WHAT YOU NEED TO KNOW CEOS CIOS CTOS 13 MIN READ

Stop Buying AI Hype and Better Prepare for True Agentic AI Investment

AUGUST 4TH, 2025

Authors

Jim Davies, Analyst and Executive Partner
George Harrison, Analyst and AI Consultant
Simon Harrison, Analyst and Executive Partner

Summary

CRM and Contact Center as a Service (CCaaS) vendors are aggressively positioning their AI platforms as “agentic”, claiming systems that can autonomously reason, adapt, and act without being programmed for every scenario. While platforms have evolved from rigid rules-based systems to more flexible workflows, most still operate deterministically. As a result a discrepancy between marketing and technical reality is widening, creating a credibility problem in enterprise AI procurement. Agentic AI capability must be validated through clearer agentic AI definition, behavioral testing, and buyer-side skill development. This research note outlines what’s required to help buyers cut through the noise and hold vendors accountable for the value true agentic AI should be delivering.

Key Take: Too many “AI-enabled” CCaaS and CRM applications are deterministic systems masquerading as autonomous agents. To avoid buyer regret, enterprises must be clear on what agentic AI means, use real-world edge-case testing, and improve internal skills to identify true agentic behaviors.

Vendors Are Rebranding Automation as Agentic AI to Gain Competitive Edge Causing Confusion for Buyers

The enterprise software landscape is undergoing a paradigm shift, especially in the CRM and CCaaS markets. In an effort to stand out vendors are rushing to label their AI platforms as “agentic”, implying a level of autonomy and adaptability that most current systems have yet to achieve. End-users are challenged with navigating this while trying to ramp up their own AI competency.

© 2026 Actionary Ltd. All rights reserved. A license is required for any external use.

- **Vendors are aggressively marketing “agentic AI” as a differentiator.** AI branding has escalated from “automated” to “intelligent” to “agentic” in the past two years, particularly in the CRM and CCaaS space. Leading platforms now claim capabilities such as self-directed decision-making, adaptive routing, and autonomous service resolution. However, most systems still depend on deterministic flows (predefined, rules-based logic with fixed outcomes) and lack the ability to interpret unstructured inputs or adapt their reasoning in real-time. The gap between promise and capability is expanding as vendors chase market position rather than maturity, setting the stage for widespread buyer disappointment.

- **Vendors favor pre-recorded demos often concealing deterministic behavior behind scripted storytelling.** Demos are typically stitched together by teams to craft a narrative often leading to subtle inconsistencies that wouldn’t even be present if just Generative AI (GenAI). Mismatched naming, casing, or spelling, suggest pre-programmed outputs rather than live decision-making. Other giveaways include awkward flow dynamics, timestamp mismatches, repetitive data use, inconsistent response formats, or input expectations shifting without cause, (even inconsistencies should be consistently inconsistent). The result is a jigsaw of capabilities that look coherent, even to some experts, but often point to a lack of true agentic behavior.

- **Inconsistency in true agency definition means buyers are at risk of buying hype over function**. As with all markets there is a lack of knowledge and experience to support buying in the early stages of its lifecycle. End-user companies, especially those outside of AI-native industries, lack the exposure to agentic capabilities to evaluate whether a solution truly offers agentic benefits. This gap in understanding can be exploited by vendors using vague or inflated terminology. The issue is compounded when RFP processes rely on surface-level feature checklists rather than empirical testing or proofs-of-concept tied to real-world complexity.

Sacrificing Storytelling to Show True Agency Leads to Diminishing Returns

Vendors are under pressure to articulate agentic AI in a way that aligns with the technology they are selling. The concept of agency is being framed through characteristics that *appear* agentic and match what vendors can demonstrate. It has become a race, and the stakes could not be higher. Securing the role of central AI solution within an enterprise is perhaps the most important position a software vendor can achieve, since all other investment decisions will be secondary to the core AI-enabled “orchestration engine.” In many cases, demonstrating true agentic capabilities is sacrificed for the sake of telling a compelling story that ensures relevance. Unless this dynamic is addressed, the gap between expectations and real-world possibility will continue to widen.

© 2026 Actionary Ltd. All rights reserved. A license is required for any external use.

- **Agentic AI is fundamentally difficult to build while vendors are under pressure to develop faster.** True agentic AI requires not only advanced machine learning models, but also orchestration frameworks capable of chaining together multiple tools, maintaining state, and reasoning across incomplete or ambiguous data. These capabilities are not trivial to develop or deploy; they demand large volumes of high-quality training data, rigorous evaluation methods, and often, the integration of multiple sub-systems like planning agents, memory architectures, and retrieval-augmented generation. The pace of the market, accelerated by competitive pressure, is forcing vendors to declare agentic readiness long before it's achieved.

- **Vendors struggle to demonstrate agentic AI, making overclaiming almost inevitable.** True agency is difficult to demonstrate in ways that cannot be replicated by deterministic capabilities. Agentic behavior, such as goal-driven planning, tool selection, and decision-making under uncertainty, emerges from reasoning chains, contextual adjustments, and prioritization in the face of incomplete or conflicting data. Demo environments typically lack access to the enterprise data, systems, and workflows required to showcase these capabilities authentically. Even leading agentic AI companies acknowledge that robust agency involves emergent behavior, which is inherently unpredictable and difficult to present effectively in short sales cycles. As a result, marketing teams often rely on narrative stitching and scripted flows to simulate intelligence and present it as autonomy. This normalizes performative AI, erodes trust across the ecosystem, and delays the realization of value that genuine agency could deliver.

- **Opacity in messaging makes due diligence nearly impossible.** Vendors frequently use ambiguous language to describe capabilities, relying on terms like "intelligent" or "self-learning" or "understands context" without describing what this actually means or how it specifically translates to agentic AI capabilities. Else, vendors may reference uniqueness derived from having large proprietary datasets that increase accuracy (data washing). This isn't necessarily to purposefully deceive buyers, even capable staff often lack the expertise to describe AI capabilities well enough, themselves. In such a scenario, there's a risk staff members may quickly find themselves diluting messaging or worse, in an indefensible position when being challenged about real-world agentic AI capabilities. It makes sense to be sufficiently vague and to defer to the "experts" when buyers show intent to buy.

Establish Evaluation Rigor Now to Avoid Disappointment in Agentic AI Investments

To avoid investment mistakes in over-promised tools, end-user companies need to assess agentic AI claims rigorously using structured evaluation methods rather than relying on vendor definitions. Apply a standardized assessment framework, validate capabilities through proofs of concept that expose ambiguity, and upskill staff to recognize genuine behaviors that ensure agentic AI investments are better aligned to enterprise goals.

© 2026 Actionary Ltd. All rights reserved. A license is required for any external use.

- **Use Actionary's Agentic AI Assessment Framework for a standardized definition of agentic AI in enterprise software.** Our practical evaluation tool assists procurement and technology teams assessing CRM and CCaaS AI capabilities. It provides a clear list of characteristics that demystify what's truly agentic, (see "Appendix: Vendor Assessment Framework for Agentic AI" for associated guidance). It will help to ensure vendors explain, in concrete terms, how its system behaves when it *isn't* told exactly what to do.

- **Build proof-of-concept tests that simulate ambiguity and edge cases**. Agentic AI introduces a new kind of capability, and therefore requires a new kind of test. A Proof of Concept (PoC) designed specifically to validate *true* agentic behavior is the only reliable way to justify investment decisions in today's uncertain AI market — no matter how advanced the demo, it simply can't reveal whether a system is reasoning on its own. Unlike conventional AI PoCs that assess performance on predefined tasks with clean inputs, agentic AI PoCs must recreate the messiness of the real world: ambiguous goals, broken APIs, shifting objectives, and incomplete or multi-intent inputs. Does the system interpret goals independently, plan multi-step actions, adapt in real time, invoke relevant context, and revise its own strategy? Can it identify missing data, ask clarifying questions, select the right tool, and adjust its behavior across interactions? An agentic AI solution doesn't follow a fixed path, it navigates uncertainty, reasons through trade-offs, and learns from outcomes. That's why instrumentation must go deeper: trace decision paths, monitor memory retrieval, capture fallback logic, and observe how it updates its approach. Success is more than task completion, it's *how intelligently and independently* the system got there. And this kind of verification can only happen when PoCs are deliberately engineered to expose reasoning and resilience, not just functionality. Getting this right is how organizations separate real agency from scripted automation, and avoid the number one blocker in enterprise AI deployment today: wasting cycles in PoC purgatory. Reference the Appendix to rationalise the list of vendors to engage with on a PoC.

- **Build organizational skills to recognize and evaluate agentic behavior.** Procurement, product, and engineering teams must be skilled in critically assessing what agentic AI looks like in practice. Most enterprise buyers haven't been given the opportunity to develop expertise, internally, to be able to differentiate between automated workflows, traditional AI, and true agentic systems. Without a foundational understanding of how agentic AI behaves—goal orientation, tool use, planning, contextual awareness—teams risk misinterpreting deterministic automation as autonomy. Upskilling should focus on three core areas: (1) conceptual literacy about agentic AI properties (e.g., memory, chaining, self-directed decision-making), (2) technical validation skills such as prompt tracing, sandboxing, and structured PoC execution, and (3) cross-functional collaboration between business units and technical teams. Training should also cover how to design evaluation prompts and simulate edge cases to expose agentic limitations or false claims. Finally, organizations should maintain a shared playbook of evaluation methods and patterns learned from real PoCs to institutionalize this knowledge across future procurement cycles.

© 2026 Actionary Ltd. All rights reserved. A license is required for any external use.

Appendix: Vendor Assessment Framework for Agentic AI

The following tables provide a practical evaluation tool for procurement and technology teams assessing CRM and CCaaS AI capabilities.

Section 1: Core Capabilities of Agentic AI

Capability	Evaluation Q1	Evaluation Q2	Evaluation Q3
Autonomous Decision-Making	What types of decisions can your AI make without explicit programming?	Can the AI adapt its actions when presented with conflicting or incomplete data?	How does it handle novel scenarios that are not covered by pre-set flows?
Dynamic Goal Handling	How does the system reprioritize actions when customer needs change mid-interaction?	Can the AI interpret open-ended goals and generate a multi-step plan to achieve them?	What mechanisms allow it to switch focus without manual rule invocation?
Error Recovery	What does the system do when it encounters unfamiliar or malformed inputs?	Can the AI ask clarifying questions autonomously to resolve ambiguity?	How does it avoid cascading failures when part of the environment (e.g., API) is broken?
Real-Time Learning	How frequently does the system update its behavior from live user interactions?	What kind of feedback loops are implemented for immediate learning vs. batch updates?	Does learning occur without model redeployment or human retraining?
Behavior Adaptation	Can the AI self-modify its workflow or logic based on past outcomes?	How does the system track which adaptations led to successful results?	Are developers required to intervene to change the logic when performance drops?

© 2026 Actionary Ltd. All rights reserved. A license is required for any external use.

Capability	Evaluation Q1	Evaluation Q2	Evaluation Q3
Context Awareness	How does the system maintain context across a multi-turn interaction or channel shift?	Can it integrate past interactions into current decision-making?	How does it resolve inconsistencies or contradictions in remembered context?

Section 2: Technical Architecture Transparency

Capability	Evaluation Q1	Evaluation Q2	Evaluation Q3
Model Type Disclosure	What model types (symbolic, ML, LLM, hybrid) are used to power decision-making?	Are different models used for different capabilities (e.g., planning vs. dialog)?	How are model choices justified based on the use case?
Decision Logic Documentation	Can you show a representative decision path or logic tree from a real customer session?	Is the reasoning chain accessible and auditable by technical users?	How does the system explain or justify its actions to end users or admins?
Learning Loop Clarity	How is user interaction data collected and labeled?	How is this data integrated into model or logic updates?	What safeguards exist to prevent the use of bad or biased data?
Error Logging & Handling	What kinds of system errors are logged and how granular is the logging?	Are errors used to automatically trigger learning or retraining?	How is error recovery performance measured and reported?
Human Override & Intervention	In which scenarios does human intervention override the AI's actions?	Is fallback behavior configurable or hard-coded?	How often does override occur in production, and how is it tracked?

© 2026 Actionary Ltd. All rights reserved. A license is required for any external use.

Capability	Evaluation Q1	Evaluation Q2	Evaluation Q3
Auditability	Can a complete decision trail be generated for any given user interaction?	Is this audit trail timestamped and version-controlled?	Can it be exported or reviewed for compliance purposes?

Section 3: Functional Testing & Pilot Design

Capability	Evaluation Q1	Evaluation Q2	Evaluation Q3
Unstructured Query Handling	How does the AI interpret inputs with multiple intents or ambiguous phrasing?	Can it identify what's missing and ask for clarification?	How does the system avoid failure when it can't match an input to a known path?
Cross-Domain Reasoning	Can the system independently shift from one functional domain to another?	How does it determine when to change context or apply a different knowledge base?	Does it need separate models or modules for each domain?
Escalation Logic	What are the triggers for escalating a case to a human agent?	Are these thresholds adaptive or static?	Can escalation criteria be changed based on real-time context or user profile?
Feedback Loop Activation	How fast can the system improve from incorrect responses or failures?	Can changes from feedback be traced back to interaction-level triggers?	Is there visibility into what changes were made after a failed session?
Personalization at Scale	How is historical user data applied to improve current interactions?	Can the system differentiate user preferences and behaviors without hard-coding?	How is privacy preserved when storing or using personalization data?

© 2026 Actionary Ltd. All rights reserved. A license is required for any external use.

Capability	Evaluation Q1	Evaluation Q2	Evaluation Q3
Recovery from API Failure	What happens when an external system (e.g., API or DB) fails during a task?	Can the AI recognize failure and select an alternative path?	Does it attempt retries, notify the user, or adjust goals automatically?

Section 4: Evaluation Criteria Checklist for Procurement Teams

Evaluation Criteria	Explanation
AI demonstrates autonomous decision-making	No step-by-step scripting is required for routine interactions
AI adapts to real-time changes in user intent or context	Dynamic updates during conversation or workflow based on new input
System learns continuously from production usage	Uses online learning or feedback signals to improve without waiting for retraining cycles
Transparent logic and auditability of decisions	Decision chains and reasoning paths are inspectable and explainable
System was tested with ambiguous or edge-case inputs during PoC	Evaluation scenarios reflect real-world complexity, not just scripted demos
Demonstrated measurable reduction in human escalations	Evidence that AI reduces manual involvement in common support or CRM cases
Handles multi-intent queries and adapts responses accordingly	Interprets more than one goal or request and resolves without escalation
Recovers gracefully from failure of external dependencies	Shows fallback logic or creative problem-solving during outages or exceptions
Supports scalable personalization while preserving data privacy	Personalizes interaction flow without overfitting or violating compliance
Operates independently under complex, variable conditions	Self-manages and navigates trade-offs with minimal manual intervention

To speak to an analyst about this research Book an Inquiry.

© 2026 Actionary Ltd. All rights reserved. A license is required for any external use.