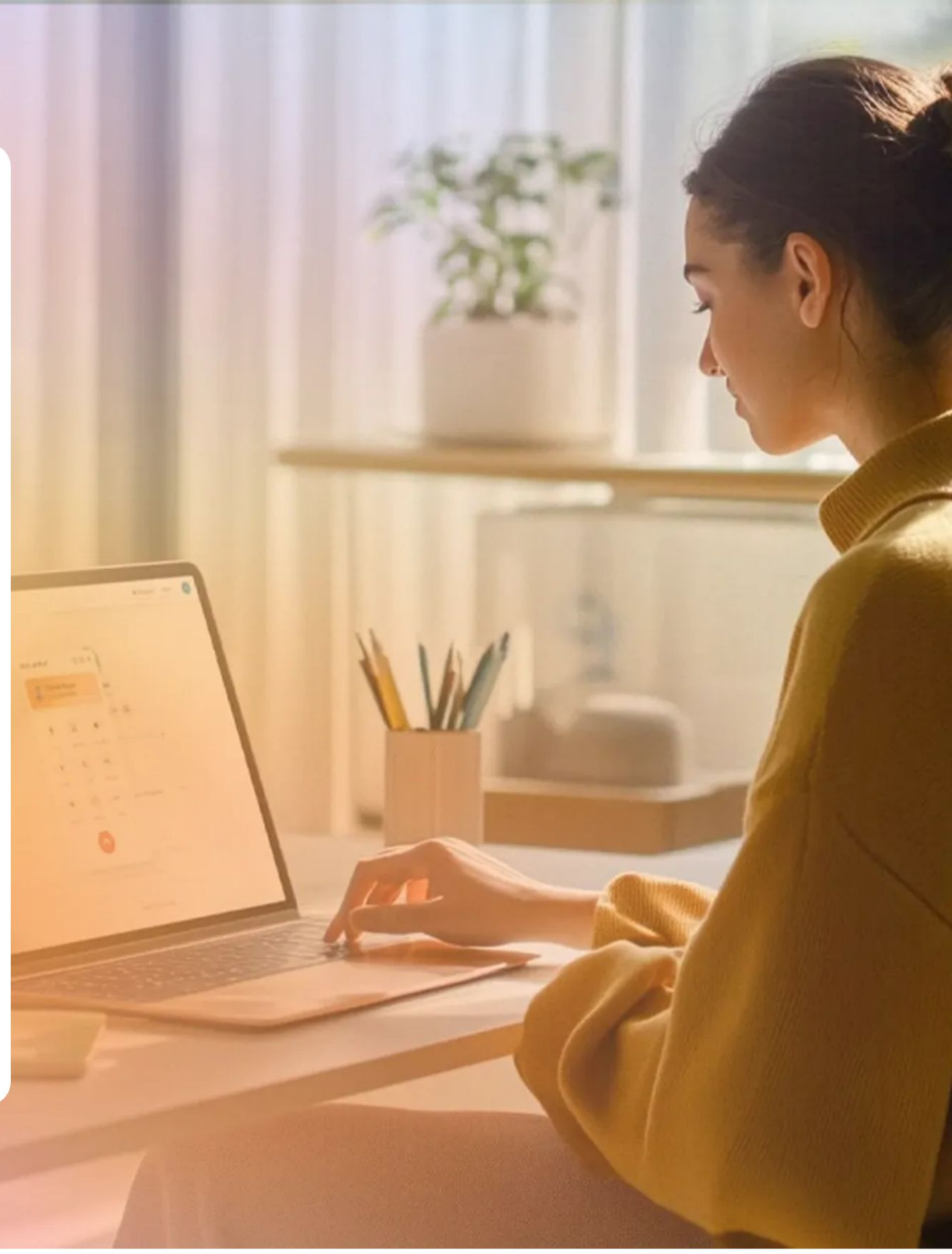
Conclusion

Conversations are becoming operational interfaces

Over the past decade, the way people interact with organizations has fundamentally changed. Increasingly, those interactions take place through conversations. What once required navigating complex portals or submitting tickets is now expected to happen in a simple, natural exchange.

Customers schedule appointments, resolve service issues, and manage accounts through voice calls, chat messages, and digital conversations. Employees rely on conversational interfaces to access internal knowledge, request IT support, submit HR inquiries, and coordinate operational tasks.

As digital communication becomes the primary interface between organizations and the people they serve, expectations for responsiveness have risen dramatically. Customers expect immediate answers regardless of channel, whether they reach out through voice, messaging, SMS, or chat. Employees expect the same immediacy from internal systems and support functions.



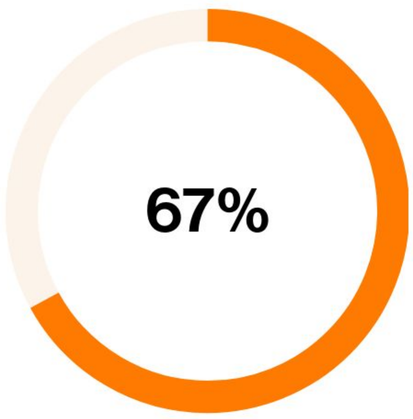
To keep pace with this shift, organizations have rapidly adopted conversational AI technologies across both customer experience (CX) and employee experience (EX) environments.

Chatbots, virtual assistants, and automated response systems now handle a growing share of routine interactions. Yet despite these advances, most conversational AI systems remain limited in scope. They are designed primarily to provide information or route requests, rather than resolve them.

A chatbot may answer a billing question, but it cannot complete the billing adjustment. A virtual assistant may help identify the right department, but it cannot finish the task itself. Even when automation begins the interaction, human teams often remain responsible for completing the underlying work.

This gap is becoming increasingly visible as organizations scale digital engagement. Answers are not good enough for customers and employees. They're seeking movement and outcomes. In fact, recent insights from RingCentral show that 67% of organizations report faster resolution times when AI is applied to customer interactions, highlighting how quickly expectations are shifting toward outcomes rather than responses.

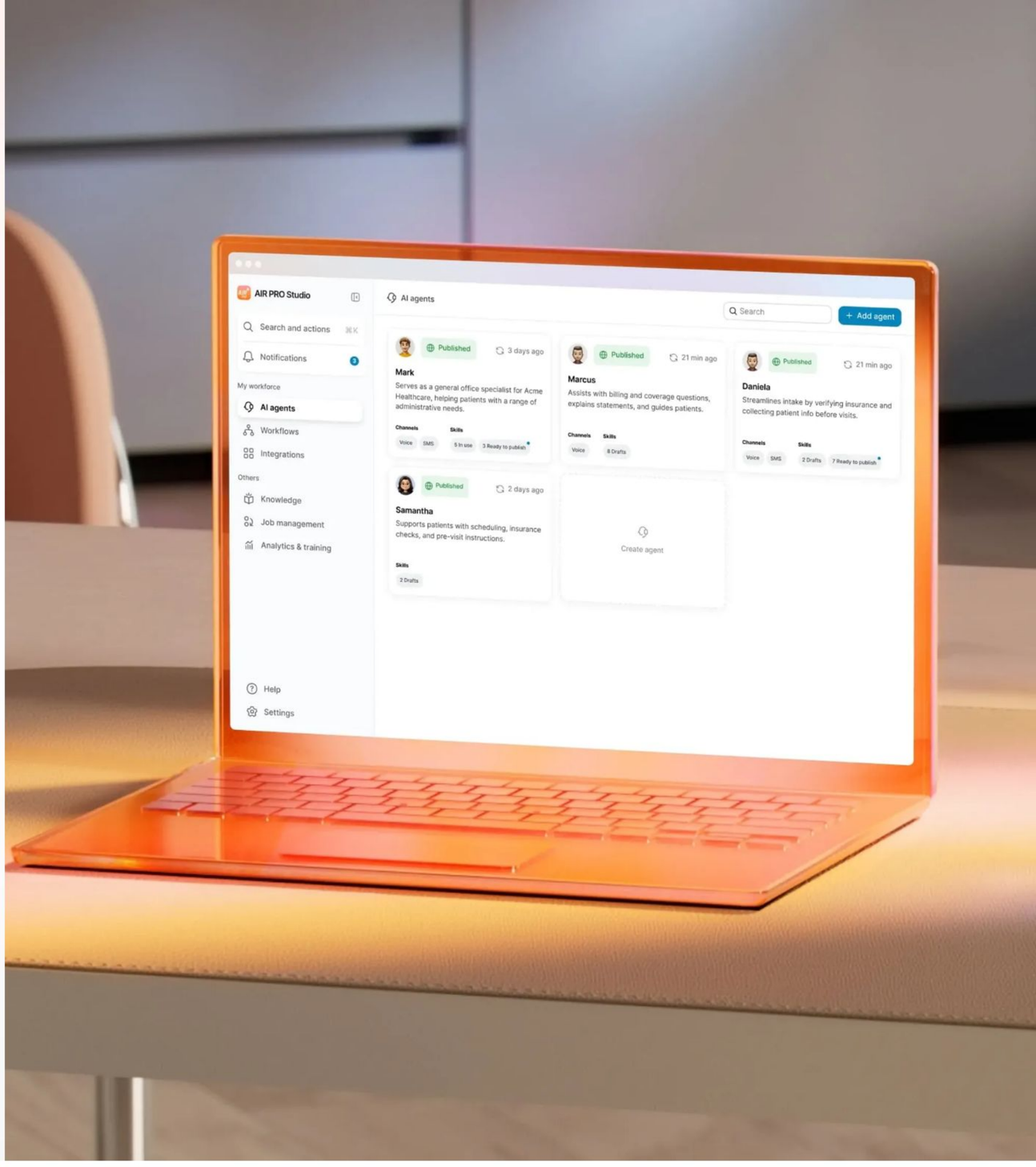
The next phase of conversational AI will therefore shift from engagement systems that respond to requests to execution systems that resolve them.



of organizations report faster resolution times when AI is applied to customer interactions.

Chapter 1

The gap between AI engagement and AI execution



The first generation of conversational AI delivered meaningful progress in automating routine interactions.

Chatbots and voice assistants helped organizations handle high volumes of questions while reducing pressure on service teams. In many environments, these systems now manage tasks such as answering FAQs, guiding through structured workflows, or routing inquiries to the correct department. For businesses facing growing interaction volumes, this level of automation provided an important first step. However, real-world interactions rarely stop at simple information exchange.

Customers may call to update an account, verify their identity, schedule an appointment, or resolve a service issue. Employees may need to access internal tools, retrieve policy information, or initiate operational processes that span multiple systems. These requests often involve verification, decision-making, and coordination across enterprise platforms.

Most conversational AI systems today are not designed to complete these workflows end-to-end.

Instead, they typically rely on a intake-only model where automation begins the interaction, but humans remain responsible for completing the underlying task. The system answers the question, gathers preliminary information, or routes the conversation to a human agent who must then finish the work.

Several structural limitations contribute to this gap

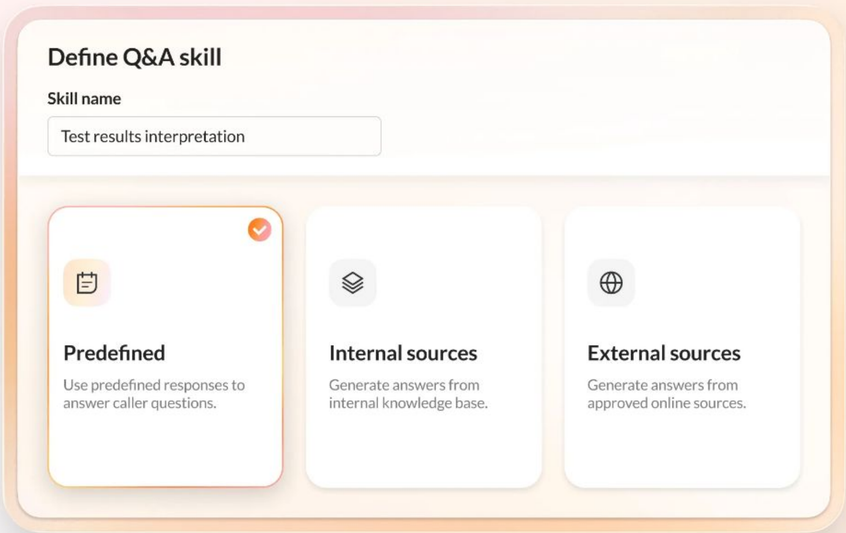
Many conversational AI systems were originally built for text-based interactions, with voice capabilities added later. Integration with enterprise systems is often limited, making it difficult for AI agents to access operational data or execute transactions. Workflow automation frequently requires developer involvement, slowing the pace at which organizations can expand automation across new use cases.

At the same time, interactions themselves have become increasingly fragmented. A customer might begin a conversation in chat, move to voice, and follow up through messaging. Without continuity across channels, context is lost, and workflows become more difficult to complete.

The outcome is a model where conversational AI can often start an interaction, but human teams must still complete the work.

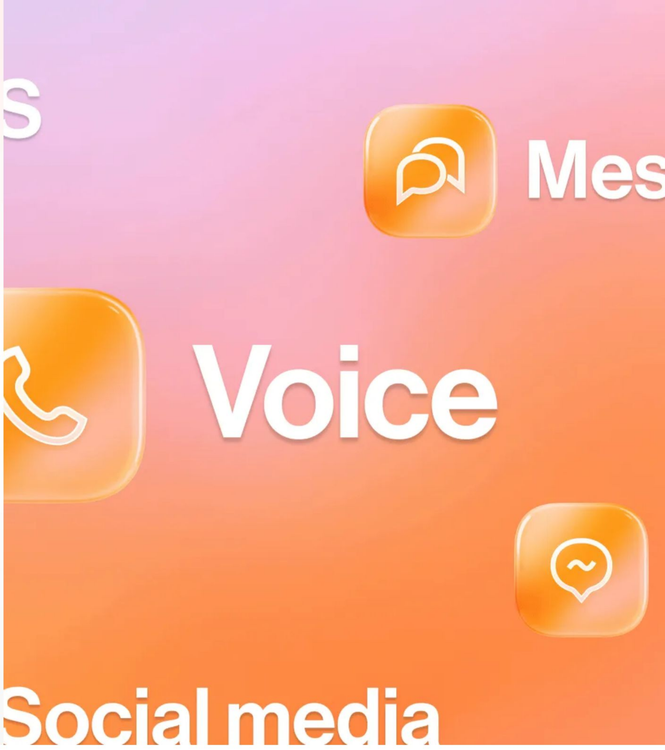
For organizations attempting to scale customer service, internal support, and operational workflows, this creates ongoing friction. Automation helps manage volume, but it does not yet deliver the full operational efficiency organizations are seeking.

Closing the engagement-to-execution gap requires a new approach to conversational automation that can execute tasks directly within the interaction itself.



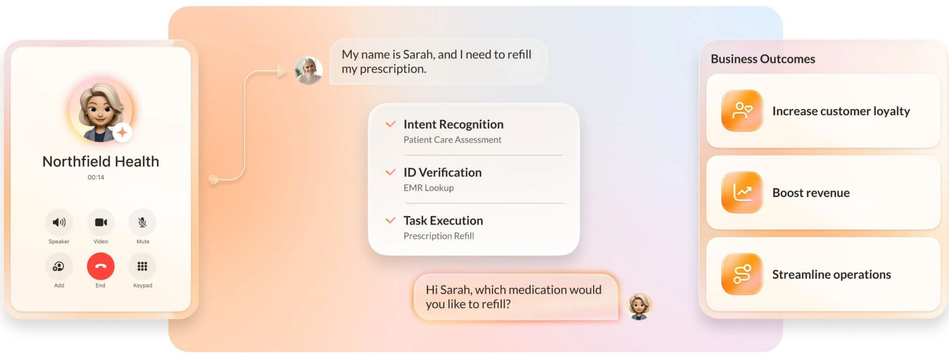
Chapter 2

What is the Agentic Voice AI Model?



The concept of AI agents has existed for years in research and early enterprise experimentation.

Today, advances in AI capabilities, reasoning models, and enterprise infrastructure are enabling these systems to operate at a new scale in real-world business environments. Agentic AI systems differ from earlier conversational technologies in a fundamental way. Rather than simply responding to prompts or generating answers, these systems combine reasoning with governed workflow execution across enterprise systems.



In practice, this means an AI system can interpret a request, determine the steps required to fulfill it, and interact with other systems to complete the task.

Instead of serving primarily as informational assistants, these systems can execute workflows and operational processes directly within the interaction.

When applied to communications platforms, this approach gives rise to what we call at RingCentral the Agentic Voice AI Model.

The Agentic Voice AI Model describes a new architectural approach to conversational systems, where AI agents operate directly within live interactions to understand intent, reason through requests, and execute workflows across enterprise systems.

Within this model, AI agents function as active participants in conversations (particularly voice interactions) capable of resolving requests and completing tasks in real time. They can understand spoken intent within natural conversation, verify identity or contextual information, and determine the appropriate sequence of actions required to fulfill a request.

From there, these agents can interact with enterprise systems such as scheduling platforms, CRM systems, service tools, or knowledge bases to carry out the necessary steps. The interaction moves beyond information exchange and becomes a mechanism for completing operational work.

Voice plays a particularly important role in this evolution because it is the most complex interaction modality. It requires real-time understanding, interruption handling, and continuous context tracking under latency constraints.

Building AI systems capable of reasoning, verifying information, and executing tasks during live voice interactions unlocks a new level of operational automation. Conversations become more than a starting point for workflows; they become the environment in which those workflows are completed.

The Agentic Voice AI Model, therefore, represents a shift from conversational systems that assist users with information to systems capable of resolving requests and executing work within the interaction itself.

Chapter 3

Why voice AI agents are different

Voice AI agents operate under different technical and operational requirements than text-based systems.

First, they rely on real-time streaming architectures that allow the system to process and respond to inputs continuously as the conversation unfolds.

Second, they depend on low-latency decision loops that continuously update intent, context, and next-best actions during the interaction.

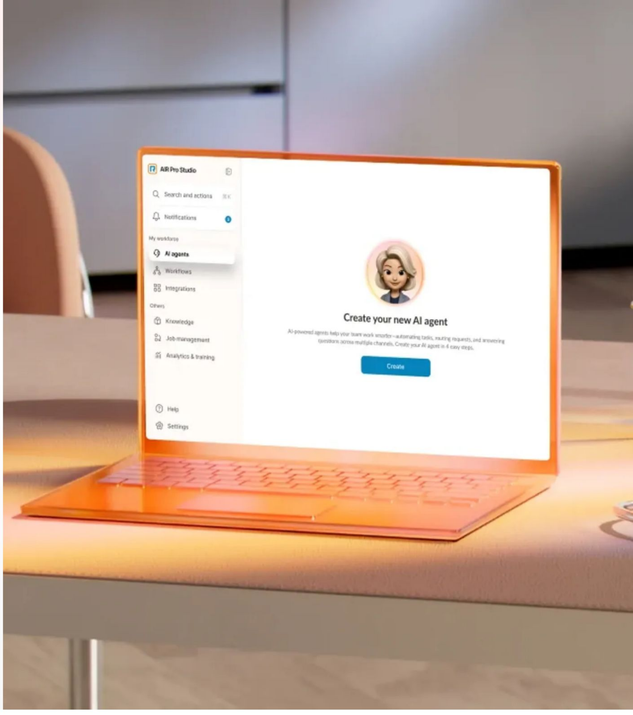
Third, they must support voice-native turn-taking, handling interruptions, pauses, and conversational shifts without breaking the flow.

Finally, these systems rely on event-driven execution rather than simple prompt chaining, allowing them to respond to changes in identity state, workflow progress, and system events in real time.

These characteristics make voice the most demanding modality for AI and one of the most valuable, particularly for complex, high-stakes interactions.

Chapter 4

Mapping the Agentic Voice AI framework



For organizations to deploy agentic voice AI successfully, conversational platforms must support a new set of capabilities.

At its core, the model requires systems designed for voice-native intelligence.

Voice interactions are inherently dynamic and often unpredictable. Unlike structured digital inputs, spoken conversations involve nuance, interruptions, clarifications, and evolving context. AI agents operating in voice environments must be able to interpret intent within natural conversation, manage turn-taking, and maintain contextual awareness throughout the interaction. Systems originally designed for text-based interactions often struggle to meet these requirements, making voice-first design essential for operational use cases.

The second critical capability is omnichannel continuity.

Modern interactions rarely occur within a single communication channel. A customer may begin a conversation through chat, continue through a voice call, and follow up via messaging. Similarly, employees may move between collaboration tools, phone conversations, and support systems while resolving an issue. For AI agents to operate effectively across these environments, they must preserve context across channels so that conversations (and the workflows associated with them) can continue seamlessly without restarting the interaction.

Equally important is action-oriented execution.

To move beyond informational assistance, AI agents must be able to interact directly with enterprise systems and operational workflows. This includes accessing data from CRM systems, updating records, scheduling services, triggering automated processes, or retrieving information from internal knowledge sources. Without the ability to execute these actions, conversational AI remains limited to providing guidance rather than delivering outcomes.

Governance can no longer be an afterthought.

It is a foundational requirement as AI agents move from answering questions to executing work. These systems must operate within clearly defined boundaries that ensure actions remain aligned to business rules, data policies, and compliance requirements. This includes purposeful swimlanes that define what each agent can access, what actions it can take, and which systems it can interact with. Role-based access controls, permissions at the knowledge and action level, and workflow policies all work together to ensure that agents operate within their intended scope. Without this level of control, the risk of incorrect actions, data exposure, or inconsistent behavior increases as automation scales. With it, organizations can confidently expand AI-driven workflows while maintaining trust, security, and operational integrity.

Organizations must also maintain business-user control over these systems.

In parallel with governance, organizations must maintain business-user control over how these systems are configured and evolve over time. Business users should be able to define goals, configure workflows, and adapt automation to evolving operational needs. This flexibility allows organizations to scale automation quickly while maintaining alignment with internal processes and policies.

Finally, the Agentic Voice AI Model depends on continuous optimization.

Organizations need visibility into how AI agents perform across interactions, including how requests are resolved, where workflows succeed, and where escalation occurs. Operational analytics allow teams to monitor outcomes, identify opportunities for improvement, and expand automation into new use cases over time. When AI agents handle more interactions, this feedback loop becomes essential for improving accuracy, efficiency, and overall experience quality.

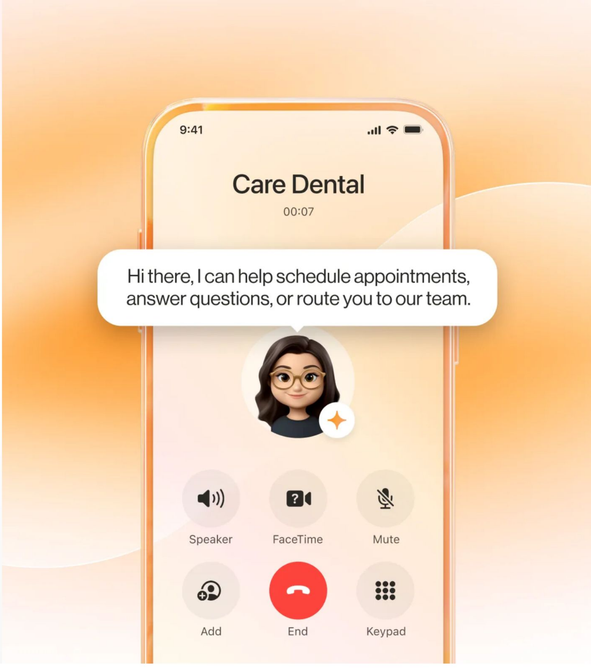
Together, these capabilities enable conversational systems to move beyond basic engagement and toward operational execution.

Rather than acting solely as interfaces that guide users to the next step, AI agents operating within the Agentic Voice AI Model become active participants in completing work during the interaction itself.

This framework lays the groundwork for the next generation of conversational platforms: systems designed to go beyond simply managing conversations, resolving requests, executing workflows, and ultimately delivering measurable outcomes.

Chapter 5

What does the Agentic Voice AI Model look like in action?



RingCentral AIR Pro applies the Agentic Voice AI Model in practice. It is a voice-first, omnichannel AI agent system built to support real-time reasoning and governed workflow execution during live interactions.

At the core of AIR Pro is a real-time voice reasoning loop. As a conversation unfolds, the system continuously interprets intent, tracks context, handles interruptions, and determines the next appropriate action. This allows the interaction to evolve dynamically rather than follow a fixed prompt-response pattern.

This reasoning layer is paired with an embedded workflow execution engine, enabling AIR Pro to trigger and coordinate actions across enterprise systems during the interaction itself. Requests can move from intake to execution without requiring a separate handoff.

Because AIR Pro is built natively on RingCentral's communications platform, it benefits from direct integration with UCaaS and CCaaS environments. This allows AI agents to operate within the same systems where interactions already take place, rather than relying on disconnected layers or external orchestration.

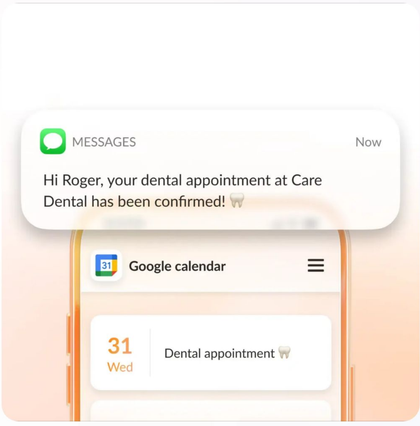
AIR Pro also supports cross-channel state persistence, allowing conversations to move across voice, messaging, and chat without losing context, so workflows can continue without restarting.

Rather than functioning solely as conversational interfaces, AI agents operate directly within governed operational processes.

During an interaction, an agent can understand spoken intent, verify identity or contextual information, and determine the actions required to resolve a request. From there, it can interact with enterprise systems (such as scheduling tools, CRM platforms, service management systems, or knowledge bases) to carry out the necessary steps.organizations can expand automation across new use cases.

As businesses matured with AIR, they saw what a simple digital agent could do and just how much more potential there was to unlock. AIR Pro extends that foundation into a model designed for real-time execution and more complex workflows.

A key component supporting this capability is AIR Pro Studio, a no-code environment that enables business teams to design and deploy AI agents without extensive developer involvement. Within this environment, users can define the outcome an agent should achieve, like resolving a billing inquiry, scheduling an appointment, or verifying account information, and the system interprets that goal to coordinate the appropriate workflow.



Because AIR Pro AI agents operate within a voice-first communications environment, they can engage users across both voice and digital channels while maintaining continuity of the interaction

Conversations can move across phone calls, messaging, or chat while preserving context, allowing requests to be resolved without forcing users to restart the process.

Just as important is the ability to monitor and continuously improve how these interactions perform. Operational analytics provide visibility into resolution rates, workflow completion, and escalation patterns, enabling organizations to refine automation and expand agent capabilities over time.

Through this approach, AI agents evolve into active participants in completing the work required to resolve questions and route requests.

Chapter 6

Industry applications: Where Agentic Voice AI delivers immediate value

The operational potential of the Agentic Voice AI Model becomes most visible when applied to high-volume interactions that are repetitive in pattern but still require verification, system access, and workflow execution.

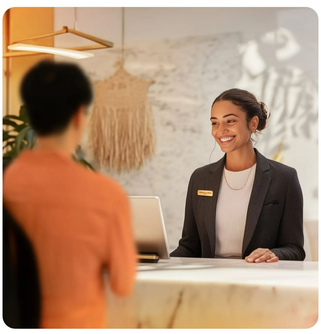
These are the types of interactions that have traditionally followed the intake-only model, in which automation captures the request, but human teams must complete the work.

Agentic voice AI changes this dynamic by enabling more of these interactions to move from intake to resolution within the conversation itself.

Customer service and support

Customer service teams handle large volumes of interactions related to account management, service inquiries, troubleshooting, and appointment coordination. Many of these interactions follow predictable patterns but still require operational steps such as verifying identity, retrieving account information, updating records, or scheduling services.

Agentic voice AI agents can perform these tasks during the conversation itself. A customer calling about a billing question, for example, can be authenticated, guided through the issue, and have the relevant account update completed within the same interaction. Instead of routing the customer to multiple representatives or systems, the AI agent can coordinate the process directly across the necessary platforms. In addition to handling inbound interactions, these systems can also support proactive engagement, such as sending reminders, follow-ups, or service notifications triggered by system events or schedules.



For service teams, this reduces operational load while allowing human agents to focus on complex or sensitive interactions that require judgment and empathy.

Employee support and internal operations

The same model also applies to internal employee support environments.

Traditional support models rely on ticketing systems or service desks, where requests are routed to specialists who must manually retrieve information or execute tasks across multiple systems.

Agentic voice AI agents can handle many of these requests directly during the interaction. Employees can access policy information, reset credentials, submit requests, or retrieve operational updates through conversational interactions that execute the necessary actions behind the scenes.

This reduces ticket backlogs while allowing support teams to focus on more complex operational issues.

Across these environments, the value of the Agentic Voice AI Model lies in its ability to transform conversations from entry points into mechanisms for completing work.

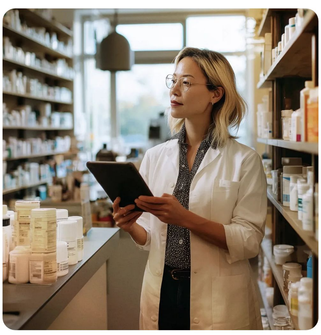
Interactions that once required multiple handoffs can increasingly be resolved within the conversation itself.

Healthcare

Healthcare organizations are among the environments where conversational interactions and operational workflows intersect most frequently. Providers manage high volumes of patient communications related to appointment scheduling, insurance verification, pre-visit instructions, and post-visit follow-up.

These interactions are often administrative in nature but require access to multiple systems, including electronic health records, scheduling platforms, and billing systems.

Agentic voice AI agents can help streamline these workflows by managing appointment scheduling, verifying patient information, answering common inquiries, and coordinating follow-up actions within a single interaction.



This reduces administrative burden for healthcare staff while improving responsiveness for patients seeking information or assistance.

Conclusion

The future of Agentic Voice AI

AI agents will handle high-volume, repeatable workflows across customer service, healthcare coordination, internal operations, and other interaction-heavy environments.

Human teams will remain central to these processes, focusing on complex cases, strategic decision-making, and interactions that require empathy, negotiation, or creativity. AI agents, meanwhile, will handle routine operational tasks that can be executed consistently and efficiently through conversational workflows.

These systems will increasingly operate within clearly defined data policies, business rules, and governance frameworks to ensure reliability, compliance, and trust at scale.

Voice will remain a particularly important channel within this evolution.

Many of the most urgent, sensitive, and operationally complex interactions still occur over calls. Systems capable of understanding intent, verifying context, and executing tasks in real time during live voice interactions unlock a level of operational automation that traditional digital interfaces cannot easily replicate. As the Agentic Voice AI Model continues to mature, organizations will increasingly design communications platforms not simply to manage conversations but to resolve them.

Conversations will become the interface through which operational work is completed, enabling faster resolution, improved efficiency, and more responsive experiences across both customer and employee environments.

RingCentral Agentic AI Trends 2026

Learn more about RingCentral AI

About RingCentral

RingCentral is a leading provider of trusted AI communications, contact center, sales intelligence, video, and hybrid event solutions. RingCentral empowers businesses with conversation intelligence and unlocks rich customer and employee interactions to provide insights and improved business outcomes. With decades of expertise in reliable and secure cloud communications, RingCentral has earned the trust of hundreds of thousands of customers and millions of users worldwide. For more information, please contact a sales representative.

Visit ringcentral.com or call 855-774-2510.

RingCentral, Inc. 20 Davis Drive, Belmont, CA 94002.

© 2026 RingCentral, Inc. All rights reserved.

RingCentral, the RingCentral logo, and all trademarks identified by the ® or ™ symbol are registered trademarks of RingCentral, Inc. Other third-party marks and logos displayed in this document are the trademarks of their respective owners.