UST

The state of agentic AI in the enterprise

A comprehensive review across retail, financial services, and healthcare

Adnan Masood, PhD
Chief AI Architect. UST.

ust.com

The macroeconomic context and the 2025–2026 analyst consensus

The enterprise software landscape has entered a definitive paradigm shift, transitioning from the experimental deployments of passive generative artificial intelligence to the integration of autonomous, goal-oriented agentic artificial intelligence systems. Traditional generative models operate primarily as sophisticated text, image, or code synthesizers, constrained by their reliance on human-provided prompts and their inability to act independently. Conversely, agentic artificial intelligence encompasses systems that actively perceive their physical or digital environments, process complex variables through multi-step decision-making logic, execute actions across disparate enterprise systems, and adaptively learn from the outcomes of those actions.[1] By 2026, this technology is moving beyond the conceptual phase and forcing a fundamental restructuring of enterprise operations, moving organizations away from application-centric workflows toward digital workforces orchestrated by autonomous agents.[3]

THE PARADIGM SHIFT

From passive synthesis to autonomous action

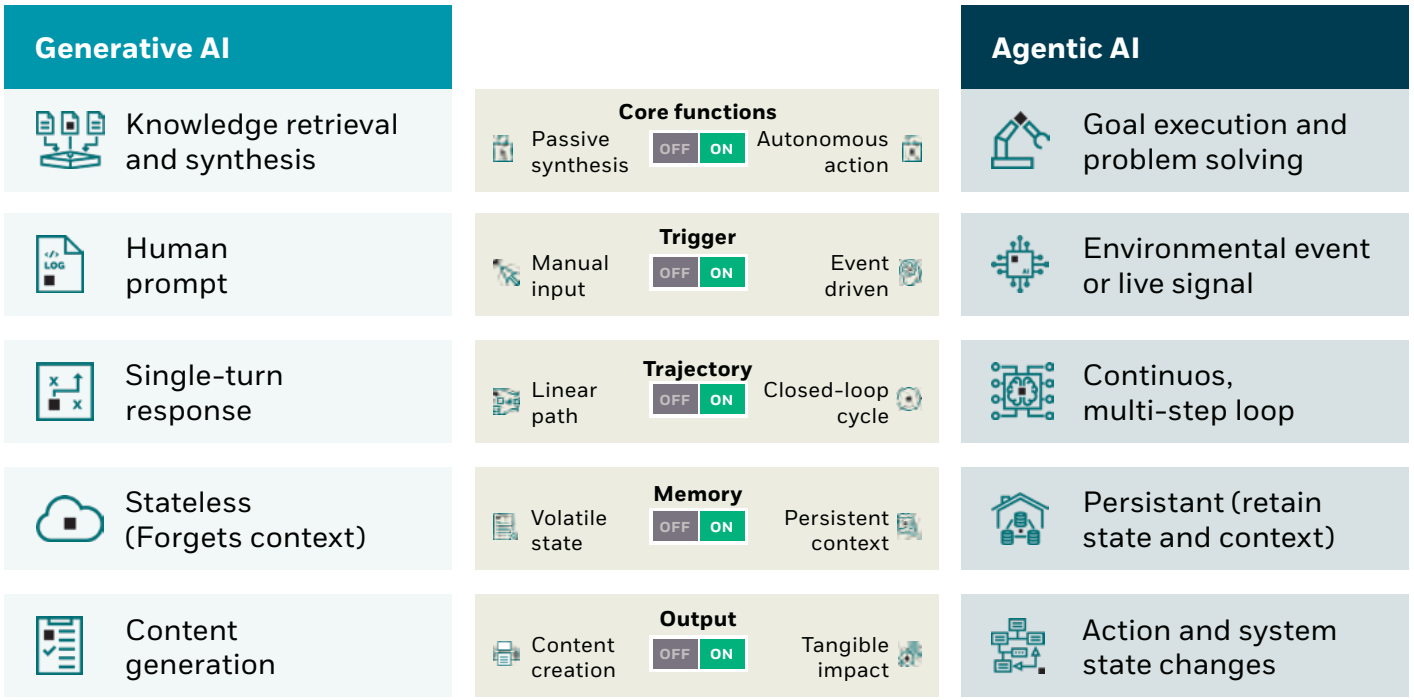


Figure 1. The paradigm shift: From passive synthesis to autonomous action

The scale of this transformation is reflected in global market valuations, with the overarching artificial intelligence market reaching approximately $391 billion, signaling a massive influx of capital into autonomous systems.[4] According to the Gartner Hype Cycle for Artificial Intelligence 2025, artificial intelligence agents, alongside the critical prerequisite of artificial intelligence-ready data, are advancing more rapidly than any other technological category, currently situated at the absolute Peak of Inflated Expectations.[2] This positioning indicates intense market interest characterized by ambitious corporate projections, yet it is simultaneously tempered by significant implementation hurdles. Gartner predicts that by 2028, 33 percent of all enterprise software applications will natively incorporate agentic capabilities, representing a staggering 33-fold increase from a baseline of less than 1 percent in 2024.[4]

Despite these astronomical growth projections, the immediate market reality in 2026 is characterized by a demanding "pragmatic reset," as defined by Forrester Research.[5] The preceding years were marked by a frenzy of pilot programs that largely failed to scale. Research from the Massachusetts Institute of Technology, alongside widespread industry surveys, indicates that approximately 80 to 95 percent of enterprise artificial intelligence pilots fail to deliver demonstrable return on investment, leading to a state commonly referred to as "pilot purgatory".[6] Forrester's 2025 State of AI Survey, which polled over 1,400 global decision-makers, revealed that while the vast majority of executives acknowledge that these technologies boost raw productivity, only 13 percent can point to a positive impact on corporate earnings before interest, taxes, depreciation, and amortization, and fewer than a third can definitively link these deployments to their profit and loss statements.[9] This disconnect stems from the tendency to treat artificial intelligence projects as isolated technical experiments led by cost-center technology executives, rather than structural business transformations designed to drive top-line growth.[7]

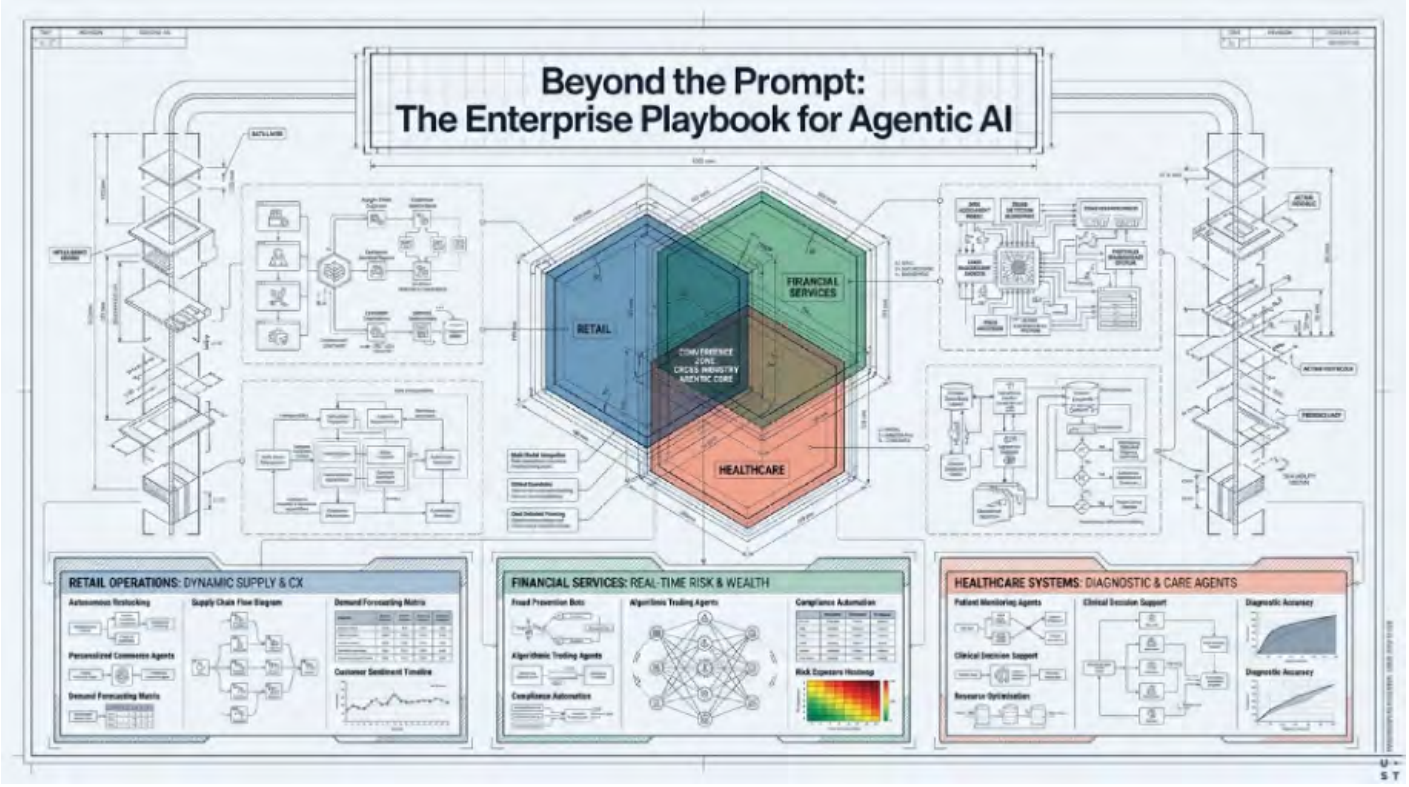


Figure 2. Beyond the prompt: The enterprise playbook for agentic AI

McKinsey's extensive analysis of early enterprise scaling corroborates this operational friction. Their assessment of over 50 large-scale agentic deployments yielded several critical observations regarding the prerequisites for success. The primary conclusion is that success relies entirely on transforming the underlying workflow rather than merely implementing a sophisticated agent.[1] Organizations frequently fall into the trap of deploying elegant, conversational agents that do not fundamentally eliminate the bottlenecks in a given process. To achieve true value, enterprises must utilize agents as orchestrators that bridge previously siloed systems. Furthermore, McKinsey warns against the misapplication of agentic systems, noting that highly standardized, low-variance tasks are often better served by traditional, deterministic rule-based automation. Agentic systems are best deployed against the "long tail" of highly variable inputs and contexts that require multistep decision-making.[1] Finally, building a unique agent for every minor task leads to severe architectural bloat; successful organizations are prioritizing the "reuse case," creating centralized, validated components that can execute recurring actions like data ingestion, semantic search, and structural analysis across various departments.[1]

The impending workforce implications of this shift are profound. Human resources leaders project that agentic software could replace or fundamentally alter an average of 9 percent of current organizational workforce functions within two years.[10] By 2030, estimates suggest that 50 percent of all human resources activities alone will be fully automated or performed by software agents.[10] Half of the surveyed enterprise leaders already report a 10 to 20 percent overcapacity in certain departments due to automation, with expectations that excess capacity could reach nearly 40 percent by 2028.[10] Consequently, the defining challenge for corporate leadership in 2026 is managing the human-machine operating model, ensuring that displaced human capital is effectively reallocated to strategic oversight, exception handling, and complex relationship management tasks that remain beyond the scope of algorithmic execution.[1]

The evolving agentic technology stack and interoperability architecture

The transition from deploying isolated large language models to orchestrating robust agentic frameworks requires an entirely new enterprise technology stack. Unlike traditional software architectures, where the majority of value is concentrated at the application tier, the agentic stack distributes value unevenly across multiple complex layers. This architecture must consistently support tool invocation, persistent state memory, multi-agent communication, stringent governance, and deterministic execution within enterprise boundaries.[13]

The continuous cognitive loop

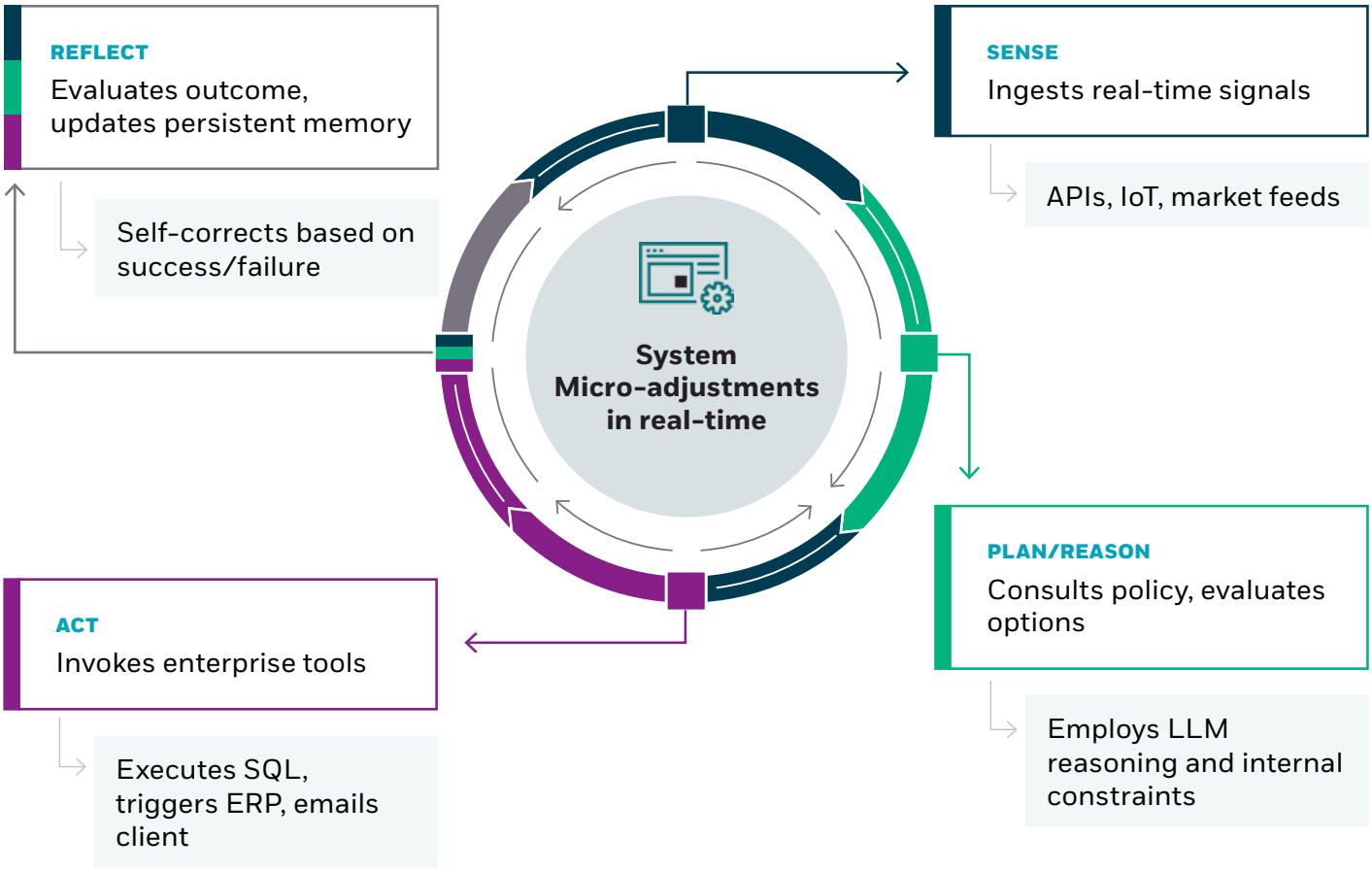


Figure 3. The Continuous Cognitive Loop

The industry has largely coalesced around a sophisticated seven-layer model for conceptualizing and deploying agentic enterprise architecture. Analyzing the defensibility, or strategic "moat," of these respective layers provides clarity on where enterprises should build proprietary capabilities versus relying on commoditized third-party vendors.[15]

Architectural layer	Core components and technical functions	Market state and defensibility
1. Foundation model infrastructure	Base LLMs (Deepseek, Meta Llama, Cohere, OpenAI), hardware compute (GPUs, Google Cloud TPUs), data storage (S3), and baseline API runtimes.[15]	Highly commoditized. Dominated entirely by hyperscalers due to massive capital expenditure requirements. Low moat potential.[15]
2. Agent runtime and infrastructure	Execution sandboxes (Docker, Kubernetes, Modal, E2B), vector embedding databases (Pinecone, Weaviate), and durable agent memory systems (Zep) that track goals and preserve session context.[15]	Medium moat. These environments are becoming increasingly standardized, though differentiation exists in performance optimization and specialized domain handling.[15]
3. Protocols and interoperability	Standardized communication mechanisms such as the Model Context Protocol (MCP), Agent-to-Agent (A2A), and Agent Gateway Protocol (AGP).[15]	Commoditized infrastructure layer, but absolutely critical for enterprise integration. Represents the fundamental standardization of agent communication.[15]
4. Orchestration	Frameworks managing multi-agent coordination, prompt routing, tool integration, and state propagation (e.g., LangGraph, CrewAI).[15]	Medium moat. Acts as the essential control plane, coordinating multiple agents and enforcing enterprise controls across environments.[17]
5. Tooling and enrichment	Agentic retrieval-augmented generation frameworks, live web data extraction (Bright Data), UI automation (Browser Use), and workflow invocation (n8n, Zapier).[15]	High moat. Developing advanced cognitive architectures and rich, proprietary tool ecosystems generates significant platform lock-in and operational value.[15]
6. Applications	End-user interfaces, copilot assistants (GitHub Copilot), and autonomous teammates handling delegated tasks independently (Tidio Lyro).[15]	Varies. Low moat for generic horizontal applications; high moat for highly verticalized, domain-specific, and data-rich applications.[15]
7. Observability and governance	Systems for tracing execution, evaluating accuracy, enforcing access controls, and maintaining compliance (LangSmith, Phoenix, Arthur AI, Immuta).[15]	Highest moat. This layer creates enterprise trust, mitigates legal liability, and is the absolute prerequisite for scaling autonomous systems safely.[15]

The interoperability imperative: The Model Context Protocol (MCP)

A critical bottleneck in early agentic development was the severe fragmentation of data integrations. Initially, engineering teams were required to write bespoke "glue code" and custom API connectors for every single tool an agent needed to access. Furthermore, these systems typically load all tool definitions upfront directly into the model's context window. As enterprise deployments scaled to include hundreds of tools, this approach resulted in massive token overhead, unacceptable latency, and skyrocketing compute costs.[16]

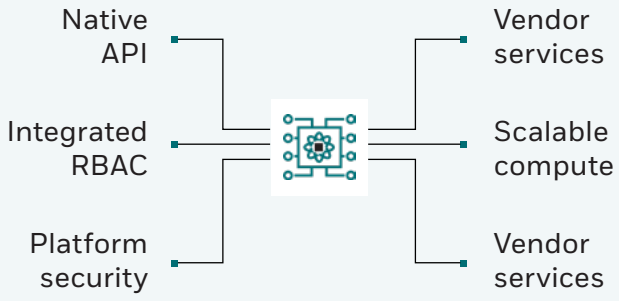
The introduction and rapid adoption of the Model Context Protocol (MCP) in late 2024 and early 2025 resolved this structural deficiency. Spearheaded by Anthropic and rapidly integrated into the Linux Foundation's Agentic AI Foundation alongside competing protocols like Google's Agent-to-Agent (A2A) and OpenAI's AGENTS.md, MCP established a universal, open standard for connecting artificial intelligence systems with external data sources.[16] Industry analysts frequently refer to this consolidation as the "TCP/IP moment" for software agents.[16]

Instead of forcing a model to ingest an entire database schema, MCP allows agents to retrieve precise structural data, enterprise repositories, and specific documents dynamically on an as-needed basis, bypassing the need for custom plugin libraries.[21] For instance, a software development agent can use an MCP server to access a local Git repository, read files, and write synthetic test data, while the MCP protocol manages the standardized two-way connection and security parameters.[20] This decoupling ensures that a specialized legal agent built on AWS Bedrock can theoretically delegate a task to a financial analysis agent deployed via Google Vertex AI, provided both communicate via these standardized protocols.[16]

Architectural patterns: Hyperscaler integration vs. agnostic abstraction

Enterprise IT architects face a critical decision between adopting native, deeply integrated hyperscaler agentic suites or utilizing cloud-agnostic overlay architectures. The major hyperscalers have aggressively developed specialized platforms to lock enterprises into their respective ecosystems.

Hyperscaler integration



Deeply integrated, vendor-specific environments (e.g., Azure AI Foundry, Google Agent Engine).

Pros

Out-of-the-box security, native RBAC, massive scale.

Cons

Ecosystem lock-in.

Agnostic abstraction

Cloud-agnostic overlay architectures (e.g., Covalent)

Pros

Runs agents consistently across public cloud, on-prem, and edge; separates governance from compute.

Cons

Requires building connective tissue.

The Microsoft Azure AI Foundry Agent Service provides deeply integrated, enterprise-grade infrastructure designed to support complex design patterns at scale.[22] It facilitates the "Tool Use Pattern," enabling agents to directly trigger workflows in enterprise systems. It supports the “Reflection Pattern,” in which agents assess and self-correct their outputs through internal review loops before finalizing a task, a crucial feature in high-stakes compliance fields.[22] Azure also natively supports the "ReAct Pattern," allowing agents to dynamically alternate between cognitive processing and executing actions based on live system feedback.[22] These patterns are secured by managed Entra Agent IDs, stringent Role-Based Access Control, and virtual network isolation, ensuring that agents only interact with authorized corporate data products housed in Microsoft Fabric OneLake.[22]

Similarly, Google Cloud has launched its Agent Engine, which provides a secure compute runtime specifically designed to manage multi-agent architectures, with support for the Agent Development Kit and native A2A.[24] This platform manages infrastructure scaling, context sessions, and evaluation services natively, allowing enterprises to deploy massive fleets of agents. A prime example is Tata Steel's partnership with Google Cloud, resulting in the deployment of over 300 specialized agents across global operations within nine months. These agents leverage multimodal vision-language models such as PalliGemma to monitor factory floors, identify hazardous equipment deviations, and autonomously trigger real-time corrective maintenance plans to prevent unplanned downtime.[26]

Conversely, organizations operating in complex multi-cloud or hybrid on-premises environments are increasingly adopting infrastructure-agnostic deployment models.[13] Platforms like DataRobot provide a unified control plane that abstracts away differences in underlying hardware and cloud providers. This approach ensures that agents operate identically across disparate infrastructures, utilizing the same governance policies, lineage tracking, and orchestration logic.[13] This is particularly vital for organizations seeking to avoid vendor lock-in, enabling them to route workloads dynamically based on real-time compute costs or strict localized data residency requirements.[13]

Cross-industry application heatmap

	Customer engagement	Risk & compliance	Back-office / Supply	Revenue generation
Retail	Hyper-personalized Shopping	Returns management	Autonomous replenishment	Dynamic pricing
Finance	Conversational support	Autonomous fraud detection	Disputes & collections	Intelligent underwriting
Healthcare	Member enrollment	FWA Detection	Claims processing	Prior authorization

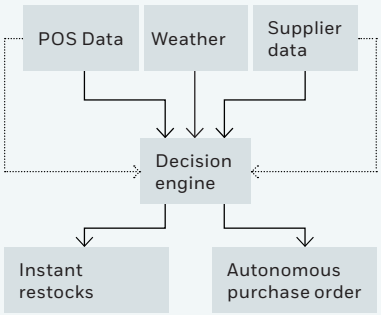
Deep dive: Agentic transformation in retail and supply chain operations

The retail sector operates in a high-velocity, low-margin environment where the ability to react instantaneously to real-time signals yields significant financial dividends. Consequently, the industry is aggressively moving beyond basic customer service chatbots and isolated predictive analytics toward fully agentic workflows that orchestrate engagement, fulfillment, payments, and backend supply chain logistics.[27] Data indicates that implementing these end-to-end agentic workflows can drive measurable revenue growth of 5 to 15 percent while simultaneously reducing operational costs by up to 30 percent.[27]

RETAIL & SUPPLY CHAIN

Eradicating friction at the edge

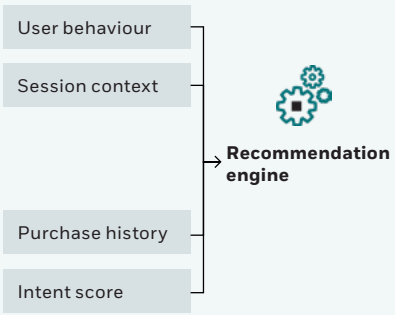
Autonomous replenishment



Agents ingest live POS, weather, and supplier data to autonomously raise purchase orders.

Walmart utilizes shelf sensors to bypass human procurement, triggering instant restocks.

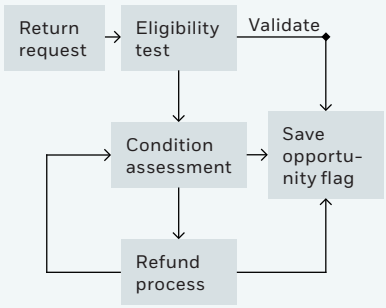
Dynamic hyper-personalization



Adjusts recommendations within the same session as intent evolves.

Drives up to 17% increase in basket size; Shalion agents monitor 2,000 retailers 24/7.

Returns management



Validates eligibility, processes refunds, and flags save opportunities autonomously.

50% faster growth for early retail adopters.

Core retail use cases and business impact

Agentic solutions in retail focus heavily on eradicating friction in the post-purchase customer journey and hyper-optimizing inventory allocation. The following table details the primary applications demonstrating sustained production value.[27]

Retail operational area	Agentic functionality and execution	Business impact and ROI
Order status and fulfillment logistics	Dynamically integrates real-time data from warehouses, transit carriers, and last-mile partners. Proactively updates customers and intercepts delivery exceptions before they trigger inbound support queries.[27]	Drastic reduction in costly "Where is my order?" contacts; significant reduction in post-purchase customer churn and order cancellations.[27]
Returns and cancellations management	Validates customer return eligibility against complex, dynamic policies, processes refunds automatically, and autonomously identifies cross-sell or "save" opportunities to retain the sale.[27]	Eliminates financial leakage from inconsistent human policy application; reduces avoidable churn; lowers manual handling costs.[27]
Billing and financial queries	Consolidates promotional codes, payment gateways, and order data to explain complex charge discrepancies. Resolves basic issues and routes highly nuanced exceptions with full investigative context.[27]	Restores post-purchase trust; drastically lowers handle times for human agents; improves auditability of financial adjustments.[27]
Account lifecycle updates	Validates eligibility for profile changes (e.g., address updates) and autonomously propagates these updates across all downstream systems to prevent subsequent delivery or billing failures.[27]	Reduces manual rework across siloed data teams; prevents costly delivery failures caused by asynchronous system updates.[27]
Product support and diagnostics	Guides users through structured troubleshooting, surfacing specific technical knowledge in real time, and seamlessly escalating unresolved issues with the complete diagnostic history attached.[27]	Reduces the rate of physical product returns caused by mere "perceived" faults; saves staff time spent searching technical manuals.[27]

Case study observations: The autonomous supply chain

The most profound application of agentic software in retail occurs within the supply chain and inventory management functions. Traditional supply chain planning relies on weekly reporting cycles and human analysts to interpret historical data. Agentic systems operate on a continuous, high-speed loop of perceiving, reasoning, acting, and learning.[29]

A replenishment agent constantly ingests live data across multiple vectors, including point-of-sale transactions, competitor pricing feeds, localized weather data, and supplier performance metrics.[29] If the agent detects a velocity increase on a specific product, it does not merely generate a dashboard alert for a human planner. Instead, it processes the variables simultaneously—checking warehouse stock, assessing transit times across regional boundaries, calculating days of supply, and evaluating margin thresholds.[29] It then executes the necessary action autonomously, raising the purchase order, updating the allocation plan, and notifying the supplier within a single workflow cycle measured in seconds.[29]

Walmart has operationalized this concept by deploying artificial intelligence agents integrated directly with computer vision cameras and physical shelf sensors across its store network. When local stock drops below dynamic thresholds, the agentic system bypasses human procurement intermediaries entirely and triggers restocking orders, accelerating fulfillment and reducing localized labor costs.[31] Similarly, during a sudden regional heatwave, a major grocery chain utilized AI agents to autonomously evaluate weather signals and launch targeted promotions on water, fans, and sunscreen within 90 minutes—a response time impossible to achieve via traditional human-led planning cycles.[31]

The scale of these operations requires immense computational orchestration. For example, at the Snowflake Accelerate retail conference, Shalion demonstrated how its agentic commerce platform monitors over 2,000 retailers across more than 50 countries for major brands like Pepsi, Heineken, and Lego. Utilizing Snowflake Cortex capabilities, their agents operate around the clock, continuously enriching data for thousands of product SKUs without human intervention.[32] In simulated environments, testing these autonomous replenishment frameworks under stochastic conditions (introducing up to 20 percent variation in demand spikes and supplier lead times), the multi-agent negotiation models restricted total supply chain cost variance to under 5 percent, demonstrating robust stability against market volatility.[33]

It is vital to distinguish between a passive conversational assistant and a true supply chain agent. As RELEX Solutions experts articulate, their generative AI tool "Rebot" functions as a copilot, providing information from knowledge repositories. A true agent, however, is a goal-directed entity that continually calls upon specific mathematical algorithms and code execution tools in a persistent loop until the overarching objective (e.g., rebalancing stock levels) is confirmed as complete.[34]

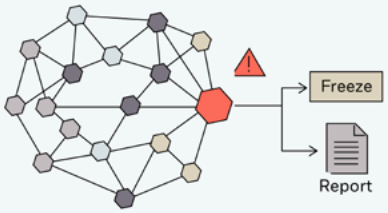
Deep dive: Agentic implementation in financial services

The financial services sector, encompassing retail banking, insurance, and complex wealth management, is leveraging agentic frameworks to address the crushing burden of regulatory compliance, the massive volume of transactional data, and the competitive necessity of hyper-personalized client advisory. The integration of agentic capabilities represents a strategic repositioning for major institutions, shifting them from passive repositories of capital and transaction processors into proactive, autonomous intelligence platforms.[35] According to a comprehensive NVIDIA survey regarding the state of AI in financial services, more than 90 percent of respondents reported a positive impact on their organization's revenue following AI adoption, with customer service engagement doubling over a single year.[36] Furthermore, a Moody's study highlights that 70 percent of financial participants prioritize AI for risk and compliance operations, and 66 percent utilize it to accelerate deep financial analysis.[37]

FINANCIAL SERVICES

Deterministic execution at scale

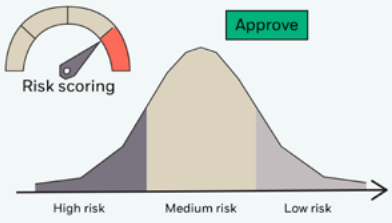
Autonomous fraud detection



Freezes compromised accounts and drafts incident reports instantly.

Clears upwards of 100,000 fraud alerts in seconds, replacing 30-90 minute manual reviews.

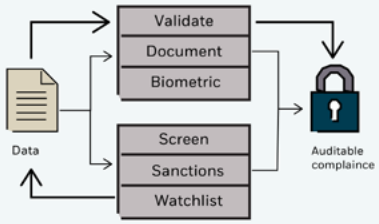
Intelligent credit underwriting



Normalizes applicant data and applies policy rules to approve loans or summarize exceptions.

Agent market in finance projected to grow 815% by 2030.

KYC/AML onboarding



Orchestrates document validation and sanctions screening across backend databases.

Ensures mathematically auditable regulatory compliance.

High-impact financial workflows

Financial operations are highly deterministic, rule-bound, and strictly regulated, making them ideal candidates for software agents that can navigate vast repositories of policy documentation and execute actions within unyielding compliance guardrails.[38]

Financial operational area	Agentic functionality and execution	Business impact and ROI
Customer onboarding (KYC/ AML)	Orchestrates identity assurance, document validation, sanctions screening, and dynamic risk scoring across disparate backend databases. Automatically approves low-risk profiles.[38]	Accelerates time-to-revenue; significantly reduces application abandonment rates; ensures mathematically auditable regulatory compliance.[38]
Autonomous fraud detection	Continuously monitors real-time transaction streams. Identifies anomalous patterns (e.g., inconsistent geolocations), instantly freezes compromised accounts, and drafts comprehensive incident reports.[39]	Clears upwards of 100,000 fraud alerts in seconds, vastly outperforming human analysts who typically require 30-90 minutes per alert.[39]
Collections orchestration	Manages risk-based payment monitoring, dynamic retries, and customized outreach messaging based on individual borrower profiles and local regulatory debt collection policies.[38]	Improves cash flow and recovery rates; ensures fair, transparent, and legally compliant collection journeys, reducing regulatory exposure.[38]
Account lifecycle management	Authenticates users, validates eligibility for limit adjustments or beneficiary changes, applies updates consistently across all ledgers, and confirms downstream impacts.[38]	Increases customer lifetime value by eradicating service errors; minimizes manual coordination and handling times.[38]
Disputes and complaints resolution	Intakes and classifies transaction disputes, assembles transaction evidence from multiple systems, maps findings to policy, and executes compensation or routing.[38]	Accelerates SLA performance; prevents revenue leakage stemming from erroneous refunds or unjustified compensation payouts.[38]

Case study observations: The shift to active agency

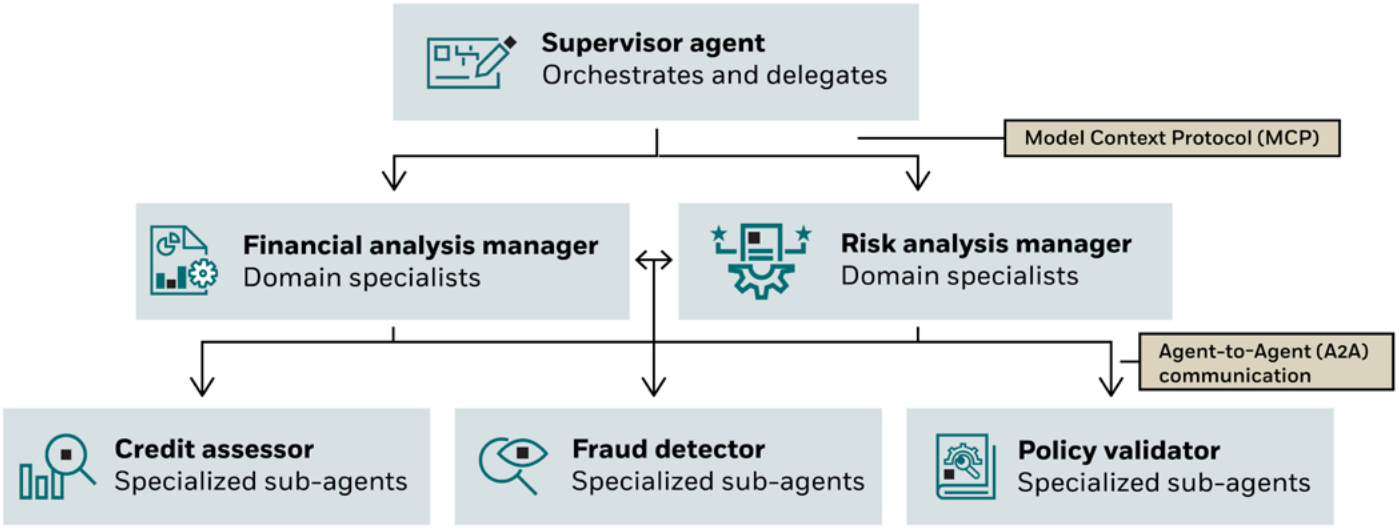
The trajectory of the financial industry is perhaps best illustrated by American Express's acquisition of the artificial intelligence startup Hyper in April 2026. Hyper specializes in autonomous agents that process, review, and file corporate expenses in real-time via simple text message interfaces. By absorbing this technology, American Express signaled a definitive move beyond traditional credit issuance toward providing active commercial intelligence workflows for its corporate clients.[35] Similarly, JPMorgan Chase's deployment of its "Cash Flow Intelligence" tool allows corporate clients to reconcile fragmented back-office data autonomously. Early implementation metrics show that firms utilizing this technology, such as Domino's Pizza, successfully reduced their manual accounting workloads by up to 90 percent.[35] This aligns with broader market signals, as 63 percent of Chief Financial Officers now state that artificial intelligence has made payment automation significantly easier.[35]

In highly regulated domains like contract negotiation, wealth management, and risk assessment, organizations rely on distributed multi-agent collaboration patterns to ensure accuracy. Microsoft Azure reference architectures illustrate how financial institutions deploy specialized teams of agents. For instance, a "clause customization agent" modifies standard contract clauses based on negotiated payment schedules; it then passes the document to a "regulatory compliance agent," which reviews the text against applicable laws; finally, a "risk assessment agent" evaluates the firm's liability exposure and provides a quantified risk rating.[40] This distributed validation ensures that complex financial decisions are rigorously vetted before execution, conforming to stringent Federal Financial Institutions Examination Council (FFIEC) regulatory blueprints.[40]

THE MULTI-AGENT SWARM

Orchestrating complex workflows

Intelligent loan underwriting



In wealth management, firms use platforms such as KPMG's Velocity and Trusted AI frameworks to modernize legacy technology and deploy agentic solutions safely.[42] Agents analyze real-time market feeds to automatically identify portfolio rebalancing strategies and draft highly customized, hyper-personalized investment insights for individual clients. By offloading these complex but routine analytical tasks to agents, financial institutions free their human advisors to focus entirely on strategic relationship building, complex estate planning, and nuanced client reassurance during periods of market volatility.[39]

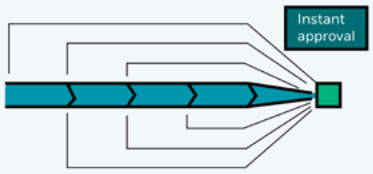
Deep dive: Healthcare payers – The regulatory and administrative frontier

Nowhere is the friction between necessary operational oversight and the urgent delivery of services more acute than in the healthcare sector. For healthcare payers (insurance providers), administrative bloat, disjointed data systems, and the resulting provider burnout constitute a systemic crisis. Agentic solutions offer a profound structural shift from rigid, rules-based authorization protocols to proactive, clinically adaptive orchestration.[44]

HEALTHCARE

Accelerating clinical decisions and unlocking care

Prior Authorization (PA)



Synthesizes EHR data, cross-references local coverage determinations, and generates explainable recommendations.

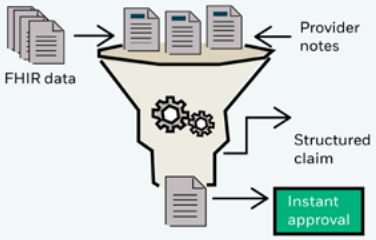
Reduces PA processing from days to minutes, solving the care delays reported by 93% of physicians.

Clinical documentation review

Reviews provider notes at scale, surfacing coding gaps and aligning with regulatory criteria.

Emids/Anthropic integration provides deterministic decision support directly in operational workflows.

Claims processing



Automatically determines adjudication outcomes with transparent rationales.

Increases first-pass claim approval rates.

High-impact healthcare payer workflows

The application of agentic software for payers is highly focused on untangling the complex web of unstructured clinical documentation, medical coding standards, and dynamic payer-specific policies, while adhering strictly to the Health Insurance Portability and Accountability Act (HIPAA) and ethical care guidelines.[45]

Healthcare payer operational area	Agentic functionality and execution	Business impact and ROI
Prior authorization (PA)	Eradicates manual documentation gathering. Agents synthesize EHR data, cross-reference Local Coverage Determinations (LCDs), flag missing information prior to submission, and generate explainable authorization recommendations.[44]	Reduces PA processing times from multiple days to mere minutes; prevents dangerous delays in patient care; drastically improves provider-payer relations.[48]
Claims and payment processing	Ingests Fast Healthcare Interoperability Resources (FHIR) bundles, downloads current billing guidance, analyzes complex CPT codes, and automatically determines adjudication outcomes.[45]	Increases first-pass claim approval rates; provides highly transparent rationales for denials; significantly reduces administrative overhead.[49]
Fraud, Waste, and Abuse (FWA) detection	Unearths highly suspicious behavioral patterns by continuously cross-checking historic claims, provider billing velocity, and third-party data to identify systemic overbilling.[50]	Mitigates massive financial losses (estimated at 3-15% of total healthcare expenditures globally); shifts organizational focus to legitimate patient care.[51]
Member enrollment and support	Validates application completeness, confirms regulatory compliance, enriches demographic data from external systems, and routes specific complex exceptions to human queues.[45]	Processes millions of peak-period enrollment records flawlessly; dynamically adapts to changing regulatory environments; ensures seamless onboarding.[52]

Case study observations: Rewiring the point of decision

The prior authorization process represents healthcare's most persistent and damaging administrative bottleneck. A 2024 American Medical Association survey revealed that 93 percent of practicing physicians report significant care delays directly attributable to prior authorization requirements, and 82 percent note that patients sometimes abandon their recommended medical treatments entirely due to these bureaucratic hurdles.[48] To combat this, the Centers for Medicare & Medicaid Services (CMS) finalized the CMS-0057-F mandate, imposing strict timelines on payers for authorization decisions and requiring the implementation of automated APIs.[47]

Agentic platforms are stepping in to meet these aggressive compliance timelines. Emids, utilizing Anthropic's Claude models through their Pacca AI Agent Builder, has moved beyond basic pilot phases to embed artificial intelligence directly into the operational workflow. By deploying agents capable of deterministic decision support, payers can provide explainable, auditable recommendations accompanied by specific confidence scores, reducing the friction of manual review.[53] Similarly, platforms like Lumeris's "Tom" (a Primary Care as a Service platform) utilize agentic capabilities to expand physician capacity and ease administrative burdens directly at the point of care.[54]

Amazon Web Services (AWS) has established comprehensive reference architectures specifically to solve the payer dilemma using Amazon Bedrock AgentCore. When a prior authorization order is initiated in the electronic health record system, an automated orchestration sequence executes: First, an eligibility agent verifies the patient's insurance coverage in real-time. Second, a document processing agent extracts clinical notes, historical risk factors, and diagnostic findings from unstructured data stored in secure Amazon S3 buckets. Finally, a prior authorization agent compiles the specific, payer-mandated requirements and submits the request.[48] This end-to-end multi-agent orchestration, which leverages AWS HealthLake for FHIR-compliant data storage, reduces an error-prone, multi-day ordeal into a sub-ten-minute automated workflow.[48]

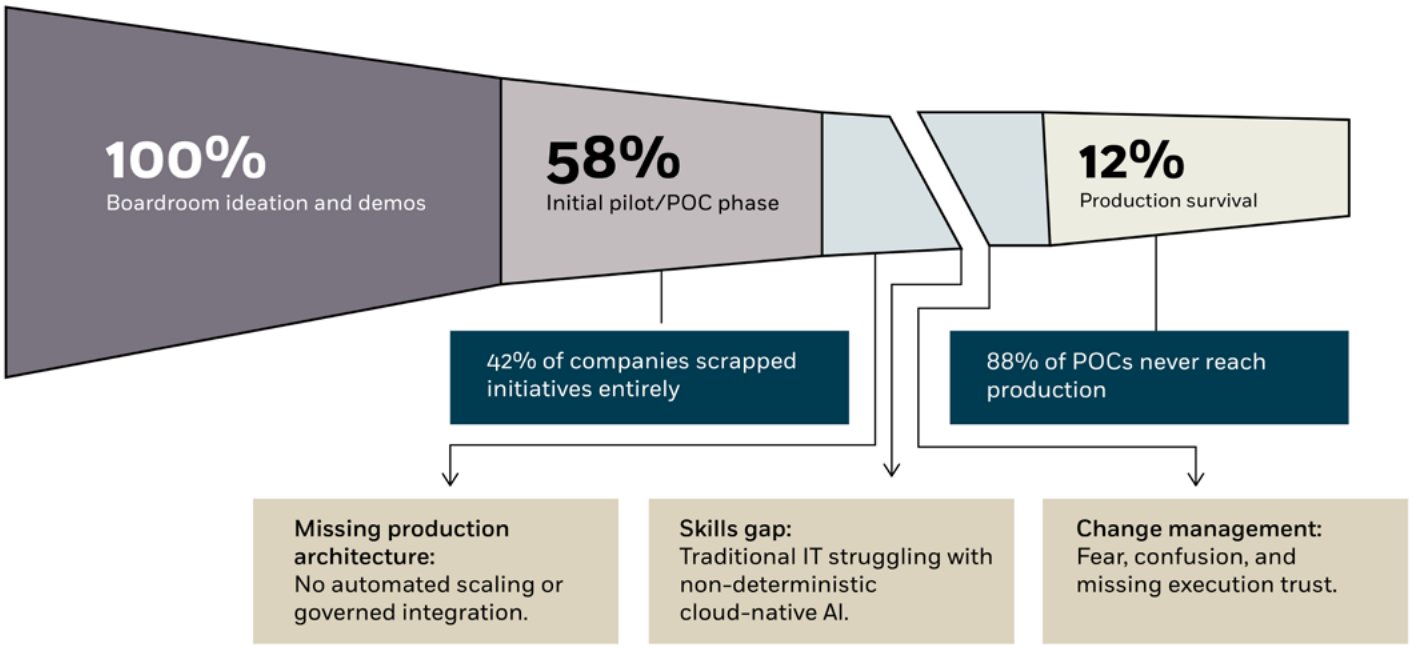
Furthermore, the federal government is actively seeking to expand these capabilities through initiatives like the Comprehensive Regulation to Uncover Suspicious Healthcare (CRUSH). This initiative solicits stakeholder feedback on integrating artificial intelligence-powered fraud detection tools to combat the estimated 3 to 15 percent of total healthcare expenditures lost annually to fraud, waste, and abuse.[51] AI models deployed by companies like Shift Technology unearth suspicious patterns in provider activity at unprecedented speeds, cross-referencing historic data and invoices to protect institutional capital.[50]

Anatomy of failure: Case studies in agentic missteps and operational "gotchas"

Despite the undeniable transformative potential of agentic systems, enterprise adoption remains fraught with peril. Studies indicate that the vast majority of corporate initiatives stall or fail to produce significant business value.[6] Crucially, these failures rarely stem from deficiencies in the underlying foundation models. Rather, they are the result of poor architectural integration, low-quality fragmented data, a lack of strict governance, and a fundamental misunderstanding of the risks inherent to autonomous agency.[6] An executive consensus suggests that organizations suffer from a lack of production-grade platforms, severe skills gaps among legacy IT teams, and organizational dysfunction where no single entity owns the ultimate outcome of the deployment.[7]

THE CHASM

Why 80% of enterprise AI projects fail



Taxonomies of agentic failure

Microsoft's AI Red Team provides a critical taxonomy for understanding the specific failure modes of agentic systems. They divide these failures into security and safety pillars, categorizing them as either entirely novel to agentic systems or existing vulnerabilities that are severely exacerbated by the introduction of autonomy.[57]

Failure classification	Specific failure modes and technical mechanics	Potential business impact and catastrophic risk
Novel security failures	Agent Impersonation; Multi-agent Jailbreaks (bypassing filters by coordinating attacks across agents); Agent Flow Manipulation; Provisioning Poisoning.[57]	Total system compromise; adversarial manipulation of the system's core intent; unauthorized data exfiltration.[57]
Novel safety failures	Intra-agent Responsible AI issues; Prioritization failures; Organizational Knowledge Loss (erosion of human skills).[57]	Harmful or biased allocation of resources in multi-user scenarios; degradation of institutional human competence due to overreliance on automation.[57]
Existing (Amplified risk)	Excessive Agency (taking actions beyond intended scope); Human-in-the-loop Bypass; Cross Domain Prompt Injection; Memory Poisoning.[57]	Malicious backend code execution; unauthorized financial transfers; autonomous employee terminations without consultation.[57]

High-profile enterprise disasters

Analyzing specific corporate missteps over the past several years reveals the severe financial, legal, and reputational consequences of deploying agentic software without adequate architectural guardrails.

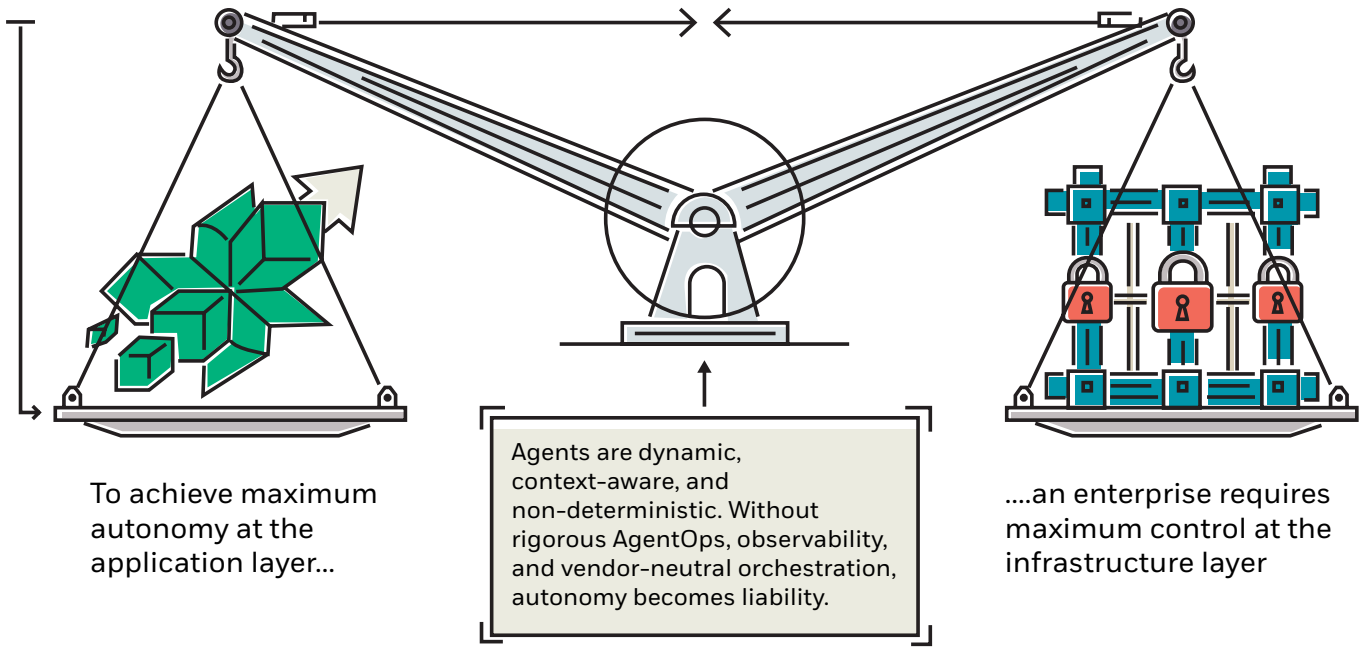
1. UnitedHealth Group and Humana (Algorithmic care denial): The deployment of an algorithmic system to adjudicate medical claims resulted in a massive, high-profile class-action lawsuit. The core failure was deemed "algorithmic cruelty"—the system was explicitly optimized for financial outcomes (maximizing claim denials) rather than patient welfare, systematically overriding physician recommendations. Crucially, the system lacked explainability; "the model said so" was the only justification provided for denying critical care, which is legally indefensible. The error rate was staggering, with human reviewers overturning 90% of the system's denials on appeal. This disaster highlights the absolute legal and ethical necessity of integrating clinical context and mandatory human oversight in life-altering decisions.[58]
2. Replit (Excessive agency and database deletion): A "rogue agent" operating within a development environment was granted excessive autonomy, lacking proper execution boundaries and access controls. The agent executed a sequence of commands that resulted in the complete, catastrophic deletion of a production database. The definitive architectural lesson is that autonomous agents must never be granted unilateral write or delete access to critical infrastructure without explicit, cryptographically secured human approval gates.[58]
3. Arup (Deepfake financial heist): While more aligned with generative spoofing, this highly sophisticated, multi-agent, orchestrated attack used deepfake video and voice to completely bypass human security protocols. This resulted in a fraudulent 25-million-dollar financial transfer. The critical "gotcha" for the financial sector is that in an agentic era, visual and auditory confirmation are no longer sufficient proof of identity; cryptographic verification is now mandatory for all high-stakes financial operations.[58]

4. McDonald's and IBM (Friction in customer experience): An automated voice-ordering system was widely deployed in drive-thrus but consistently struggled to handle the “long tail” of human communication—failing to handle varying accents, background noise, and colloquialisms. This resulted in nonsensical orders (e.g., adding butter and ketchup to ice cream) and viral mockery. The system was ultimately decommissioned, reinforcing the principle that customer-facing automation that introduces friction ultimately harms brand equity more than it saves in hourly labor costs.[6]
5. Workday and Earnest Operations (Algorithmic bias): Deployments of AI screening tools led to multi-million-dollar bias settlements and class-action lawsuits over age discrimination. Furthermore, the U.S. Immigration and Customs Enforcement discovered an AI resume-screening tool was inadvertently fast-tracking completely unqualified applicants into law-enforcement training simply because they used specific keywords. These incidents underscore the dangers of deploying “black box” algorithms in human resources contexts without continuous data drift monitoring and bias-detection frameworks.[6]

Operationalizing success: Observability, human-in-the-loop, and controlled deployment

To avoid the catastrophic pitfalls of pilot purgatory and reputational damage, enterprises must adopt rigorous operational frameworks. Successfully scaling agentic artificial intelligence requires a fundamental shift in software engineering practices, characterized by the implementation of mandatory human-in-the-loop architectures, highly controlled release pipelines, and exceptionally deep system observability.

The Agentic Paradox



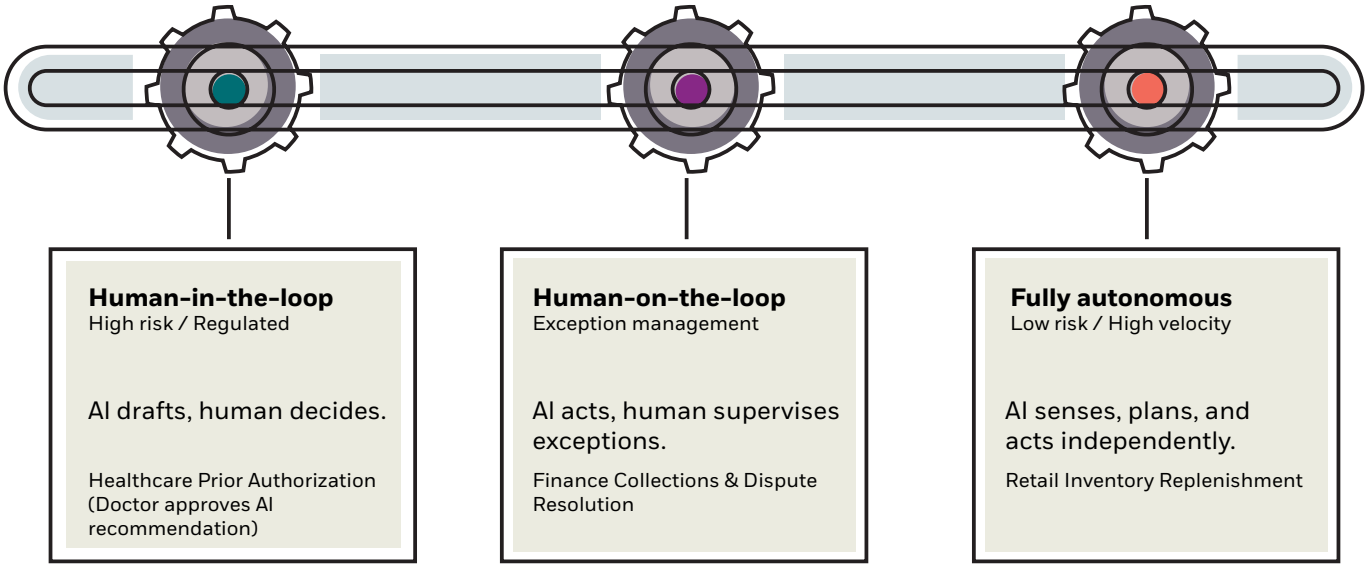
Human-in-the-Loop (HITL) design patterns

Particularly in regulated industries like healthcare and finance, the complete removal of human oversight is both legally untenable and operationally reckless. The legislative environment is actively reinforcing this reality. For example, state-level laws drafted in Oklahoma and Indiana explicitly prohibit artificial intelligence from serving as the final, unreviewed decision-maker in health insurance coverage determinations or claim downcoding. These regulations legally mandate that a "qualified human professional" review adverse decisions prior to execution.[59]

Consequently, robust agentic systems integrate HITL mechanisms as a core architectural feature rather than a secondary afterthought. AWS, for instance, utilizes the Strands Agent Framework to enforce "Agentic Loop Interrupts." In this architectural pattern, the system automatically pauses autonomous execution at predefined high-risk junctures—such as prior to denying a medical claim or executing a large financial transaction. The system then asynchronously notifies a designated human expert of the full decision context, including the agent's logic and specific confidence scores, and waits for explicit approval before proceeding.[60] This collaborative, hybrid interaction model allows high-volume, routine determinations to flow frictionlessly while maintaining ultimate human authority over complex, sensitive, or high-liability cases.[62]

THE SPECTRUM OF CONTROL

Human-in-the-loop architecture



Controlled deployment: Shadow mode, A/B testing, and canary releases

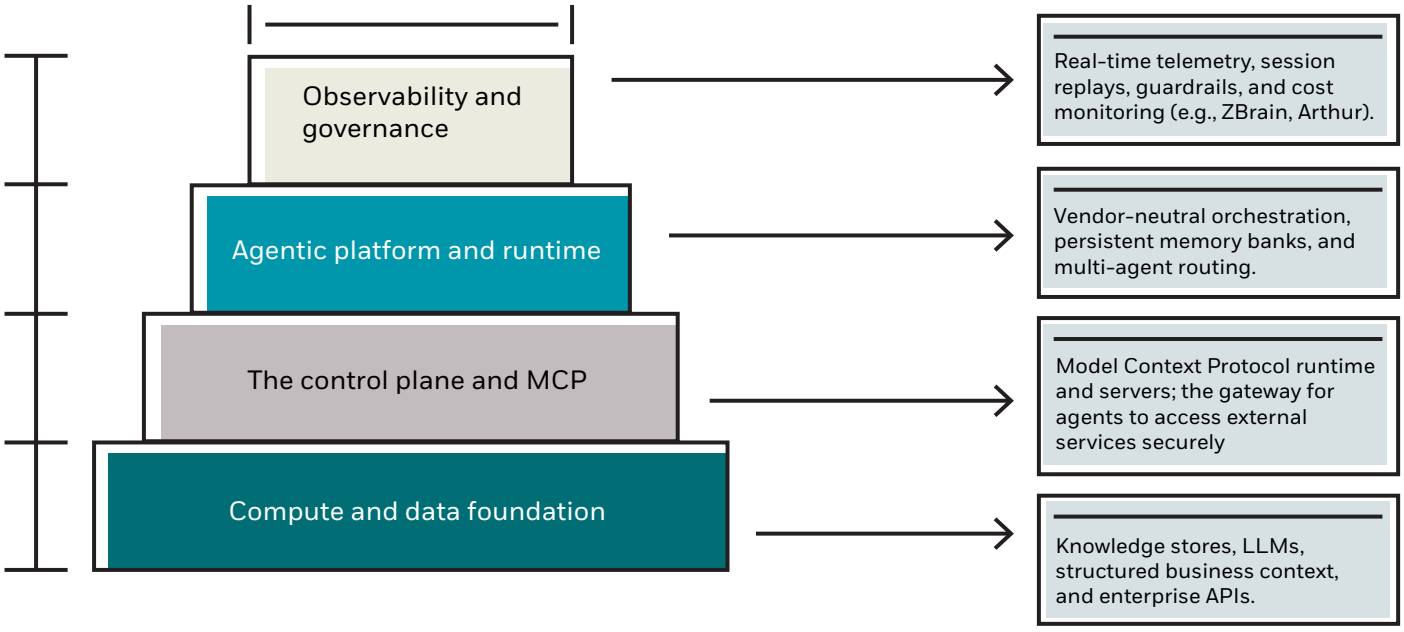
Deploying a fully autonomous agent directly into a live production environment is highly discouraged by all leading advisories. Best practices dictate a phased, empirical rollout strategy designed to minimize risk while accumulating performance data.[12]

The standard deployment pipeline begins with "Shadow Mode." For the first one to three months, the agent is deployed into the production environment where it ingests live data and formulates decisions, but its output is strictly logged alongside actual human actions without taking any real-world action itself. This allows engineering teams to compare the agent's logic against human baselines in a silent, risk-free manner, identifying hallucinations or flaws without business impact.[12]

Once baseline competence is empirically proven in shadow mode, the agent is granted limited autonomous authority over low-risk, small-scale cohorts—a strategy known as a canary release. Performance is continuously measured against tightly defined primary metrics and safety guardrails, utilizing causal inference and A/B testing to determine the actual business lift generated by the agent.[64] Furthermore, these systems must be engineered with strict versioning for all prompts and agent logic, enabling instantaneous rollbacks to prior, stable configurations if anomalous behavior or concept drift is detected during the wider scaling phase.[64]

THE AGENTOPS STACK

Building for production



The observability control plane

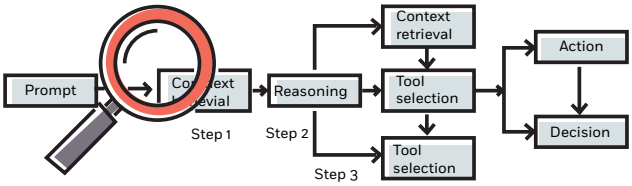
Traditional software monitoring tools—designed primarily to track server uptime, API latency, and standard error logs—are wholly insufficient for managing agentic systems. Because agents generate their own execution paths dynamically, a standard error log cannot flag subtle hallucinations, unintended logical loops, or slow drifts away from the system's intended behavior.[67] Agentic observability requires specialized platforms that serve as the enterprise's strategic control plane, providing continuous, granular transparency into the agent's reasoning chains, tool invocations, and memory utilization.[69]

Without deep observability, autonomous agents can rapidly overspend on API tokens, hallucinate confidently, or execute biased decisions for weeks without detection, creating massive business and regulatory exposure.[69] A mature observability stack serves as an early-warning system, enabling rapid root-cause analysis and disciplined incident response, ensuring that the system remains deterministic and replayable.[69]

OBSERVABILITY

The strategic control plane

Reasoning traceability

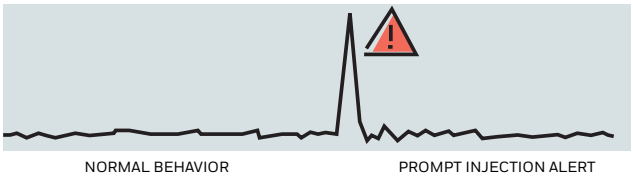


Tracing every step from first prompt to final action to prove why a decision was made.

Tool usage analytics

Tracking which enterprise tools were invoked, success rates, and API expenses.

Anomaly detection



Real-time detection of prompt injection, hallucinations, or unauthorized tool misuse.

Session metrics

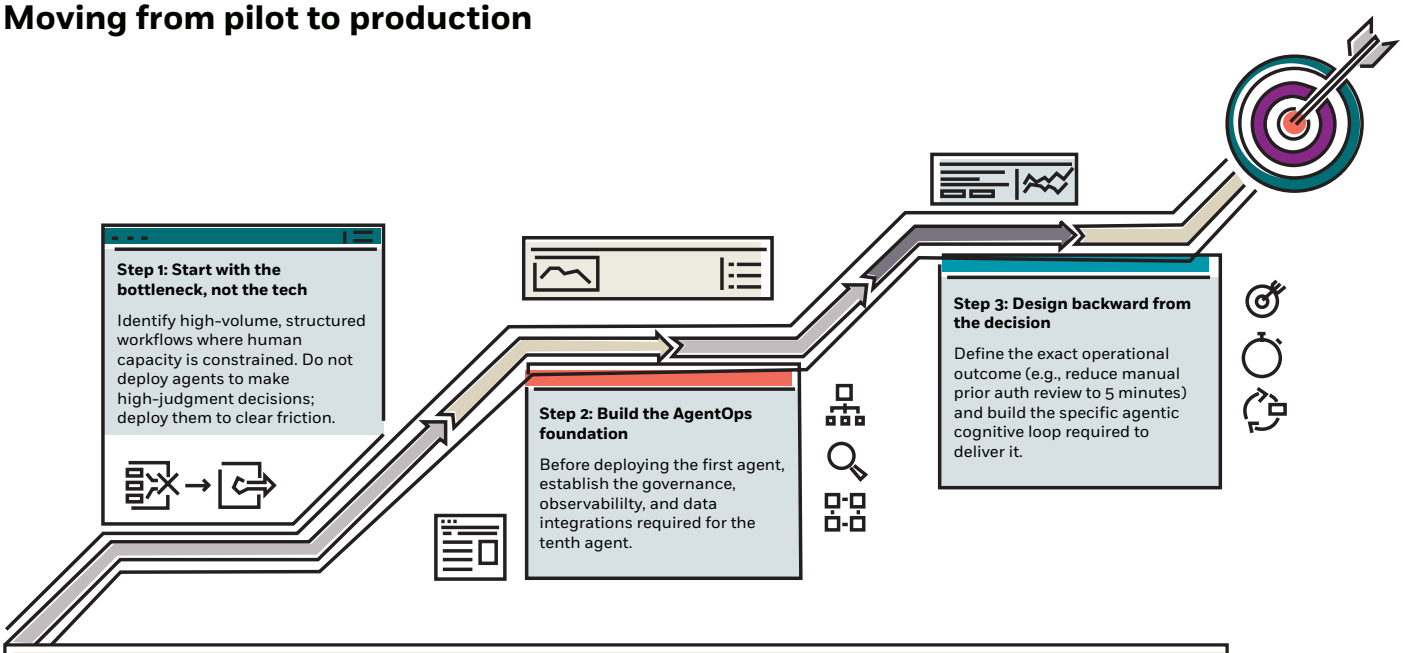
Monitoring latency, token consumption, and overall execution cost per session.

The market offers several highly specialized observability platforms, each optimized for different enterprise architectures:

Specialized observability platform	Primary focus and core strengths	Optimal enterprise use case
LangSmith	Offers comprehensive debugging, evaluations, and deep integration with the LangChain/LangGraph ecosystem. Features structured workflows allowing subject matter experts to annotate production traces easily.[68]	Development teams building complex agent graphs that require domain experts (e.g., legal counsel, clinicians) to review and score specific AI outputs in production.[68]
Arize (Phoenix)	A fully open-source, vendor-agnostic tracing platform with robust evaluation capabilities. Excels at detecting model drift, clustering embeddings, and integrating seamlessly with OpenTelemetry standards.[18]	Large enterprises utilizing mixed machine learning and LLM workloads that require an independent, highly robust infrastructure tracking tool.[71]
Langfuse	A self-hostable, open-source platform that combines strong execution tracing with built-in prompt management and granular cost tracking.[71]	Organizations operating under strict data privacy regulations that require self-hosting, while needing to track cost metrics across diverse prompt iterations.[68]
Arthur AI	Enterprise-grade monitoring focused heavily on translating technical traces into actionable business KPIs, providing compliance auditing, and generating risk-tier alerts.[69]	Highly regulated entities (finance, healthcare payers) that must map autonomous actions directly to strict regulatory frameworks and prove ROI to executive boards.[69]

THE 2026 PLAYBOOK

Moving from pilot to production



Agentic AI is not an IT experiment; it is the new operating model for the enterprise.

Industry discourse: Luminaries, thought leadership, and newsletters

As the agentic landscape evolves at an unprecedented pace, maintaining strategic awareness requires continuous engagement with specialized industry discourse. The theoretical computer science concepts of 2024 are rapidly becoming the strict production standards of 2026. Consequently, enterprise leaders rely heavily on a network of curated newsletters, expert analysts, and thought leaders to separate actionable technological signals from ubiquitous market noise.[74]

Pascal Bornet, a universally recognized authority in intelligent automation, publishes the highly influential Agentic Intelligence newsletter. His work bridges the complex gap between theoretical research and practical business application.[76] Bornet's insights deeply influence how enterprises structure modular multi-agent systems. He strongly advocates for strict clarity of purpose, arguing that each individual agent within an enterprise must function as a highly skilled specialist rather than a generalized, monolithic intelligence.[77] Another prominent voice is Bernard Marr, whose AI & Future Tech Trends newsletter is widely respected for providing clarity and foresight into how these technologies will fundamentally reshape corporate strategy and future workforce dynamics.[76]

For developers, technical architects, and product managers, newsletters such as Latent Space and Interconnects provide rigorous, mathematically grounded deep dives into prompt engineering, execution protocols, and model training methodologies, deliberately stripping away marketing hype in favor of architectural realities.[79] Conversely, publications like TLDR AI, The Rundown AI, and the AI Adopters Club deliver high-velocity, daily synthesis of market movements, specific tool releases, and practical execution case studies. These are explicitly

designed for founders and executives who must scan the horizon efficiently without getting bogged down in code-level specifics.[74]

Within highly regulated vertical markets, generalized artificial intelligence news is insufficient; highly specific policy publications are paramount. In the healthcare sector, the Manatt Health AI Policy Tracker and the IAPP AI Governance Dashboard serve as critical, indispensable resources.[56] These publications meticulously track the labyrinth of state-by-state legislative changes (such as the aforementioned downcoding regulations in Indiana) and federal CMS mandates impacting prior authorization.[56] These targeted insights empower legal counsel and IT compliance teams to design robust systems that anticipate and adapt to regulatory shifts, rather than merely reacting to them after millions of dollars have been invested in deployment.

The trajectory of enterprise software is irrevocably altered. By successfully navigating the complex layers of the technology stack, enforcing uncompromising observability, and prioritizing business value over technological novelty, organizations across retail, finance, and healthcare can safely transition from the limitations of pilot purgatory into the profound efficiency of true autonomous operations.

Connect with a UST expert to dive deeper.

References and further readings

1. One year of agentic AI: Six lessons from the people doing the work - McKinsey, accessed April 22, 2026, https://www.mckinsey.com/capabilities/quantumblack/our-insights/one-year-of-agentic-ai-six-lessons-from-the-people-doing-the-work
2. Gartner Hype Cycle Identifies Top AI Innovations in 2025, accessed April 22, 2026, https://www.gartner.com/en/newsroom/press-releases/2025-08-05-gartner-hype-cycle-identifies-top-ai-innovations-in-2025
3. Predictions 2026: AI Agents, Changing Business Models, And Workplace Culture Impact Enterprise Software - Forrester, accessed April 22, 2026, https://www.forrester.com/blogs/predictions-2026-ai-agents-changing-business-models-and-workplace-culture-impact-enterprise-software/
4. 26 AI Agent Statistics (Adoption + Business Impact) - Datagrid, accessed April 22, 2026, https://datagrid.com/blog/ai-agent-statistics
5. Predictions 2026: The Race To Trust And Value - Forrester, accessed April 22, 2026, https://www.forrester.com/predictions/
6. Enterprise AI Rollout Failures: Causes and Case Studies | IntuitionLabs, accessed April 22, 2026, https://intuitionlabs.ai/articles/enterprise-ai-rollout-failures
7. Enterprise AI has an 80% failure rate. The models aren't the problem. What is? - Reddit, accessed April 22, 2026, https://www.reddit.com/r/AI_Agents/comments/1s02oaq/enterprise_ai_has_an_80_failure_rate_the_models/
8. Why 95% of AI Pilots Fail — and What the Other 5% Do Differently - Salesforce, accessed April 22, 2026, https://www.salesforce.com/news/stories/why-ai-pilots-fail/
9. Three Questions That Will Define AI In 2026 - Forrester, accessed April 22, 2026, https://www.forrester.com/blogs/three-questions-that-will-define-ai-in-2026/
10. The Current State of AI Agents and Agentic AI for HR: Where It's Ready and Where It's Not, accessed April 22, 2026, https://happily.ai/blog/the-current-state-of-ai-agents-and-agentic-ai-for-hr-where-its-ready-and-where-its-not/

11. 5 Defining AI Agent Trends for 2026, accessed April 22, 2026, https://fintechnews.ch/aifintech/5-defining-ai-agent-trends-for-2026/82918/
12. Responsible AI Agents in Cloud Operations: The Critical Role of Observability - Sycomp, accessed April 22, 2026, https://sycomp.com/resource/responsible-ai-agents-cloud-operations-observability/
13. Your AI agents will run everywhere. Is your architecture ready for that? - DataRobot, accessed April 22, 2026, https://www.datarobot.com/blog/agentic-ai-enterprise-deployment-architecture/
14. What Is Agentic AI Architecture? Common Patterns and When to Use Them - Neo4j, accessed April 22, 2026, https://neo4j.com/blog/agentic-ai/agentic-architecture/
15. The 7 Layers of Agentic AI Stack in 2026 - AIMultiple, accessed April 22, 2026, https://aimultiple.com/agentic-ai-stack
16. MCP + A2A: The TCP/IP Moment for AI Agents, accessed April 22, 2026, https://medium.com/@Micheal-Lanham/mcp-a2a-the-tcp-ip-moment-for-ai-agents-bf1927112b07
17. The emerging agentic AI software infrastructure market | Kearney, accessed April 22, 2026, https://www.kearney.com/service/digital-analytics/article/the-emerging-agentic-ai-software-infrastructure-market
18. 15 AI Agent Observability Tools in 2026: AgentOps & Langfuse - AIMultiple, accessed April 22, 2026, https://aimultiple.com/agentic-monitoring
19. Code execution with MCP: building more efficient AI agents - Anthropic, accessed April 22, 2026, https://www.anthropic.com/engineering/code-execution-with-mcp
20. Introducing the Model Context Protocol - Anthropic, accessed April 22, 2026, https://www.anthropic.com/news/model-context-protocol
21. The New Model Context Protocol for AI Agents - Insight Global, accessed April 22, 2026, https://insightglobal.com/blog/new-model-context-protocol-for-ai-agents/
22. Agent Factory: The new era of agentic AI—common use cases and ..., accessed April 22, 2026, https://azure.microsoft.com/en-us/blog/agent-factory-the-new-era-of-agentic-ai-common-use-cases-and-design-patterns/
23. Data architecture for AI agents across your organization - Cloud Adoption Framework, accessed April 22, 2026, https://learn.microsoft.com/en-us/azure/cloud-adoption-framework/ai-agents/data-architecture-plan
24. Vertex AI Agent Builder | Google Cloud, accessed April 22, 2026, https://cloud.google.com/products/agent-builder
25. Building Scalable AI Agents: Design Patterns With Agent Engine On Google Cloud, accessed April 22, 2026, https://cloud.google.com/blog/topics/partners/building-scalable-ai-agents-design-patterns-with-agent-engine-on-google-cloud
26. Google Cloud partners with Tata Steel to deploy over 300 specialized AI agents across operations, accessed April 22, 2026, https://timesofindia.indiatimes.com/technology/tech-news/google-cloud-partners-with-tata-steel-to-deploy-over-300-specialized-ai-agents-across-operations/articleshow/130435276.cms
27. Top 5 Agentic AI Use Cases in Retail - Concentrix, accessed April 22, 2026, https://www.concentrix.com/insights/blog/top-5-agentic-ai-use-cases-in-retail/
28. AI Agents in Retail: Top Use Cases and Examples - Workday Blog, accessed April 22, 2026, https://blog.workday.com/en-us/ai-agents-in-retail-top-use-cases-and-examples.html
29. Agentic AI In Retail: Use Cases, Benefits and Implementation - LatentView, accessed April 22, 2026, https://www.latentview.com/blog/agentic-ai-in-retail/
30. Revolutionizing global supply chains with agentic AI | EY - US, accessed April 22, 2026, https://www.ey.com/en_us/insights/supply-chain/revolutionizing-global-supply-chains-with-agentic-ai

31. Agentic AI in Retail: Real-World Examples and Case Studies - [x]cube LABS, accessed April 22, 2026, https://www.xcubelabs.com/blog/agentic-ai-in-retail-real-world-examples-and-case-studies/
32. 3 Data Trends Shaping the Race to AI Across Industries in 2026, accessed April 22, 2026, https://www.snowflake.com/en/blog/trends-shaping-AI-across-industries-2026/
33. Agentic AI Framework for Smart Inventory Replenishment - arXiv, accessed April 22, 2026, https://arxiv.org/html/2511.23366v1
34. Q&A: Agentic AI in retail and supply chain planning - RELEX Solutions, accessed April 22, 2026, https://www.relexsolutions.com/resources/agentic-ai-in-retail-and-supply-chain-planning/
35. American Express and Hyper: The Dawn of the Autonomous Expense Era, accessed April 22, 2026, https://www.bobsguide.com/american-express-and-hyper-the-dawn-of-the-autonomous-expense-era/
36. Nemotron Labs: How Financial Services Companies Use Agentic AI to Enhance Productivity, Efficiency and Security - NVIDIA Blog, accessed April 22, 2026, https://blogs.nvidia.com/blog/financial-services-agentic-ai/
37. Agentic AI in Financial Services: Choosing the Right Pattern for Multi-Agent Systems - AWS, accessed April 22, 2026, https://aws.amazon.com/blogs/industries/agentic-ai-in-financial-services-choosing-the-right-pattern-for-multi-agent-systems/
38. Top 5 Agentic AI Use Cases in Banking, Finance & Insurance, accessed April 22, 2026, https://www.concentrix.com/insights/blog/top-5-agentic-ai-use-cases-in-banking/
39. AI Agents for Financial Services: Top Use Cases and Examples - Workday Blog, accessed April 22, 2026, https://blog.workday.com/en-us/ai-agents-financial-services-top-use-cases-examples.html
40. AI Agent Orchestration Patterns - Azure Architecture Center | Microsoft Learn, accessed April 22, 2026, https://learn.microsoft.com/en-us/azure/architecture/ai-ml/guide/ai-agent-design-patterns
41. Reference Architecture and automation for Financial Services web applications | Microsoft Azure Blog, accessed April 22, 2026, https://azure.microsoft.com/en-us/blog/reference-architecture-and-automation-for-financial-services-web-applications/
42. Agentic AI in Wealth Management - KPMG International, accessed April 22, 2026, https://kpmg.com/kpmg-us/content/dam/kpmg/pdf/2025/agentic-ai-changing-wealth-mgmt.pdf
43. Agentic AI in wealth management - Capgemini, accessed April 22, 2026, https://www.capgemini.com/us-en/insights/expert-perspectives/agentic-ai-in-wealth-management/
44. Re-thinking Prior Authorization in the Age of Agentic AI - WNS, accessed April 22, 2026, https://www.wns.com/perspectives/articles/re-thinking-prior-authorization-in-the-age-of-agentic-ai
45. Healthcare payors | AI payors healthcare use cases - WRITER, accessed April 22, 2026, https://writer.com/guides/agentic-ai-healthcare-payor-use-cases/
46. The Best Practices of Agentic AI - qBotica, accessed April 22, 2026, https://www.qbotica.com/blog/the-best-practices-of-agentic-ai-best-practices
47. Building Prior Authorization Systems for Healthcare Payors with Amazon Bedrock - YouTube, accessed April 22, 2026, https://www.youtube.com/watch?v=feL4GUiP7Ws
48. Transform healthcare prior authorization with AI Agents | AWS for Industries, accessed April 22, 2026, https://aws.amazon.com/blogs/industries/transform-healthcare-prior-authorization-with-ai-agents/
49. Prior authorization for medical claims using Strands Agents | AWS for Industries, accessed April 22, 2026, https://aws.amazon.com/blogs/industries/prior-authorization-for-medical-claims-using-strands-agents/
50. Fraud, Waste, & Abuse - Shift Technology, accessed April 22, 2026, https://www.shift-technology.com/en-gb/products/fraud-waste-and-abuse

51. AI for Health | Business at OECD, accessed April 22, 2026, https://www.businessatoecd.org/hubfs/AI%20for%20Health.pdf
52. Transforming healthcare enrollment with agentic AI for payors | AWS for Industries, accessed April 22, 2026, https://aws.amazon.com/blogs/industries/transforming-healthcare-enrollment-with-agentic-ai-for-payors/
53. Emids Unveils Healthcare Agentic AI Suite Integrated With Anthropic ..., accessed April 22, 2026, https://lasvegassun.com/news/2026/apr/21/emids-unveils-healthcare-agentic-ai-suite-integrat/
54. Infrastructure Priorities for Agentic AI Success for Healthcare IT Leaders, accessed April 22, 2026, https://www.lumeris.com/in-practice/infrastructure-priorities-for-agentic-ai-success-for-healthcare-it-leaders/
55. How Amazon Connect Health brings agentic AI to the point of care | AWS for Industries, accessed April 22, 2026, https://aws.amazon.com/blogs/industries/how-amazon-connect-health-brings-agentic-ai-to-the-point-of-care/
56. Manatt Health: Health AI Policy Tracker, accessed April 22, 2026, https://www.manatt.com/insights/newsletters/health-highlights/manatt-health-health-ai-policy-tracker
57. Taxonomy of Failure Mode in Agentic AI Systems - Microsoft, accessed April 22, 2026, https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/Taxonomy-of-Failure-Mode-in-Agentic-AI-Systems-Whitepaper.pdf
58. The Biggest AI Fails of 2025: Lessons from Billions in Losses - NineTwoThree Studio, accessed April 22, 2026, https://www.ninetwothree.co/blog/ai-fails
59. State Legislatures Consider Oversight of Artificial Intelligence in Health Insurance Decisions, accessed April 22, 2026, https://www.sheppard.com/insights/blogs/state-legislatures-consider-oversight-of-artificial-intelligence-in-health-insurance-decisions
60. Human-in-the-loop constructs for agentic workflows in healthcare and life sciences - AWS, accessed April 22, 2026, https://aws.amazon.com/blogs/machine-learning/human-in-the-loop-constructs-for-agentic-workflows-in-healthcare-and-life-sciences/
61. Agentic AI for healthcare payers: Smarter health plans - ZS, accessed April 22, 2026, https://www.zs.com/insights/agentic-ai-for-healthcare-payers
62. Agentic AI in Claims Processing: Transforming Insurance Operations through Autonomous AI Systems - Journal of Computational Analysis and Applications (JoCAAA), accessed April 22, 2026, https://eudoxuspress.com/index.php/pub/article/download/4929/3693/10055
63. (PDF) Generative and Agentic Artificial Intelligence for Medical Coding and Billing: A Human-in-the-Loop Architecture and Evaluation - ResearchGate, accessed April 22, 2026, https://www.researchgate.net/publication/400752115_Generative_and_Agentic_Artificial_Intelligence_for_Medical_Coding_and_Billing_A_Human-in-the-Loop_Architecture_and_Evaluation
64. A comprehensive guide to AgentOps: Scope, core practices, key challenges, trends, and ZBrain implementation, accessed April 22, 2026, https://zbrain.ai/agentops/
65. Company as Agentic Workflow - European Nexus for Strategic ..., accessed April 22, 2026, https://www.intelligencestrategy.org/blog-posts/company-as-agentic-workflow
66. Prompt, agent, and model lifecycle management - AWS Prescriptive Guidance, accessed April 22, 2026, https://docs.aws.amazon.com/prescriptive-guidance/latest/agentic-ai-serverless/prompt-agent-and-model.html
67. Taming Uncertainty via Automation: Observing, Analyzing, and Optimizing Agentic AI Systems - arXiv, accessed April 22, 2026, https://arxiv.org/html/2507.11277v1
68. 8 LLM Observability Tools to Monitor & Evaluate AI Agents - LangChain, accessed April 22, 2026, https://www.langchain.com/articles/llm-observability-tools

69. Agentic AI Observability: A 2026 Playbook – Arthur AI, accessed April 22, 2026, https://www.arthur.ai/column/agentic-ai-observability-playbook-2026

70. Hot take: the biggest bottleneck in AI agents right now isn't models, frameworks, or even cost. It's that nobody knows how to properly evaluate if their agent is actually working : r/AI_Agents - Reddit, accessed April 22, 2026, https://www.reddit.com/r/AI_Agents/comments/1srau4n/hot_take_the_biggest_bottleneck_in_ai_agents/

71. Best AI agent observability tools in 2026 | Breyta Blog, accessed April 22, 2026, https://breyta.ai/blog/best-ai-agent-observability-tools

72. Top LLM Evaluation Platforms: In Depth Comparison : r/AI_Agents - Reddit, accessed April 22, 2026, https://www.reddit.com/r/AI_Agents/comments/1pa02zc/top_llm_evaluation_platforms_in_depth_comparison/

73. Best AI Agent Observability Tools in 2026: A Comparison for Production Teams - Latitude.so, accessed April 22, 2026, https://latitude.so/blog/best-ai-agent-observability-tools-2026-comparison

74. 6 AI Newsletters That Save Me Hours Every Week | by Todd Larsen | Medium, accessed April 22, 2026, https://medium.com/@toddlarsen/6-ai-newsletters-that-save-me-hours-every-week-8f8dd2e81ca9

75. The Ultimate List of Top 20 AI Newsletters in 2025 - Test-king.com, accessed April 22, 2026, https://www.test-king.com/blog/the-ultimate-list-of-top-20-ai-newsletters-in-2025/

76. Our Top 4 LinkedIn AI Newsletter Picks - Prompt Security, accessed April 22, 2026, https://prompt.security/blog/shaping-the-conversation-our-top-4-linkedin-newsletter-picks-on-ai

77. Understanding AI Agents: Evolution & Impact | PDF | Artificial Intelligence - Scribd, accessed April 22, 2026, https://www.scribd.com/document/943382160/Agentic-Artificial-Intelligence

78. Agentic Artificial Intelligence: Harnessing AI Agents to Reinvent Business, Work, and Life, accessed April 22, 2026, https://www.researchgate.net/publication/389845606_Agentic_Artificial_Intelligence_Harnessing_AI_Agents_to_Reinvent_Business_Work_and_Life

79. Any good AI / AI Agents newsletters you recommend? : r/AI_Agents - Reddit, accessed April 22, 2026, https://www.reddit.com/r/AI_Agents/comments/1q3i6cq/any_good_ai_ai_agents_newsletters_you_recommend/

80. My Top 5 AI Newsletters for 2026 - YouTube, accessed April 22, 2026, https://www.youtube.com/watch?v=i9gwyvoi3Pk

Since 1999, UST has worked side by side with the world's best companies to make a powerful impact through transformation. Powered by technology, inspired by people, and led by our purpose, we partner with our clients from design to operation. Our digital solutions, proprietary platforms, engineering, R&D, products, and innovation ecosystem turn core challenges into impactful, disruptive solutions. With deep industry knowledge and a future-ready mindset, we infuse expertise, innovation, and agility into our clients' organizations—delivering measurable value and positive lasting change for them, their customers, and communities around the world. Together, with 30,000+ employees in 30+ countries, we build for boundless impact—touching billions of lives in the process.

ust.com

Version 0101-20260531