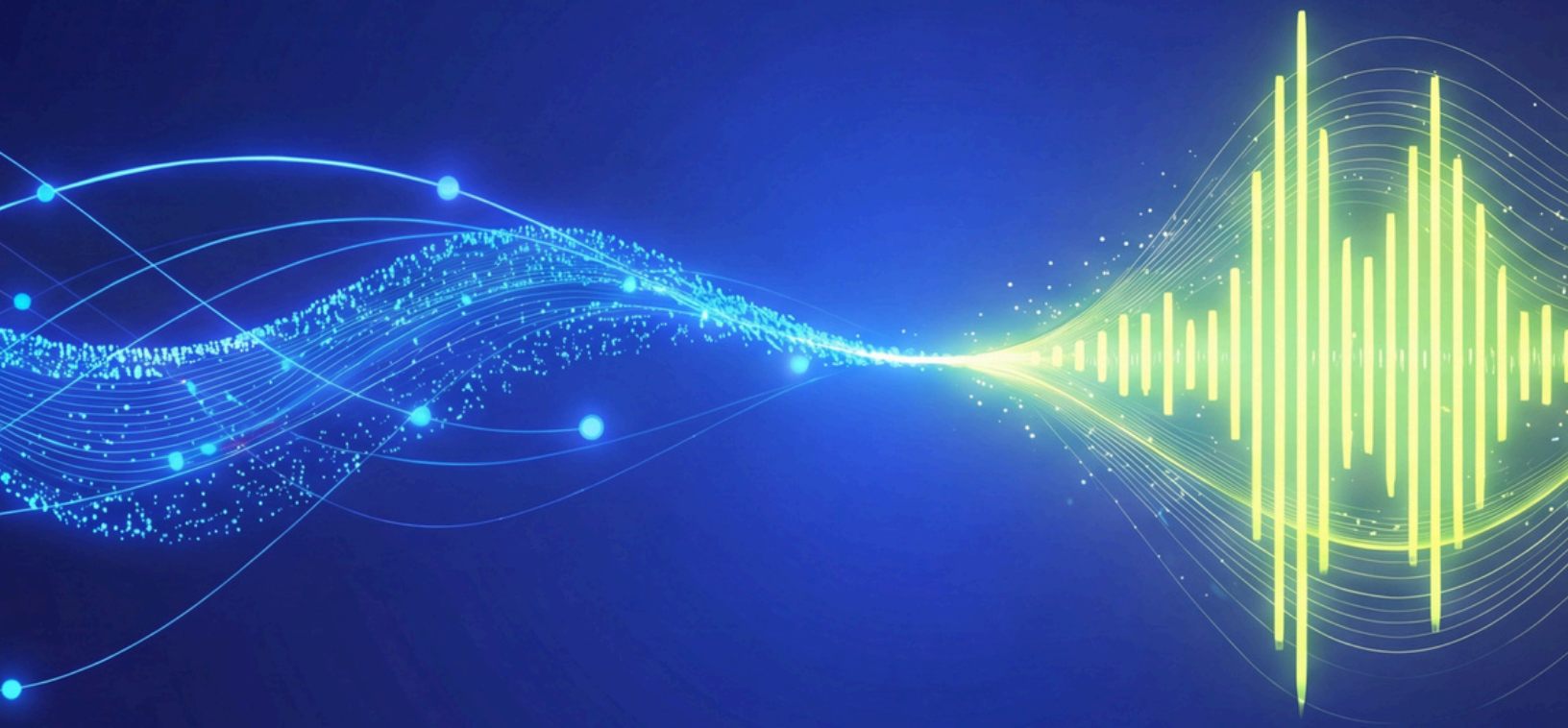
Buyer's Guide

The Enterprise Guide to Voice AI Agents

A buyer's guide for CX, IT and finance leaders evaluating the next generation of Agentic customer self-service experience.



Five9

Contents

Foreword

01 Voice is your most critical channel. And it's not scaling.

02 A new generation of agentic voice AI is the superpower to leapfrog your CX performance.

03 What a great voice experience looks like in 2026

04 What the next generation of voice AI agents actually looks like.

05 What enterprise-grade voice AI agents require.

Foreword

Customers don't want another bot. They want resolution.

That gap — between what self-service promises and what it actually delivers — is the most expensive line item in your contact center today. Customers get stuck in rigid flows, repeat themselves, wait for an agent anyway, or abandon the journey altogether. Volume climbs. Cost climbs with it. Loyalty leaks out one frustrating call at a time.

A new generation of voice AI is closing that gap. Not another bot. Not another script. AI that listens, reasons, takes action, and hands off to a human the moment it matters — at enterprise scale.

This guide is for leaders weighing what to do next. You'll find a clear-eyed view of what's broken with legacy voice automation, what's now possible with the next generation, what to demand from any vendor on your shortlist, and an evaluation framework to use across the buying committee.

Read it in fifteen minutes. Make a better-defended decision afterward.

1 Voice is your most critical channel. And it's not scaling.

Voice is the costliest channel to serve. Failed automation just adds repeat calls and cost. Call volumes keep rising. Budgets stay flat. And the tools most teams rely on weren't built for what customers expect in 2026.

What's actually broken

- **Legacy bots have hit the containment ceiling.** Scripted flows handle the easy 30%. Everything else escalates.
- **Customers still end up shouting "AGENT."** The bot couldn't reason, so they did the only thing left.
- **Repeat contacts keep driving costs higher.** Containment looks good on a dashboard. The same customer calling back twice tells a different story.
- **Rigid flows demand long build cycles.** New use cases sit in the backlog for quarters.
- **Most AI never leaves pilot mode.** The demo was great. Production never happened.





The pressure is compounding. Voice is your most-used channel, your most expensive channel, and the channel where every customer decides — in the first thirty seconds — whether your brand is worth their loyalty.

The mandate

The convergence of large language models, real-time speech, and agentic AI has created a once-in-a-generation opening. Self-service stops being a containment play. It starts driving loyalty and growth.

The question is no longer whether to adopt agentic CX. It's whether you move fast enough to lead.

WHY NOW

Gartner estimates that AI Agents will resolve 80% of common customer service interactions autonomously by 2029. The contact centers betting on agentic architecture today are positioning to win the next decade of CX. The ones still iterating on scripted IVR will spend that decade catching up.

2 A new generation of agentic voice AI is the superpower to leapfrog your CX performance.

For a decade, "voice AI" relied on intent recognition and predefined conversational flows. The customer said something. The system matched the expected intent. If the match was good, the conversation followed a scripted path. Effective for common requests, but anything more complex issues required a transfer.

Agentic voice AI is a different category entirely.

Instead of responding, it resolves. Instead of matching intent to a fixed script, it understands context and plans the next action. Instead of stopping at "I don't understand," it asks a clarifying question, switches context and finishes the task — across the systems where the answer actually lives.

The shift is from answering questions to delivering outcomes.



The shift, in plain terms

Legacy Bots	Agentic Voice AI
Follow a script	Understand and adapt
Struggle with unexpected requests	Handle conversations naturally
Sound robotic	Sound human
Get confused by interruptions	Handle interruptions seamlessly
Lose context	Remember context
Escalate only when stuck	Escalate intelligently
Handle one request at a time	Resolve complex, multi-step requests

What's now possible

- **AI that understands intent and context in real-time** — even as the conversation shifts.
- **Multi-turn, non-linear conversations** with full continuity across the interaction.
- **Real action across business systems** — without forcing the customer to repeat themselves to a human.
- **Intelligent escalation** only when human judgment is required.

Sounding human isn't enough. AI has to understand the moment.

The contact centers winning on CX in 2026 aren't the ones running the most pilots. They're the ones investing in agentic architecture — the platform layer that turns isolated AI experiments into production systems that scale.

3 What a great voice experience looks like in 2026.

This is the customer-facing half of the story. What should change for the person on the other end of the line.

Conversations that feel human

The bar has moved. Customers expect natural, low-latency interactions that understand intent the first time — no rigid scripts, no robotic prompts. The AI should adapt instantly when the caller changes direction, and respond with conversational flow.

Low-latency streaming. Interruption handling. Real-time context switching.
The kind of interaction your highest-value customers expect — and the kind your competition is still building toward.

Customers don't want scripts.
They want help that understands the moment.

Intelligence that reasons, acts, and resolves

The next generation goes far beyond FAQs. It reasons through problems, handles interruptions gracefully and shifts context on the fly to complete complex tasks across business systems.

The result: fewer transfers, higher first-contact resolution and measurable improvement in CSAT.

Seamless human and AI collaboration

When a human touch is required, the transition should be effortless. Full context should transfer across the handoff. Customers should never have to repeat themselves.

AI should augment human expertise, not replace it. The conversations that need a human get one — better prepared, with more context, in less time. And every conversation should make the system sharper.

The best AI gets better with every conversation.

The outcomes that follow

- Higher containment rates — without the customer-experience tax
- Improved CSAT and NPS
- Reduced repeat contacts
- Faster resolution times
- Lower cost-to-serve

CUSTOMER PROOF POINT

"Five9 solved critical challenges around noise handling, turn detection, and hallucination prevention, enabling reliable, context-aware conversations at scale. Just as importantly, they were a true partner in helping us move from concept to production. By year-end, we expect Five9 Voice AI Agents to handle more than 100,000 service calls this year while improving the customer experience."

-Ruthu Raj, VP of Architecture, IT Infrastructure and Operations at PODS.

Want to see this play out across specific use cases — from password resets to insurance claims intake? Read the **The Agentic Continuum: Voice AI Use Cases Across Every Stage of Resolution**.

4 What the next generation of voice AI agents actually looks like.

Customer experiences like the ones above don't come from incremental upgrades to legacy bots. They come from a different category of architecture. Three things define it.

1. A new execution layer

Not generative AI bolted onto yesterday's IVR. A platform built from the ground up for multi-agent orchestration, advanced AI prompting and integrated guardrails. The chassis itself is agentic — designed for non-linear conversations and real action across business systems. Instead of mapping every path a customer might take, you define what the agent needs to accomplish in plain language and let it reason its way to resolution. No rigid flow diagrams. No developer dependency for every update.

2. Built by voice experts

Voice is its own discipline. Latency thresholds, interruption handling, multilingual fluency — these capabilities aren't transferable. They're built by teams that treat voice as the core, not an afterthought. The vendors who treat voice as a side feature lose to the ones who don't.

3. Designed for the buying committee

Enterprise AI decisions aren't made by one person. The CIO needs confidence in security, governance and integration. The CFO needs a clear line from deployment to measurable cost reduction. The CCO needs containment, CSAT and customer experience outcomes that justify the investment. The right platform doesn't ask you to trade one off against another.

Bots respond. Agentic AI acts.

When you see this generation in a demo, the cues are unmistakable: the AI handles the messy parts. Topic shifts don't break it. Interruptions don't reset the call. Multi-step requests are completed end-to-end. New use cases ship in days, not quarters.

This is the architecture Five9 Voice AI Agents are built around. The rest of this guide unpacks what to expect from it — for your customers, your operations team, your CIO and your CFO.



Back to TOC →

5 What enterprise-grade voice AI agents require.

This is the half of the story your CIO and CFO will care about most.
The non-negotiables.

Purpose-built for complexity

The biggest risk in voice AI isn't the demo. It's what happens when a real customer says something you didn't script for. Most automation breaks the moment the request gets messy. A customer changes their mind mid-sentence. A refund requires three backend systems. An intent shift mid-call sends a legacy bot to a dead end.

Enterprise-grade voice AI should handle multi-step intent, real action across business systems, and shifting context — at scale, without losing the thread. The architecture should be designed for complexity, not in spite of it.

Real moments are messy. Your AI should be ready for that.

Speed to ship, from idea to live in days, not quarters.

The fastest way to lose to a competitor on AI is to spend a quarter on a use case they shipped in a week. The next generation should let your team define agents in natural language, reuse what works across journeys, and ship at the cadence the business actually moves.

Trust, built in

AI only scales when leaders trust it. The platform should bring enterprise-grade guardrails to every interaction — governance, real-time monitoring, sensitive data redaction, continuous compliance checks. Every interaction controlled, auditable, and on brand. You don't find out an agent is underperforming from a customer complaint — you find out from the dashboard, and fix it before the next call.

AI that can't be trusted won't scale.

A dial of autonomy you control

Not every contact center is ready for full automation on day one. The platform should let you calibrate autonomy by use case — scripted, assisted, supervised, fully agentic — moving along the dial at your pace, on your terms.

Automate what's ready. Hold the rest. Move forward with confidence.



EVALUATION CHECKLIST:

What to ask every voice AI agent vendor.

BEFORE YOU START: A NOTE FOR BUYERS

Most vendors will answer yes to every question on this list. The difference is in the follow-up.

For each capability, ask to see it live in a demo environment — not a recorded walkthrough.

Ask for a reference customer in your industry who is in production today, not in beta.

Use the scoring guide below to compare vendors objectively.

Scoring guide

Score each vendor 1 to 3 per question. Total scores at the bottom of each category.

- **1 —** Not available or requires a custom build.
- **2 —** Partially available or on the product roadmap.
- **3 —** Available and proven in production today.

Quick-reference scorecard

Track scores across categories. Detailed questions follow.

Category	Vendor A	Vendor B	Vendor C
1. Conversation quality and human-likeness	/3	/3	/3
2. Agentic architecture and orchestration	/3	/3	/3
3. Tool calling and enterprise integration	/3	/3	/3
4. Safety guardrails and governance	/3	/3	/3
5. Testing, simulation, and continuous improvement	/3	/3	/3
6. Human-AI collaboration and handoff	/3	/3	/3
7. Omnichannel continuity	/3	/3	/3
8. Deployment and scalability	/3	/3	/3
9. Commercial model and total cost of ownership	/3	/3	/3
10. LLM and model flexibility	/3	/3	/3
11. Vendor maturity, support, and time to value	/3	/3	/3

1. Conversation quality and human-likeness

- Does the AI handle interruptions, topic shifts, and multi-turn context without breaking?
- Is latency low enough to feel natural, not just functional?
- Does the AI understand messy, real-world language — or only clean, scripted inputs?
- Can the customer talk like a person — or do they still have to talk like a script?

2. Agentic architecture and orchestration

- Is this a purpose-built agentic architecture, or GenAI bolted onto a legacy flow engine?
- Can multiple specialized agents coordinate on a single customer journey?
- How does the vendor handle context across agent handoffs?

3. Tool calling and enterprise integration

- Which CRM, ERP, ticketing, and billing systems are pre-integrated today?
- How many production integrations does the vendor have — and at what scale?
- What happens to your existing investments? Replace, extend, or both?

4. Safety guardrails and governance

- Hallucination tracking, context adherence, sensitive data redaction — built in or bolted on?
- Can you require explicit runtime approval before the AI triggers a sensitive workflow, like a refund or account change?
- What does the audit trail look like for a regulated industry deployment?
- Can your CISO get the controls they need without a custom build?

5. Testing, simulation, and continuous improvement

- Can you simulate edge cases before they reach customers?
- How is AI performance measured in production?
- What's the cadence for tuning and improving deployed agents?

6. Human-AI collaboration and handoff

- When the AI escalates, what context does the human agent receive?
- Are AI Agents and Human Agents on the same platform — or two integrations away from each other?
- How does the system learn from human handoffs and improve over time?

7. Omnichannel continuity

- Does the same agent reasoning work across voice, chat, SMS, and messaging?
- Are intents and conversation summaries reusable across channels?
- How is context maintained as a customer moves between channels and sessions?

8. Deployment and scalability

- Can the platform scale to peak volumes without re-architecting?
- What's the deployment model — fully managed, hybrid, single- or multi-tenant?
- What does a phased rollout look like across regions or business units?

9. Commercial model and total cost of ownership

- Is pricing usage-based, seat-based, or outcome-based?
- What's included in the base SKU versus add-ons?
- What are the hidden costs — implementation, tuning, integrations, support tiers?

10. LLM and model flexibility

- Is the platform LLM-agnostic — or locked to a single model provider?
- Can you swap the underlying model without re-implementing your agents?
- How does the platform handle model upgrades, deprecations, and new releases?

11. Vendor maturity, support, and time to value

- How many enterprise customers are in production today?
- What's the average time from contract to first live agent?
- What does the implementation, enablement, and support model look like?

TIP FOR THE BUYING COMMITTEE

Score each vendor 1 to 3 on every dimension. Anything below a 2 is a red flag worth raising before you sign.

CONCLUSION:

The contact centers that lead are moving now.

The technology has arrived. The business case is clear. The gap between AI that demos well and AI that scales is real — and the contact centers that figure out which side of that gap they're standing on first will set the pace for the next decade.

When you pick a partner, look for three things

- **Architecture over experimentation.** A pilot is a project. A platform is a strategy.
- **Platform over point solutions.** AI Agents and Human Agents belong on the same foundation, not two integrations away from each other.
- **Voice expertise.** Voice is its own discipline. The vendors who treat it as a side feature lose to the ones who treat it as the core.

Five9 Voice AI Agents were built around these principles — by voice experts, for the agentic era, on a single enterprise-grade platform that brings AI Agents and Human Agents together.

Self-service stops being a containment play. It starts driving loyalty and growth.

Ready to go deeper?

- **Read the companion handbook.** The Agentic Continuum: Voice AI Use Cases Across Every Stage of Resolution **walks through what voice AI looks like in production — from simple-intent self-service through autonomous resolution of complex workflows.**

- **See it in action.** Request a guided demo of Five9 Voice AI Agents handling real-world scenarios.





About Five9

Five9 empowers organizations to create hyper-personalized and effortless AI-driven customer experiences that deliver better business outcomes. Powered by Five9 Genius AI and our people, the Five9 Intelligent CX Platform is trusted by 2,500+ customers and 1,400+ partners globally. The New CX Starts Here and it's at the heart of every winning experience.

For more information, visit www.five9.com

Five9 and the Five9 logo are registered trademarks of Five9 and its subsidiaries in the United States and other countries. The names and logos of third party products and companies in this document are the property of their respective owners and may also be trademarks. The product plans, specifications, and descriptions herein are provided for information only and subject to change without notice, and are provided without warranty of any kind, express or implied.
Copyright © 2026 Five9, Inc. 07025