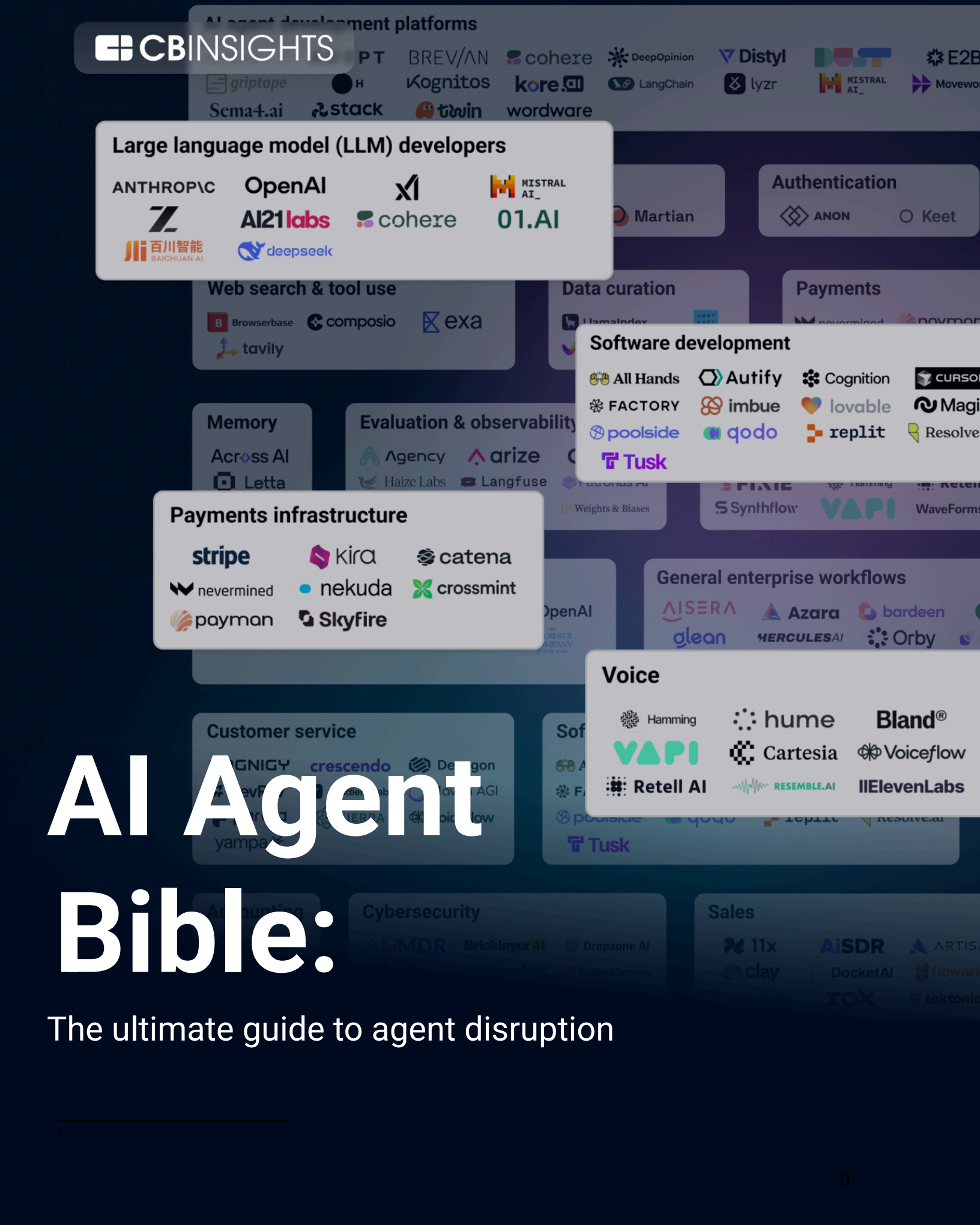
CBINSIGHTS

AI Agent Bible:

The ultimate guide to agent disruption



The AI agent landscape is evolving rapidly, with 500+ startups founded since 2023. Here is our top AI agent research — featuring the trends, key players, and 6 predictions to watch all based on CB Insights predictive intelligence.

AI agents are defining the next wave of tech innovation.

Every big tech company and a rapidly growing private market landscape are building agent offerings targeting horizontal use cases and industries from financial services to manufacturing.

For enterprises across sectors, one question is becoming unavoidable: **Which AI agent strategies will separate market leaders from those left behind?**

Enterprises are under pressure to build and implement agents as these LLM-based systems change how companies operate, hire, and scale.

Across 9 reports, discover where startup innovation is pointing, promising partnership and acquisition targets, and key trends to watch based on CB Insights predictive intelligence.

Foreword

Manlio Carrelli
CEO of CB Insights

Seven years ago, I watched companies struggle with primitive chatbots that broke the moment customers went off-script.

Today, those early systems have given way to **AI agents, which in 2 short years have moved from experiment to enterprise priority**. I'm seeing agent mentions on earnings calls 10x since 2023. The velocity is unlike anything we've tracked before.

What strikes me most: agents are climbing the value chain faster than any technology I've seen. They gained traction in customer service – which made sense, given the sector's early AI adoption. 82% of organizations we surveyed in June 2025 said they will be using AI agents in customer support in the next 12 months. But now they're supporting clinical decisions at hospitals, assessing financial risk at banks, and drafting legal memos at law firms.

This is why 25% of AI agent startups founded since 2023 are already deploying, achieving levels of commercial distribution that previously took startups 5+ years, according to our Commercial Maturity scores.

The momentum is accelerating. Among the 1,500+ tech markets CB Insights tracks, five of the top 10 by deal activity in 2025 are agent-related. One in five new unicorns is building agent technology. Half of the top 20 revenue-generating agent startups didn't even exist three years ago.

But here's what I keep telling our enterprise clients: **agents are only as good as the data they run on**. The real competitive advantage isn't in the AI – it's in operationalizing accurate, comprehensive data and the orchestration layers that make agents actually useful.

That's where the roadblocks are. Enterprises understand agents matter, but they're struggling with integration complexity, security concerns, and vendor sprawl. The opportunity is massive for those who can navigate this transition.

The research and data that follows breaks down this rapidly evolving landscape by vertical, maturity, and momentum. Remember: in the age of AI agents, the edge doesn't live in the model. It lives in connecting purpose-built agents to purpose-specific data.

Those who get this right will shape the economy agents are creating.

Predictive Intelligence on Private Companies

With CB Insights, you see your next move – first. Rare data straight from companies building the future. Powerful AI that automates the analysis.

Start free trial

Trusted by the world's smartest companies

Microsoft 3M WELLS FARGO jack henry IBM citi Boehringer Ingelheim

GDPR and SOC2 Compliant

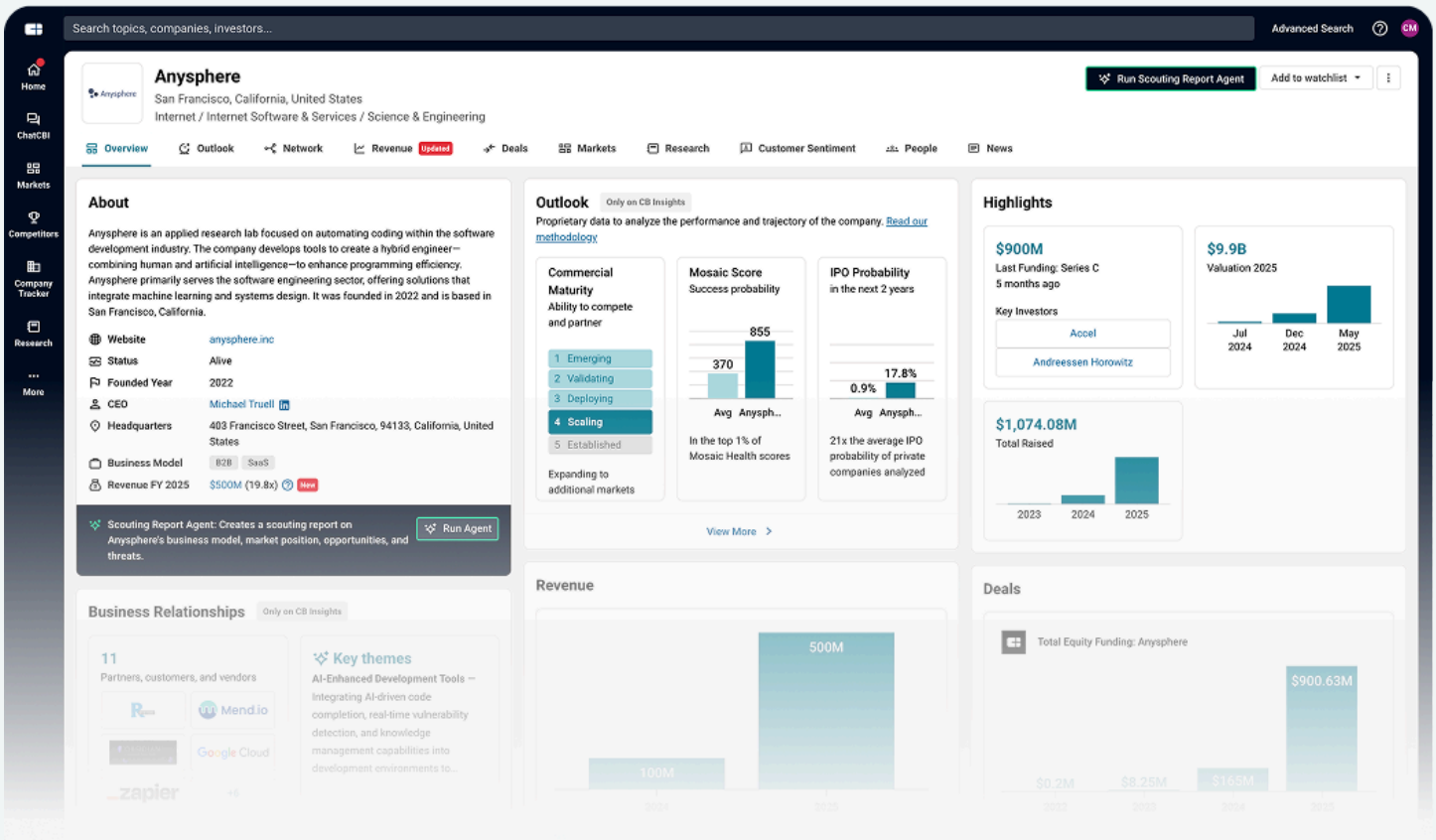


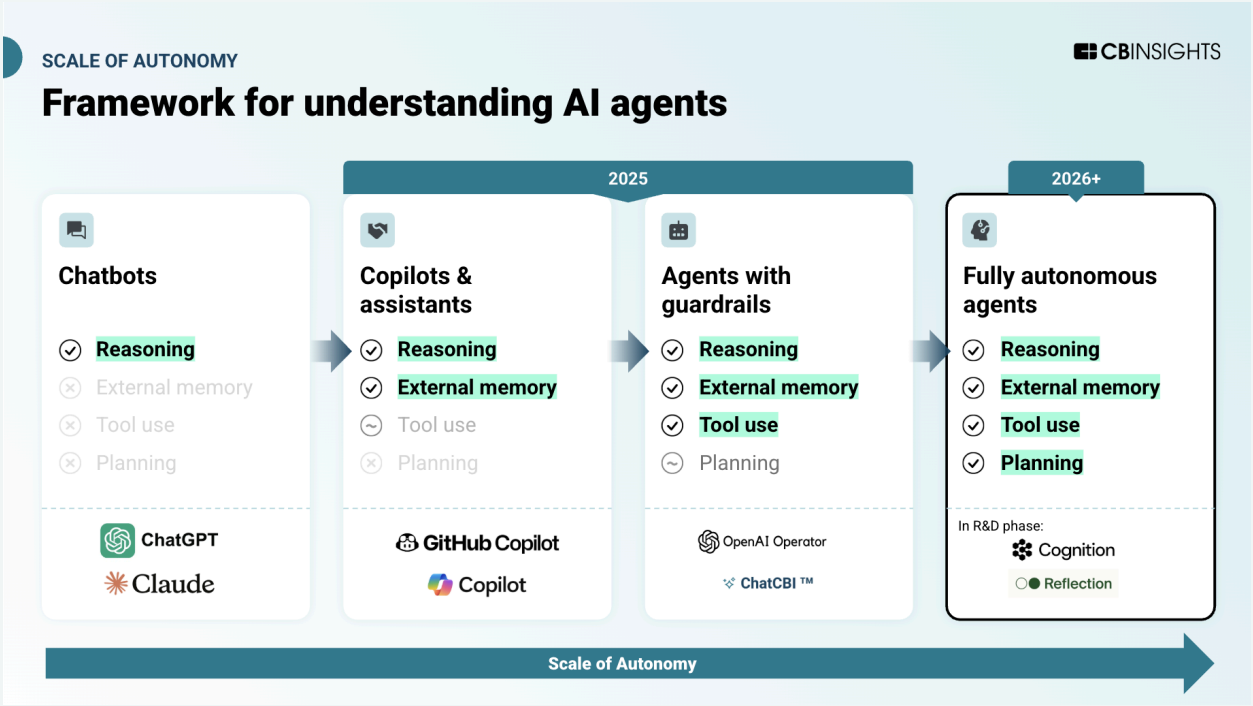
Table of Contents

Outlook on AI agents

2025: Agents with guardrails

An AI agent is an LLM-based system designed to independently perform tasks on a user's behalf by reasoning, planning, leveraging memory, and interacting with external tools and other agents.

Most agents today operate within constrained environments, utilizing structured workflows and "guardrails" to complete specific objectives while retaining some decision-making control.



As foundation model capabilities improve, agents are expected to become increasingly autonomous.

Beyond 2025: Fully autonomous agents

Agents will operate without human involvement, with more sophisticated decision-making, adaptability, and task execution.

Looking into 2026, watch for new form factors outside of the copilot/chatbot interface that will push the boundaries of what an "agent" is.

Early indications of this include “AI-native” workspaces — tools and platforms built from the ground up around AI capabilities, rather than layering AI features on top of a traditional product. Other examples include web browsing interfaces and voice-driven applications.

“

AI agents are nowhere near their final form. What we mean is this: today we’re in a “copilot” era where we get outputs via AI chat. But agents are already transforming into something more like a superpowered instrument or tool. Tools shape the work we do, so an AI supertool will allow our work to go in all kinds of previously impossible directions.

In enterprise, the copilot model has helped in areas where knowledge work is both routine and high-stakes, like support, training, onboarding, and compliance. **But the next profound shift will be for agents to cease waiting for input, instead acting as proactive collaborators and extensions of our minds. Interacting with an agent will not just help solve one problem; it will actually shape how an organization functions and reacts to change.**

Over time, enterprises that embrace AI supertools as core learning systems will find their cultures and capabilities transformed in ways that go far beyond efficiency gains; they will unlock entirely new ideas, solutions and ways of working.

Michael Mignano
Partner at Lightspeed

6 AI agent predictions looking into 2026

1) Voice AI sprints ahead

Early-stage genAI companies with the fastest-growing headcounts are concentrated in AI agent applications — and more specifically in voice AI development.

Companies are now positioning themselves for a future where humans interact with AI via conversation rather than text interfaces. For agentic AI development, implications are significant across customer service, sales, and IT support where voice agents can handle complex conversations without human intervention.

AI agents and voice applications sprint ahead

Top early-stage generative AI startups based on 12-month headcount growth

Company	Category	12 month headcount change	Mosaic (overall)
Sapien	Data labeling	+1,150%	796
wordware	AI agent development	+975%	843
Lovable	Coding AI agents	+907%	865
NURIX	**Voice AI agents**	**+611%**	**752**
etched	AI chips	+479%	742
bolo.ai	Industrial AI agents	+440%	702
VAPI	**Voice AI development**	**+427%**	**846**
rhino.ai	System modernization	+425%	834
sintra	Enterprise AI agents	+343%	N/A
Retell AI	**Voice AI development**	**+340%**	**692**
CodeRabbit	AI code review	+340%	797
HappyRobot	Supply chain AI agents	+300%	750
Dnotitia	Model development	+286%	829
Wordsmith	Legal AI	+286%	803
Synthflow	**Voice AI development**	**+267%**	**800**

State of AI Q2'25
Note: Analysis features companies with a minimum 25 headcount. As of 7/1/2025

CBINSIGHTS

Watch voice AI development startups for partnership, investment, and acquisition opportunities. Signaling the potential for increasing consolidation, Meta acquired voice AI startups Play AI and WaveForms AI in 2025.

2) Wave of AI M&A comes for agents

Agentic solutions led the top AI exits to kick off the year in Q1'25. Moveworks, Weights & Biases, and OfferFit secured the 3 largest deals among 85 acquisitions — establishing agents as the primary focus of industry consolidation.

Overall, there have been 35+ acquisitions in the broader AI agent & copilot space in 2025 so far. In the last quarter, that includes Orby AI (acquired by Uniphore), Windsurf (acquired by Cognition), Cognigy (acquired by NICE), and Flowise and Sana Labs (acquired by Workday).

Enterprise buyers are increasingly seeking to build comprehensive agent solutions to gain a competitive edge.

Two further areas of consolidation we predict:

1. Sales and marketing AI agents offer low-hanging fruit for SaaS leaders' agent plans.
2. The coding AI agent & copilots market is ripe for consolidation amid explosive growth, soaring valuations, and a fractured landscape. Partnerships with major tech and AI players set apart top coding M&A targets.

3) The margin squeeze hits beyond coding

AI agent startups across verticals will face the same economic pressures that reshaped coding AI margins in 2025. Reasoning models, which sparked the rise of "vibe coding" (giving a high-level goal and delegating multi-step implementation to the AI), also translated into higher compute costs as they inflated output-token volume roughly 20x.

Expect other agent categories with growing usage to rework pricing and contracts as reasoning costs mount. For example, in May 2025, Salesforce's Agentforce shifted prices from $2 per conversation to a hybrid-usage Flex Credits system, tying credits to necessary actions for an outcome.

As margins compress, struggling players will seek exits — see for example Cognition's acquisition of Windsurf after the startup's licensing deal with Google.

4) The foundations for agentic commerce solidify

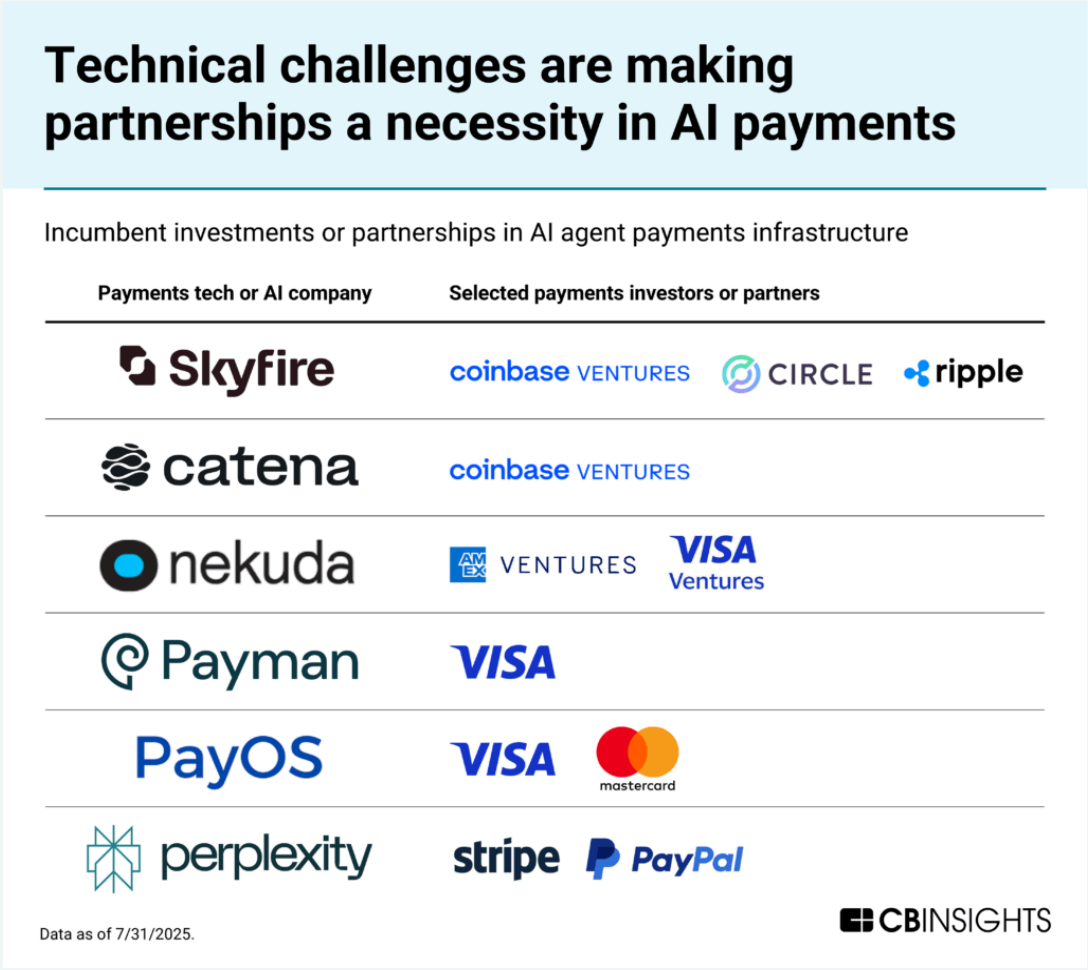
One of the biggest barriers to fully autonomous shopping is enabling secure, real-time transactions. A new class of startups is tackling this challenge head-on, building AI-native payment rails and digital wallets that let users authorize — and limit — spending by AI agents.

Overall, the AI agent payments infrastructure market is one of the more nascent markets in the AI agent tech stack, with an average Commercial Maturity of 2.4 (validating) and an average Mosaic score of 697.

Mosaic score

Evaluates the overall health and growth potential of private companies based on performance, financial stability, market conditions, and management strength. It combines these factors into a single score (out of 1,000).

In an indication of the tech's potential and the technical hurdles that must be overcome, established card and payment networks are investing and partnering with startups in the market (see graphic).



Stripe meanwhile announced in September 2025 it's launching an API for agentic payments. The $107B payments giant also launched with OpenAI the Agentic Commerce Protocol, which will eventually enable AI agents to complete purchases on behalf of users by providing a standardized communication framework between buyers, AI agents, and businesses.

As fintech giants, AI startups, and commerce platforms converge on the agentic payments challenge, this early collaboration will shape how future transactions will happen: securely, autonomously, and inside the AI interfaces consumers already trust.

4) "Data moat wars" reshape enterprise software

Software incumbents are walling off access to their customer data as AI agents become more capable. Salesforce's new Slack API rate limits in 2025, which prevent external apps from accessing chat data in bulk or storing it long-term, are a clear example. Atlassian similarly announced additional rate limiting for its Confluence and Jira APIs in 2025.

This creates problems for AI startups that depend on accessing data across multiple systems of record to automate workflows (e.g., knowledge management platforms like Glean). But it also creates friction for enterprises who want to integrate data across their apps.

A counter movement is emerging. In September 2025, Snowflake launched a consortium with over a dozen providers (including Salesforce) committing to standardized data formats that enable AI to access information across multiple applications. While representing a small fraction of the market, it signals enterprise demand for openness over vendor lock-in.

This tug-of-war will play out in a few ways. We expect incumbents will continue restricting API access to protect their AI product revenues. Enterprises that want more control over their data will push back, driving demand for solutions that help them own their data infrastructure directly. Some startups will go "up the stack" to build agents that work despite limited API access (e.g., via computer use), while others will move closer toward where data is stored and processed.

5) Agent monitoring tools emerge as an essential enterprise category

AI agent reliability remains a major challenge in the landscape. Agents that fail, hallucinate, or behave unpredictably create immediate operational problems.

This is driving activity in oversight capabilities to manage agent risk. Analyzing the 7 early-stage deals ($30.9M total funding) in the AI agent observability, evaluation, & governance market so far this year highlights emerging technical needs:

- **Voice agent testing:** Cekura ($2.4M seed) and Coval ($3.3M seed) both focus on voice AI testing and simulation, signaling that conversational agents require distinct observability approaches.
- **Synthetic user generation:** Multiple companies (Cekura, Coval, Janus) focus on synthetic user generation for agent testing.
- **AI productivity measurement:** Larridin's $17M seed round led by Andreessen Horowitz shows enterprise demand for quantifying AI agent ROI, with a focus on measuring human-AI workforce productivity.

2025 early-stage deals to AI agent observability, evaluation, and governance companies

Companies	Latest round	Deal size	Deal date	Round investors
Cekura	Seed VC - II	$2.4M	7/1/2025	Alan Rutledge, Allen Wu, Allport Capital
Traceloop	Seed VC	$6.1M	5/27/2025	Grand Ventures, Ibex Investors, Samsung NE
Larridin	Seed VC	$17.0M	4/2/2025	Andreessen Horowitz, Bloomberg
Janus	Convertible note	$0.5M	4/1/2025	Y Combinator
LangWatch	Pre-Seed - II	$1.1M	2/25/2025	Antler, Passion Capital, Volta Ventures
Coval	Seed VC - II	$3.3M	1/23/2025	Fortitude Ventures, General Catalyst
Lucidic AI	Convertible note	$0.5M	1/1/2025	Y Combinator

Source: CB Insights Business Graph. Data as of 8/26/2025.

Companies deploying agents at scale will need proper oversight tools to avoid costly failures and maintain operational control.

> **As security agents become more prevalent, the next generation of cybersecurity professionals will evolve into "generals" managing armies of AI agents and "special forces" handling tier-3 incident response and large cross-functional architecture changes.** To support these roles and address explainability challenges in critical security decisions, most AI security agents leverage LLMs not only to perform tasks but also to convert complex actions and enrichment into easy-to-understand reports with citations on key information sources.
>
> **Edward Wu**
> Founder & CEO, Dropzone AI

The AI agent ecosystem

Mapping the landscape of startups, infrastructure providers, and emerging revenue leaders reshaping the market

The AI agent market map

We mined the CB Insights database to map 170+ AI agent startups across 26 categories. We also provide an outlook on AI agents' progress, limitations, and future directions.

Published 3/6/2025 — Link to report

KEY TAKEAWAYS

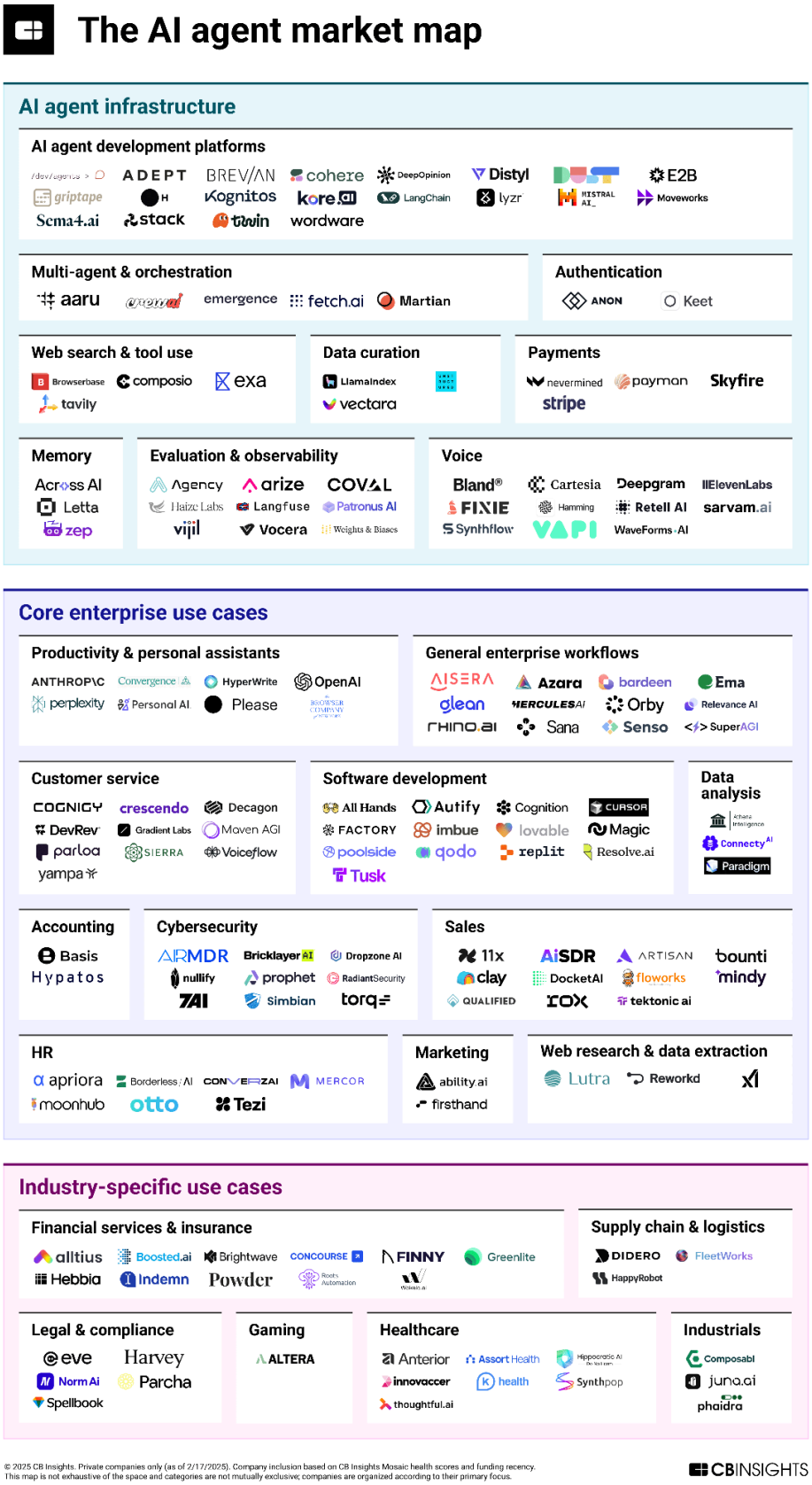
- **AI agent funding is exploding as "digital coworkers" become reality:** AI agent startups raised $3.8B in 2024 (nearly tripling 2023's total), with every major tech player now developing agents or agent tooling. The shift from AI copilots to autonomous agents with greater decision-making scope is happening quickly, promising to alter workforce composition and maximize operational efficiency through task automation.
- **Customer service and software development lead commercial adoption:** These two horizontal markets show the strongest traction based on median Mosaic health scores (714 and 737 respectively), with two-thirds of surveyed organizations already using or planning to use AI agents in customer support within 12 months. These markets benefit from well-defined workflows and testable environments that make agent deployment more straightforward.
- **Trust remains the biggest barrier to full autonomy:** Despite rapid growth, fully autonomous agents remain limited due to reliability, reasoning, and access issues. Startups are addressing this through 5 primary trust-building methods: transparency, human oversight, technical safeguards, security & compliance, and continuous improvement. Most current applications still operate with "guardrails" within constrained architectures rather than true autonomy

We selected companies for inclusion based on Mosaic health scores (500+) and/or funding recency (since 2022). We included private companies only and organized them according to their primary focus. This market map is not exhaustive of the space.

Want to be considered for future AI agent research? Brief our analysts to ensure we have the most up-to-date data on your company.

CBI customers: Track 250+ AI agent companies

The AI agent market map



CBINSIGHTS

AI agent infrastructure

This segment covers companies building agent-specific infrastructure. (We excluded general genAI infrastructure markets like foundation models and vector databases from the map.)

Development tools

A diverse ecosystem of tools has emerged to support agents' development. These range from memory frameworks like Letta that enable persistent, retrievable memory across interactions; to tools that allow agents to take action via integration (e.g., Composio), authentication (e.g., Anon), and browser automation (e.g., Browserbase).

Another set of companies is giving agents more utility across payments (which includes companies developing crypto wallets for agents as well as virtual cards) and voice (development platforms and tools for testing AI voice applications as well as speech models).

Meanwhile, demand for simplified, comprehensive deployment options is driving the rise of AI agent development platforms — the most crowded infrastructure market on our map.

LLM developers including Cohere (with its North AI workspace) and Mistral have launched their own agent development frameworks, while Amazon, Microsoft, Google, and Nvidia all offer AI agent development tooling. With many enterprises favoring established vendors due to lower risk, big tech companies have significant advantages here.

Trust & performance

Concerns around reliability and security have helped establish a market for agent evaluation & observability tools. Early-stage companies are targeting applications such as automated testing (e.g., Haize Labs) and performance tracking (e.g., Langfuse).

Multi-agent systems, where specialized sub-agents work together to complete tasks, also show promise in improving accuracy. Insight Partners-backed CrewAI's multi-agent orchestration platform is reportedly already used by 40% of the Fortune 500.

> **Multi-agent orchestration will evolve into an OS for enterprises — governing thousands of agents with rules for autonomy, oversight, and auditability.** Pre-built agents may help early adopters, but long-term, enterprises will favor platforms that let them compose and govern custom agents at scale.
>
> **João Moura**
> CEO, CrewAI

Vendors are also tackling reliability concerns directly. Based on our briefings with 20+ AI agent startups in Q1'25, companies are using 5 primary methods to build user trust:

1. Transparency
2. Human oversight
3. Technical safeguards
4. Security & compliance
5. Continuous improvement

How are AI agent startups building trust?

Based on 20+ briefings with AI agent startups in Q1'25

Method	Description
Transparency	Showing plans, reasoning, and real-time progress; allowing inspection of generated code
Human oversight	Configurable approval workflows; human-in-the-loop for critical decisions; ability to override recommendations
Technical safeguards	Multi-model consensus mechanisms; constraint layers; protection against hallucinations
Security & compliance	Data encryption; role-based access; adherence to standards (SOC 2, ISO 27001, HIPAA, GDPR)
Continuous improvement	Learning from feedback; performance measurement against business KPIs; gradual autonomy as trust builds

Source: CB Insights Analyst Briefings (as of 2/28/2025).

CBINSIGHTS

Horizontal applications & job functions

Horizontal AI agent startups make up nearly half of the map and overall landscape.

This segment primarily features startups targeting enterprises, with industry-agnostic applications across job functions like HR/recruiting, marketing, and security operations. Companies in the productivity & personal assistants market, including OpenAI with its Operator agent, are targeting consumers and employees directly.

The AI agent markets with the most traction — based on companies' median Mosaic health scores — are customer service and software development (which includes coding and code review & testing agents). These markets are also among the most crowded due to the value agents bring to well-defined workflows and testable environments.

Promising horizontal AI agent markets

Software development and customer support show strong potential

Based on median Mosaic score (a measure of private company health) by category*

Category	Median Mosaic score (out of 1,000)	Number of companies
Software development	737	18
Customer service	714	18
Productivity & personal assistants	687	12
Sales	636	18
HR	619	10
General enterprise workflows	618	22
Cybersecurity	592	12
Web research & data extraction	512	7

Source: CB Insights. Note: Medians based on Mosaic scores (as of 2/19/2025) for private AI agent companies.
*Excludes categories that had under 5 AI agent startups.

CBINSIGHTS

We see this reflected in adoption, particularly at the customer service layer: Among 64 organizations surveyed by CB Insights in December 2024, two-thirds indicated they are using or will be using AI agents in customer support in the next 12 months.

Overall, horizontal AI agent applications are more commercially mature compared to the infrastructure and vertical segments, with over two-thirds of the market deploying or scaling their solutions based on CBI Commercial Maturity scores.

Vertical (industry-specific) applications

We expect increasing verticalization as startups carve out niches by solving industry-specific customer problems, especially in areas with strict regulatory scrutiny and data sensitivity.

This category features companies catering to industries including:

- **Financial services & insurance:** The most crowded vertical category on the map with 11 companies, startups here are targeting a variety of finserv workflows such as financial research (Boosted.ai and Wokelo), insurance sales & support (Alltius and Indemn), and wealth advisory prospecting & operations (Finny AI and Powder).
- **Healthcare:** Solutions in this market aim to reduce the volume of manual tasks for healthcare professionals across use cases like clinical documentation, revenue cycle operations, call centers, and virtual triage. Solutions from companies like Thoughtful AI (revenue cycle operations) and Hippocratic AI (staffing marketplace) are targeting end-to-end healthcare workflows.
- **Industrials:** These companies look to optimize processes and equipment — including control systems, robots, and other industrial machines — without relying on consistent human intervention. For example, Composabl launched an agent platform in May 2024 that uses LLMs to create skills and goals for agents that can control industrial equipment. Public companies like Palantir are also active in this space. Learn more in our industrial AI agents & copilots market map.

> “
>
> The promise of autonomous agents cannot be fulfilled without solving context engineering. **Agents need the right context at the right time and in the right format to make reliable decisions.** Enterprises also require provenance and governance: knowing where information originated, how it evolved, and ensuring compliance across sensitive domains. Current approaches to agent memory still treat each user interaction in isolation, with little connection to other touchpoints such as email, support cases, CRM data, or SaaS usage. **Enterprises will need context systems that unify these signals, track evolving relationships and business states, and maintain controls and audit trails that make agents trustworthy at scale.**
>
> **Daniel Chalef**
> Founder and CEO, Zep AI

The AI agent tech stack

We mined the CB Insights Business Graph to map the 135+ startups across 17 markets forming the new AI agent tech stack. Discover which markets are gaining momentum, identify investment opportunities, and find partners for building AI agents.

Published on 8/22/2025 — Link to report

KEY TAKEAWAYS

- **Voice AI emerges as a new battleground** with an average Mosaic score of 756 and $400M raised in 2025, while Meta acquired its first companies since 2022 (PlayAI and WaveForms AI) in this space.
- **Tech giants race to own agent communication standards** with Anthropic's MCP, Google's A2A Protocol, and IBM's Agent Communication Protocol all launching within a year. Professional services firms and cloud providers are rapidly adopting these protocols, making them table stakes across the value chain.
- **Critical white space opportunities emerge across monetization, marketplaces, and cost management** as the AI agent stack matures

In under a year, the AI agent landscape has grown from roughly 300 players to thousands. Agents are making their way into workflows across verticals, from e-commerce to industrials.

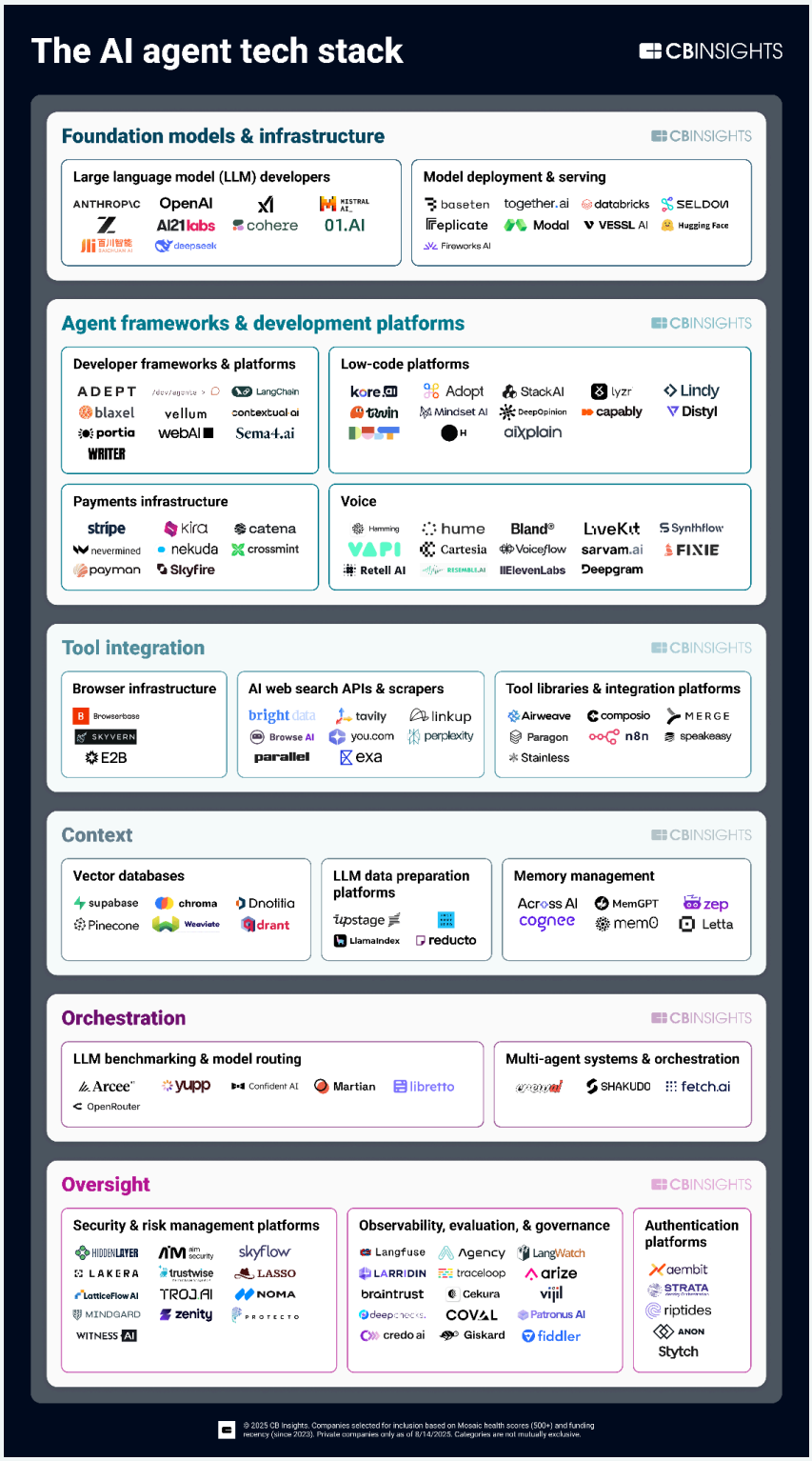
Underpinning this momentum is an emerging tech stack. Infrastructure layers — from foundation models to oversight — are helping enterprises build, deploy, and manage AI agents more effectively.

Using the CB Insights Business Graph and proprietary signals, we mapped 135+ promising private companies building infrastructure for AI agents.

Below the map, we outline the emerging markets and trends investors and strategy leaders should be watching.

CBI customers: Track 250+ companies building AI agent infrastructure

We selected companies for inclusion based on Mosaic health scores (500+) and funding recency (since 2023). Includes private companies only, organized according to their primary focus. Excludes general enterprise workflow automation platforms and non-pure-play LLM developers. This market map is not exhaustive of the space.



Outlook & key takeaways

Private market momentum points to payments, voice, and security as key markets to watch

The AI agent tech stack is a high-momentum landscape, based on CB Insights Mosaic startup health scores. Private companies across the markets outlined below have an average Mosaic score of 768 — more than double the average of 370 for all private companies. They also have an average Commercial Maturity Score of 3, indicating widespread solution deployment.

A deeper dive into these scores, partnerships, and funding reveals 3 emerging markets to watch:

- **Voice AI is the new battleground for the next wave of AI agents**: With an average Mosaic score of 756 and nearly $400M in funding in 2025 so far, voice AI development platforms are building momentum. Big tech also recognizes voice as an essential AI building block — Meta's first acquisitions since 2022 this year were PlayAI and WaveForms AI, both operating in audio and voice AI.
- **AI agent security startups see rapid momentum growth:** AI agents create new attack surfaces and data breach risks, driving urgency for agent security startups. Companies in the market averaged a 56-point Mosaic score growth over 12 months, with Zenity, WitnessAI, and TrojAI each gaining 100+ points. The companies with the highest jumps in Mosaic score are partnering with larger tech firms and cybersecurity leaders. Public and established companies have also entered the conversation, with identity leader Okta and cybersecurity giant Palo Alto Networks both building agent security into their platforms.
- **AI agent payments startups get backing from incumbents:** Agent payments infrastructure is one of the more nascent markets in this tech stack, with an average Commercial Maturity of 2.4 (validating) and an average Mosaic score of 697. The barrier to entry in payments is high, requiring complex technical and regulatory infrastructure. In an indication of the tech's potential, established card and payment networks are investing and partnering with startups in the market: Coinbase backed Skyfire and Catena, Visa invested in Payman, and American Express participated in Nekuda's recent seed round. Others like Crossmint and pre-funding PayOS have partnered with Visa and Mastercard.

Major LLM providers and tech incumbents all try to own a piece of the open standards pie

The growth of AI agents and development platforms has created a need to facilitate communication between agents and access to context. LLM developers and major tech companies are competing to own these standards.

In less than a year:

- Anthropic launched Model Context Protocol (MCP), standardizing how AI agents connect to external tools and data sources
- Google created the Agent-to-Agent (A2A) Protocol that allows agents to collaborate with each other, regardless of underlying framework
- IBM introduced Agent Communication Protocol, which enables inter-agent communication across technologies and systems within a local environment

These protocols have quickly become table stakes across the AI agent value chain. Professional services firms like Accenture, McKinsey, Deloitte, and KPMG contributed to Google's A2A, and big tech companies like Microsoft and AWS support MCP. Meanwhile, startups in the tool libraries & integrations platform market like Speakeasy and Stainless are helping companies build MCP-compatible interfaces for their APIs (known as MCP servers), enabling AI agents to interact with their services.

Big tech pushes deeper into AI agent development

While the above market map highlights the private landscape, tech giants and incumbents are also active across the AI agent infrastructure landscape. The top 3 global cloud providers — Amazon, Microsoft, and Google — are expanding their AI agent offerings across development tooling, hosting, orchestration, and more.

The cloud titans' AI agent playbook

Top cloud providers expand their AI agent offerings

Strategic focus ■ High ■ Medium ■ Low

	amazon	Alphabet	Microsoft
Summary	Amazon focuses on enterprise integration through Bedrock Agents and custom orchestrators, while expanding into consumer-facing agentic AI	Alphabet (Google) leverages its proprietary Gemini models to attract developers with comprehensive building tools and a marketplace strategy	Microsoft's offering combines pre-built business agents and robust development tools, creating an end-to-end ecosystem that addresses immediate business needs
Product offering			
Proprietary foundational model for agents	Amazon Nova	Gemini	Magma
Direct agent offerings	Nova Act agent Amazon Q Developer CLI	Deep Research Idea Generation Project Mariner Gemini Code Assist	Pre-built agents* Autonomous agents in Dynamics 365 GitHub Copilot
Agent development tools	Amazon Bedrock Agents	Vertex AI Agent Builder Agent Designer in Agentspace Agent Development Kit	Microsoft Copilot Studio Azure AI Agent Service
Dedicated agent marketplaces	AWS Agent Marketplace	AI Agent Space	Microsoft Agent Store
Agent orchestration	Amazon Bedrock Agents custom orchestrator	Agent Development Kit	AutoGen Magnetic-One (built on AutoGen)
Agent interoperability	Adopted Google's A2A protocol	Agent2Agent (A2A) protocol	Adopted Google's A2A protocol

*Analyst, Researcher, Security Copilot

Data as of 5/7/2025. Note: Strategic focus based on recency and depth of offerings, complemented by CB Insights' data such as number of partnerships and earnings transcripts insights

CBINSIGHTS

With many enterprises favoring established vendors, big tech companies have significant advantages in AI agent development. Similarly, enterprise software incumbents like Salesforce (Agentforce) and ServiceNow (AI Agent Marketplace) have launched agent platforms and marketplaces targeting their installed bases.

Yet startups across the stack are carving out defensible positions by solving specific technical challenges and pushing the boundaries of what agents can do across areas like multi-agent orchestration (CrewAI) and enterprise data preparation (LlamaIndex). In the crowded AI agent development market, end-to-end platforms like WRITER and Dust are differentiating with vertical-specific implementations and promising speedy deployments.

> “
>
> **High-value agents deeply understand the nuances of a business and its compliance requirements.** They can access relevant contextual information about their environment such as user identities, business rules, knowledge graphs, and historical data. They can use different reasoning patterns to make decisions and adapt their approach when problems change. And they can remember their decisions across sessions and improve based on feedback. Building these agents on a platform like WRITER, **where you can set guardrails for agent activities and control the systems they have access to, is how businesses achieve next-level scale and ROI.**
>
> 
>
> **May Habib**
> CEO, WRITER

Autonomous agents drive the need for an oversight layer

AI agent reliability remains a major challenge in the landscape. Agents that fail, hallucinate, or behave unpredictably create immediate business risk.

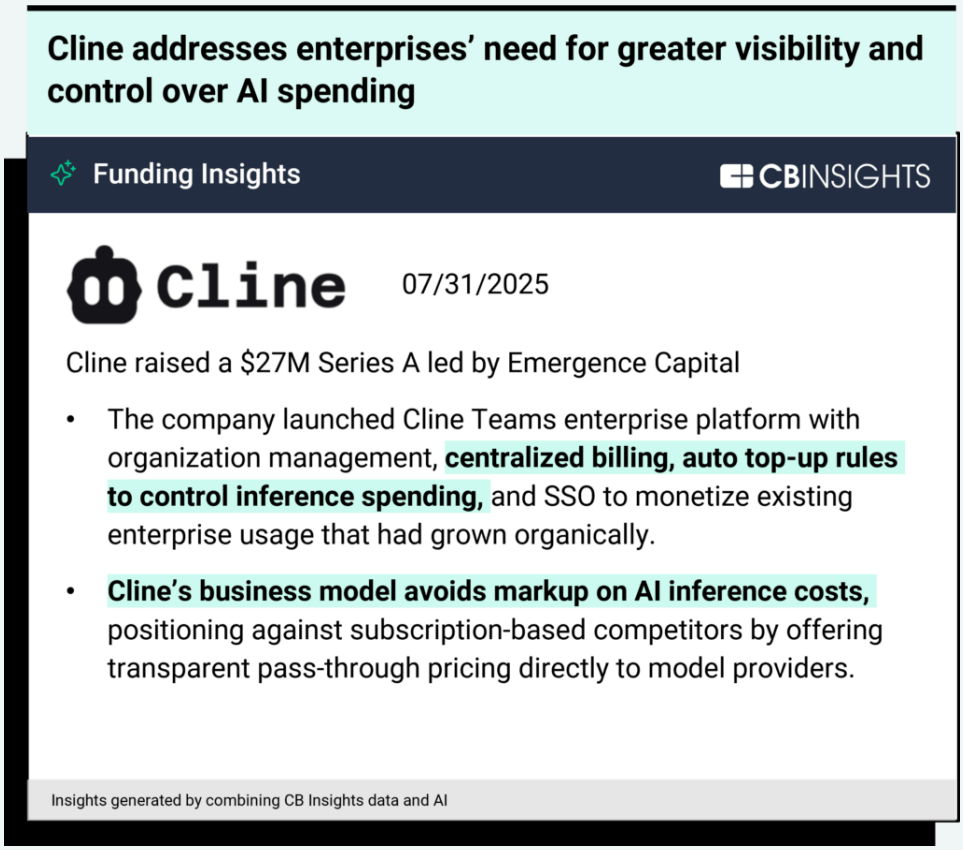
This is driving activity across observability, evaluation, and governance applications. The market has already seen 2 acquisitions in 2025 YTD. Early-stage activity highlights emerging technical needs, such as voice agent testing, with both Cekura ($2.4M seed) and Coval ($3.3M seed) focusing on evaluating and monitoring voice AI agents via simulated conversations.

Securing agents is a growing priority across the stack. Based on one-year funding activity, the AI agent security & risk management market is the fastest-growing cybersecurity segment we track as agents proliferate across enterprise environments.

White space opportunities for the AI agent ecosystem

As the AI agent tech stack matures, we predict the following areas will attract increasing innovation based on early-stage activity and recent product launches:

- **AI agent marketplaces:** Distribution is a competitive advantage, with all major cloud providers launching dedicated AI agent marketplaces, including AWS in July 2025. Companies like Olas and Agent.ai are looking to differentiate through specialized agent discovery and customization.
- **AI agent monetization:** Monetization emerges as an untapped opportunity, with companies like Paid giving visibility into AI agent costs and profit opportunities, and AGI Open Network tokenizes AI agents as tradable assets on blockchain networks.
- **Cost management:** At the end of the AI agent value chain, cost monitoring & productivity measurement will become more important as agents operate autonomously. For example, a16z-backed Larridin aims to give organizations visibility into AI spend and tool effectiveness. Other companies like coding AI agent Cline are building cost control solutions directly into their platforms to manage AI inference expenses.



Source: CB Insights Deal Agent

Category overview

Click into each market to view the full description and market players on the CB Insights platform.

Foundation models & infrastructure

Large language models (LLMs) form the cognitive core of AI agents. This layer also covers the compute, hosting, and inference systems required to serve models at scale.

- LLM developers
- Model deployment & serving

Agent frameworks & development platforms

Companies in this layer provide the software frameworks, SDKs, and low-code environments used to design, build, and deploy AI agents across different modalities and use cases.

- AI agent development platforms (including developer frameworks & platforms and low-code platforms)
- Voice
- Payments infrastructure

Tool integration

AI agents leverage "tools" to interact with external systems and perform real-world actions, such as browsing the web. This includes Model Context Protocol (MCP) implementations that standardize how agents connect to data sources and tools.

- Browser infrastructure
- AI web search APIs & scrapers
- Tool libraries & integration platforms

Context

This layer supplies agents with structured data, embeddings, and memory systems so they can retain, retrieve, and apply relevant information over time.

- LLM data preparation platforms
- Vector databases
- Memory management

Orchestration

This is the coordination layer that manages complex workflows involving multiple AI agents or models.

- Multi-agent systems & orchestration
- LLM benchmarking & model routing

Oversight

Companies here target authentication, security, monitoring, and governance functions that ensure agent actions remain safe, compliant, and aligned with intended outcomes.

- Authentication platforms
- Security & risk management platforms
- Observability, evaluation, & governance

> **Enterprise AI's value multiplies when specialized agents collaborate across vendors and organizations. The bottleneck isn't smarter individual agents — it's the missing infrastructure that lets them discover each other, verify identities, and coordinate securely at machine speed.** Think early internet before TCP/IP: brilliant isolated systems that couldn't talk. The Linux Foundation AGNTCY project solves this with open components for agent discovery, identity, messaging, and observability. We're building the plumbing that turns agent islands into collaborative workflows, because enterprises need specialized agents from multiple vendors working together, not vendor lock-in.

Guillaume de Saint Marc
VP Engineering, Outshift by Cisco

The AI agent revenue race — September's top earners show coding dominates commercialization

AI agent companies continue to grow revenue at lightning speed. We used CB Insights revenue data to rank the top private startups offering AI agents as of this month.

Published 9/22/2025 — Link to report

KEY TAKEAWAYS

- **Coding AI agents are racing ahead in commercialization**, with 6 software development agents making the top rankings, including market leaders like Anysphere's Cursor ($500M in ARR) and Replit ($150M). These startups have demonstrated that they're the most capital-efficient category, averaging $1.4M revenue per employee (compared to $594K per employee across all top agent categories). We recently identified the leaders of coding AI by market share — read more here.
- **Customer service AI agents command the highest valuation premiums**, averaging 219x revenue multiples (compared to 80x across all top revenue-generating AI agents). This valuation gap reflects investor confidence in the sector's applicability and the expectation that businesses will rapidly replace human support teams with AI agents.
- **These revenue leaders average just 3.8 years old.** Despite their youth, CB Insights' Commercial Maturity data shows the majority are already deploying or scaling their products, demonstrating the industry's compressed timelines from startup to commercial success.

AI agent startups are showing no signs of slowing down.

Since we published our first agent revenue ranking in July, the sector has continued to gain significant momentum: Harvey reached the $100M revenue threshold, category leaders like Sana Labs are exiting (Workday's $1.1B acquisition), and startups like Sierra are raising mega-rounds ($350M Series D at a $10B valuation).

Using CB Insights revenue data, we identified the top private startups generating $10M+ in revenue that offer AI agents as their primary offering and analyzed their revenue performance across categories (see the graphic below).

If you are an AI agent startup and want to submit your company's revenue data, please reach out to analyst@cbinsights.com.

The top 20 AI agents startups by revenue

We used CB Insights revenue data to rank the top private startups offering AI agents as a primary product.

CBINSIGHTS

Rank[3]	Company	Market Founded year	Annual revenue[1] ▼	Revenue multiple[2]	Revenue per employee
– 1	CURSOR (Anysphere)	Software development 2022	$500M	20x	$3,205K
▲ 2	replit	Software development 2016	$150M	20x	$758K
▼ 3	glean	Enterprise workflow 2019	$100M	72x	$82K
▲ 3	Harvey	Legal 2022	$100M	50x	$148K
▼ 3	Lovable	Software development 2023	$100M	18x	$2,222K
▼ 3	MERCOR	HR 2023	$100M	20x	$1,333K
▼ 7	crescendo	Customer service 2024	$91M	6x	$47K
NEW 8	manus	Enterprise workflow 2022	$90M	N/A	N/A
NEW 9	Cognition	Software development 2023	$73M	140x	$702K
▼ 10	StackBlitz	Software development 2017	$40M	N/A	$1,143 K
▼ 11	clay	Sales 2017	$30M	103x	$1,000 K
▼ 12	torq	Security operations center (SOC) 2020	$24M	N/A	$75K
▼ 13	SIERRA	Customer service 2023	$20M	500x	$65K
▲ 14	Fyxer.ai	Enterprise workflow 2024	$17M	4x	$236K
▼ 15	Nabla	Healthcare 2018	$16M	12x	$149K
▼ 16	Hebbia	Enterprise workflow 2020	$13M	54x	$96K
NEW 17	sintra	Enterprise workflow 2023	$12M	N/A	$333K
▼ 18	11x	Sales 2023	$10M	35x	$21K
▼ 19	Decagon	Customer service 2023	$10M	150x	$68K
NEW 19	emergent	Software development 2024	$10M	N/A	$154K
▼ 19	Robin	Legal 2019	$10M	N/A	$49K

Data as of 9/16/2025. [1] Annual revenue based on latest annual revenue or latest annual recurring revenue available. [2] Calculated using latest valuation available divided by latest annual revenue available. [3] Indicates change in rank since the last version of this list.

If you are an AI agent startup and you want to submit your company's revenue data, please reach out to analystbriefing@cbinsights.com.

CBINSIGHTS

See the top AI agent startups ranked by revenue here

AI agents make inroads across enterprise workflows

From Y Combinator's latest batch to cloud giants' strategies, how agents are transforming coding, customer service, and backend operations

Y Combinator's 2025 Spring batch reveals the future of agentic AI

We analyze Y Combinator's Spring 2025 batch to uncover the top trends in agentic AI, spanning workflow automation, software development, and applications in highly regulated industries.

Published 6/20/2025 — Link to report

KEY TAKEAWAYS

- **Software development agents are evolving beyond coding to include critical guardrails:** With 11 companies in this category and 2025 funding (for the broader market) already outpacing 2024 by 3x, startups are moving beyond basic coding assistance to provide testing, QA, code review, and debugging solutions. Over half focus specifically on making "vibe coding" less risky through browser-based testing agents and automated review systems, addressing reliability concerns that have hindered wider enterprise adoption.
- **Web-browsing agents gather steam beyond general-purpose use:** Y Combinator backs over 50% of the existing web-browsing agent market, but these startups are differentiating through targeted applications like legacy system integration, software testing, and quality assurance. This shift toward specialized browsing agents provides more contextual data access, improving decision-making and autonomy compared to general-purpose alternatives.
- **Highly regulated industries are becoming prime targets for vertical AI agents:** Healthcare and financial services represent 19% of agentic AI companies in this batch, with 32% of verticalized agents actively deploying solutions and another 45% in emerging/validating stages. Beyond traditional customer service applications, startups are tackling industry-specific workflows like mortgage processing and healthcare operations, with some beginning to handle autonomous research functions that could eventually replace human roles entirely.

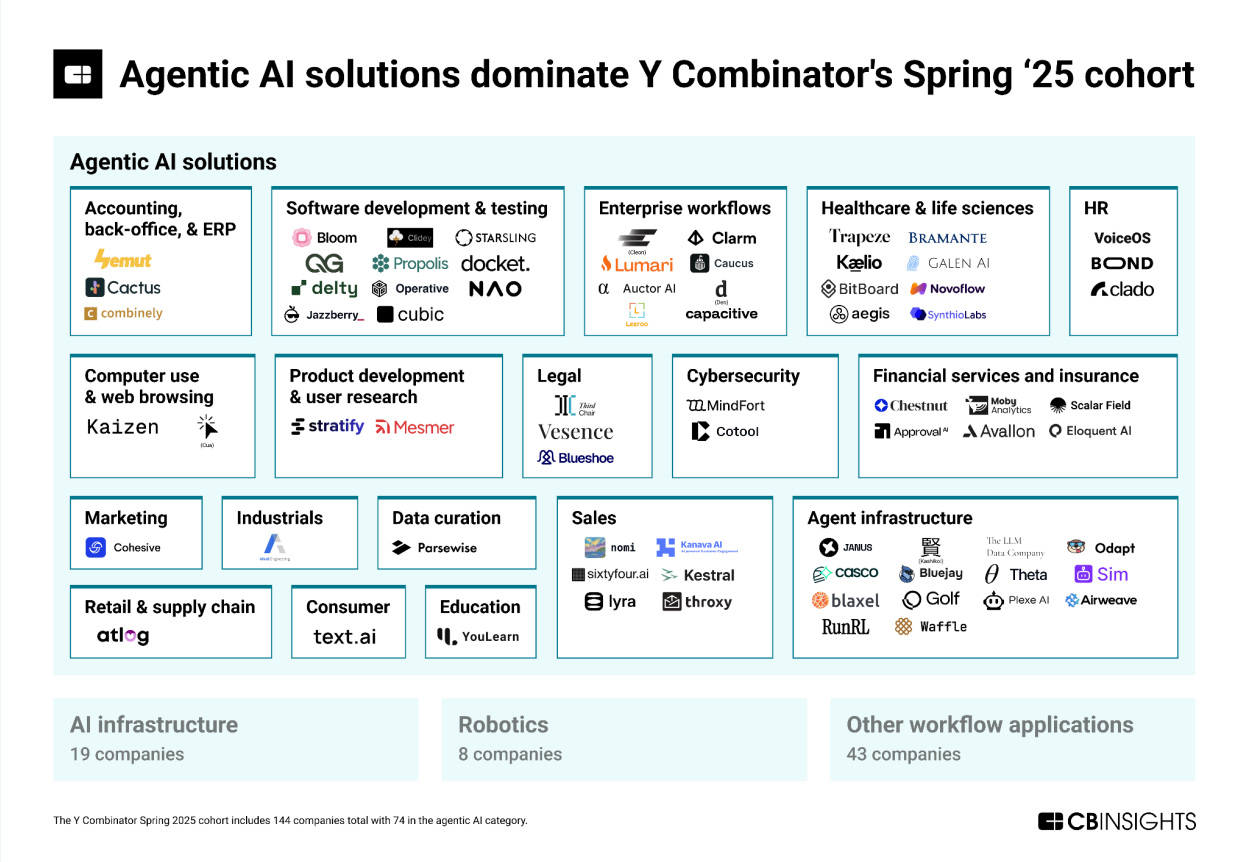
Y Combinator's Spring 2025 batch is a preview of agentic AI's future: over half of the 144 companies are building agentic AI solutions, providing valuable insights for enterprise AI strategies.

The accelerator that spotted OpenAI, Airbnb, and Stripe before they became household names is now placing bets across 4 key agentic AI areas: software development guardrails

that de-risk “vibe coding”, web-browsing agents, backend workflow automation, and vertical agents penetrating highly regulated industries.

For strategy teams, this represents both a roadmap of where agentic AI is heading and a curated list of potential acquisition targets, partners, and competitive threats.

Using CB Insights, we mapped the 70+ agentic AI companies in the Y Combinator’s 2025 Spring batch across 18 different categories.



Check out the 140+ startups in Y Combinator’s newest cohort

De-risking AI software development

Software development and testing is the second-largest agentic AI category of this batch, with 11 companies. This reflects the fact that software development AI agents are still booming, with 2025 funding already outpacing 2024 by 3x. Yet this cohort goes beyond coding AI agents, providing engineering support, QA, and guardrails to make vibe coding less risky.

A key focus of these companies is to make vibe coding less risky. For example, over half of the startups in this category focus on testing and review. Operative deploys browser agents that can test coding agents. Docket and Propolis use web agents to QA code and products. Startup Cubic reduces code review bottlenecks, and Jazzberry debugs code, both of which are issues becoming more prominent with the rise of vibe coding.

A handful of companies are developing solutions to support software engineers in vibe coding and automated code generation. Delty is an AI agent that helps with system design and architecture based on deep codebase understanding, and StarSling provides AI agents to augment DevOps.

These new tools will accelerate the growth of more reliable AI software development, boosting existing leaders in the space such as Cursor who could look to acquire them.

Web-browsing agents gather steam beyond general-purpose use

Y Combinator's dominance in web-browsing agents – backing over 50% of the existing market — signals this emerging category's potential to become critical infrastructure for agentic AI. LLM giants like OpenAI are already building their own browser agents, but this isn't deterring startups from entering the space

The Spring 2025 batch reveals how these startups are differentiating themselves by targeting high-value, specific applications rather than building general-purpose agents.

For example, Kaizen provides browser agents that enable outdated, legacy systems to connect with websites without the need for an API. Operative and Propolis are pioneering the use of browsing agents for software testing and quality assurance, areas where automation has historically struggled.

Agents capable of accessing and browsing the web can access more data and information than what is typically available in a company's systems. This helps provide more context to agentic systems, improving decision-making, and ultimately autonomy.

Agents are coming for the backend

Today, most AI agents focus on frontend interactions and applications, with customer service and enterprise workflow being 2 of the most well-funded AI agent markets. This Y Combinator cohort signals how agents are moving to the backend.

Cactus, Combinely, and Hemut are building back-office systems in areas like accounting and reporting. Caucus and Cohesive developed agent-based CRMs that go beyond the traditional enterprise space to target small businesses and government. Odapt allows custom application development in areas like finance and marketing, built on top of existing tools and systems. Cleon and Auctor AI are automating system implementations.

Currently, these companies focus on narrowly defined, specialized backend workflows. Expanding into more end-to-end workflows will require greater trust in agentic AI applications.

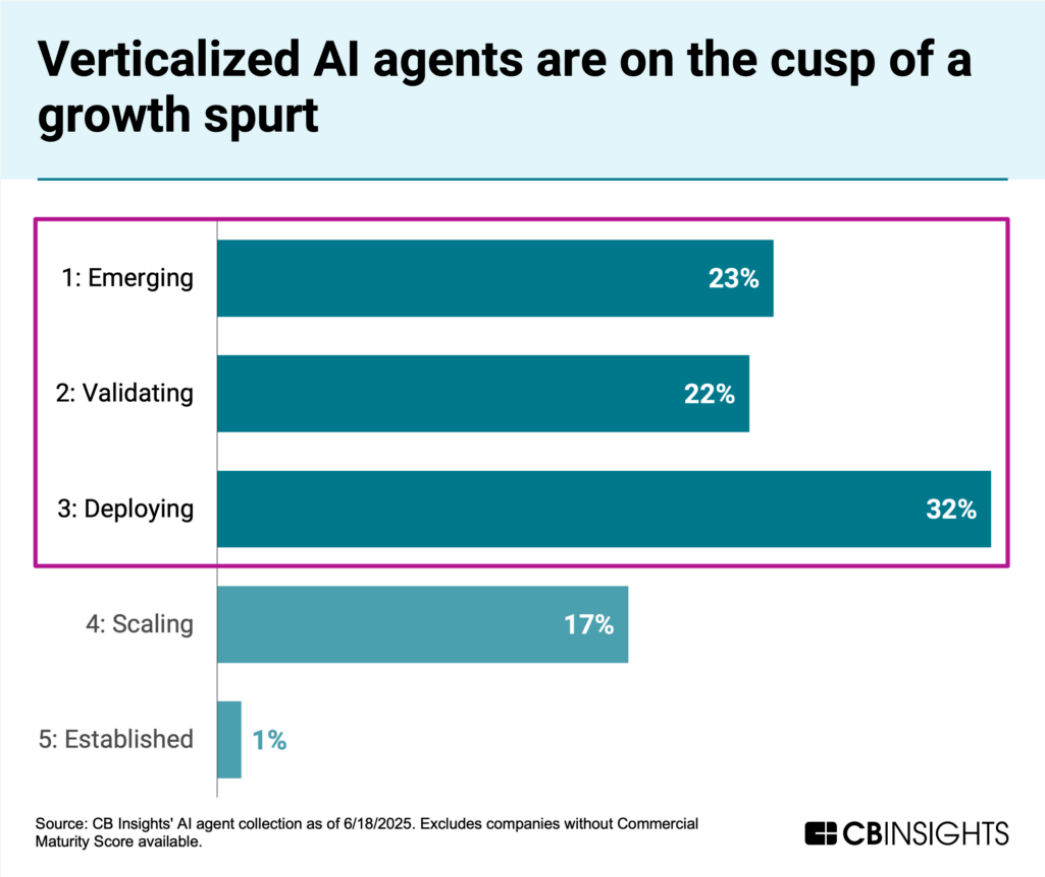
This trust can be partially built through the ability to benchmark AI agent performance. Kashikoi, Janus, and The LLM Data Company – all part of this Spring cohort – are working on this today.

AI agents keep making inroads in highly regulated industries

Once an obstacle for new AI applications, the most highly regulated industries have emerged as targets for agentic AI startups. 32% of verticalized AI agent companies are actively deploying solutions, and 23% and 22% are emerging and validating, respectively, suggesting an oncoming growth spurt.

This impending growth is fully displayed with this batch of Y Combinator companies, particularly in healthcare and financial services, which represent 19% of the agentic AI companies in this year's Spring cohort.

Customer service and engagement are common areas of focus within these verticals, with companies like Eloquent AI (financial services), Trapeze (healthcare), and Kaelio (healthcare). Other startups are delving deeper into industry-specific workflows, like Chestnut Mortgage and Approval AI (lending and mortgage), and Bitboard (healthcare operations).



We expect the next generation of industry-focused AI agent companies to go beyond operational support and handle research autonomously.

A handful of companies in this batch tackle research assistance today, like Bramante Biologics and SynthioLabs in healthcare and Scalar Field for investment research. These startups lay the foundation for a future in which AI agents can proactively source, digest, and deliver information to human users or automate their roles altogether.

Building the agent economy: How cloud leaders are shaping AI's next frontier

The top 3 cloud providers globally are racing to support AI agents and the companies enabling them. This report analyzes their strategies — based on partnerships, startup investments, and internal projects — to see how they're capitalizing on this next wave of monetization in AI.

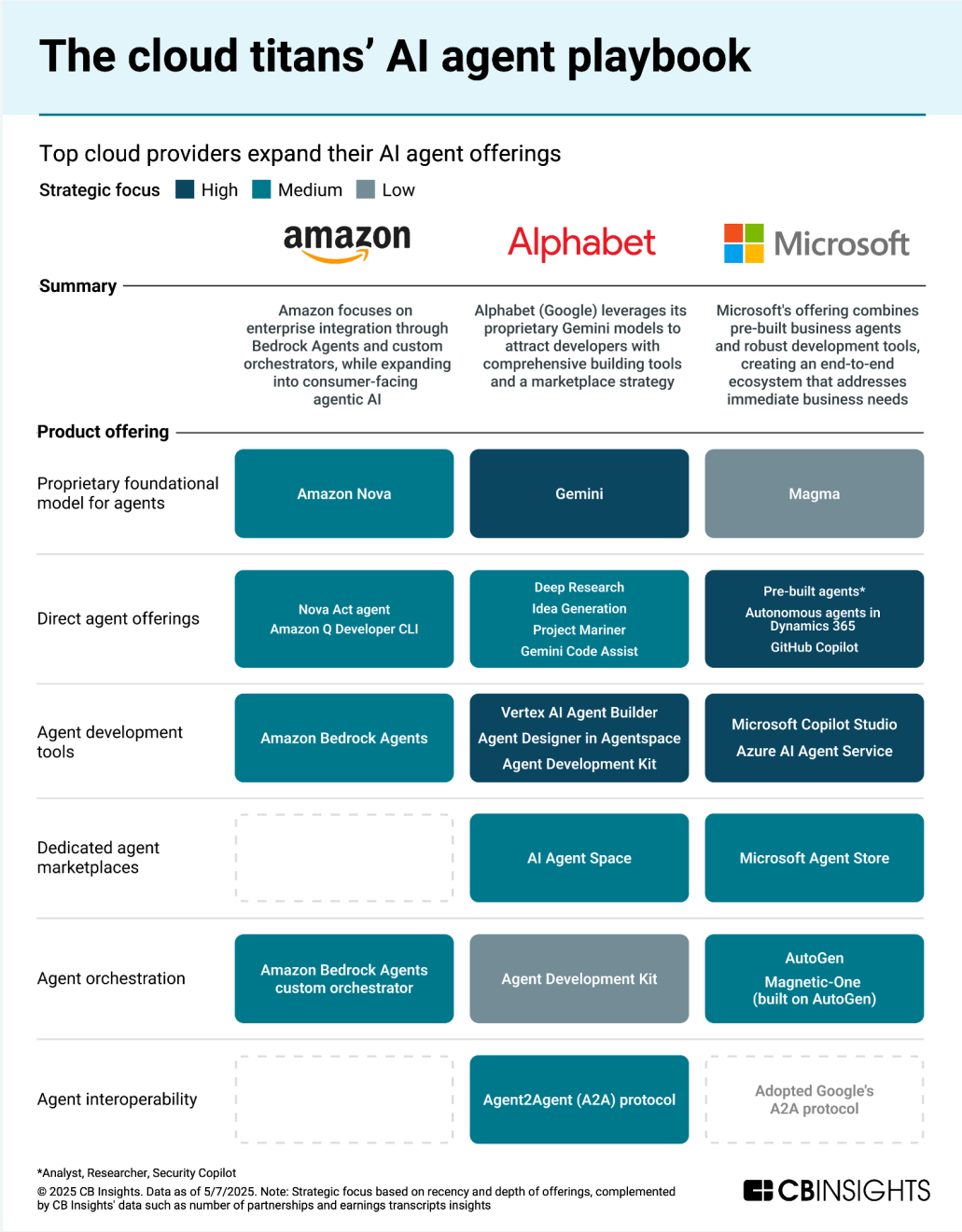
Published 5/7/2025 — Link to report

KEY TAKEAWAYS

- **Amazon positions itself as a neutral infrastructure layer for the agentic ecosystem,** betting on in-house chips and seeding its ecosystem with 16 total investments in agent startups. These investments, which primarily take the form of cloud credits rather than equity, form a low-risk, high-volume strategy to lowering the barriers to building on AWS. Amazon's approach combines enterprise enablement with strategic consumer-facing investments, signaling potential integration with its broader ecosystem.
- **Google's agentic offering centers around its Gemini foundation models, creating an open marketplace** for partner-built agents leveraging its technological leadership — supported by 46 total partnerships and its Agent2Agent protocol. This partner-centric approach means Google can quickly populate its marketplace with specialized agents while maintaining Gemini as the differentiating foundation technology.
- **Microsoft emphasizes a pre-built suite of agentic solutions to drive enterprise adoption,** embedding Copilot agents throughout its productivity ecosystem. The company recently achieved 15M GitHub Copilot users, up 4x YoY, and recorded 1M custom agents created on its SharePoint and Copilot Studio platforms. Microsoft's expansive enterprise client base gives it an in-built audience for new agent products.

As the AI boom accelerates, the top 3 global cloud providers — Amazon, Microsoft, and Google — are racing to capture a larger share of enterprise AI spend. Central to this shift is the rise of AI agents: intelligent systems capable of performing multi-step tasks, interacting autonomously with tools and data, and automating business workflows.

Drawing on CB Insights' Business Graph, which links data across private investments, business relationships, and public company disclosures, we surface key signals on how each cloud player is positioning itself in the agentic AI space and what their next move could be.



Note: Since publication, Amazon also adopted Google's A2A protocol and launched its own dedicated agent marketplace.

While all 3 providers are investing heavily in infrastructure to support agentic AI, they are taking distinct paths to monetization and market control — from proprietary models and low-code build tools to strategic partnerships and go-to-market accelerators.

Understanding these differences will prove critical in evaluating cloud alignment, competitive positioning, and agent-enabled product strategy.

Amazon positions itself as the neutral infrastructure layer for the agentic ecosystem

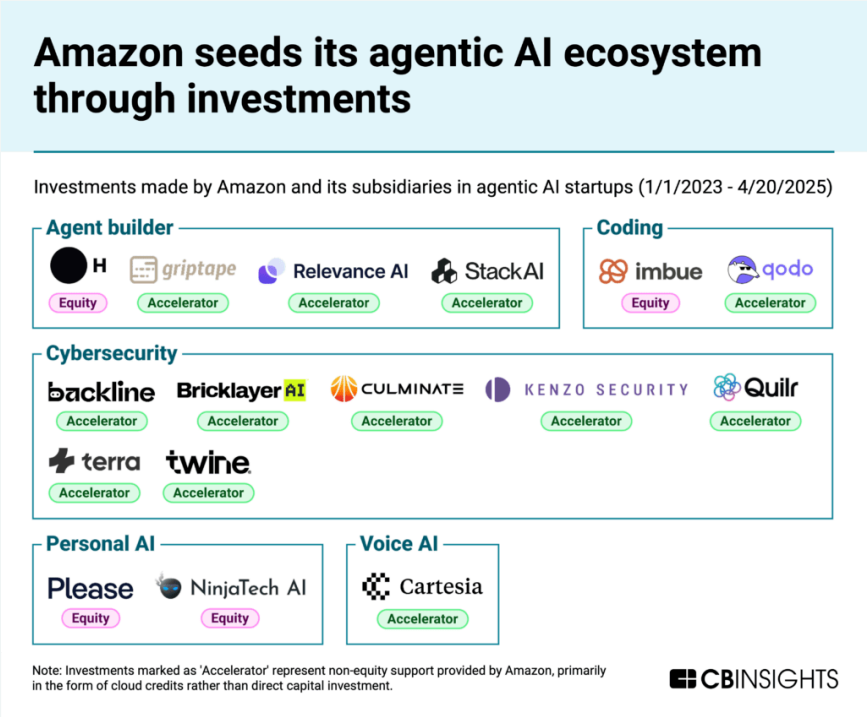
Amazon is approaching the agentic AI landscape as a pragmatic infrastructure provider, with a strategic bias toward enabling partners rather than competing with its own agent suite. This partner-first approach comes at a critical time, as Amazon has been playing catch-up in the agentic race — although recent moves, including forming a dedicated agentic AI group in March 2025 and unveiling its Nova foundational models in December 2024, signal a growing focus.

Amazon is betting on its in-house chips, Trainium and Inferentia2, to attract agentic AI workflows, as these chips can help reduce the cost and energy consumption associated with AI model training and inference. Amazon has already formed several partnerships with agentic AI startups such as Poolside and NinjaTech for them to train and run their AI agents on its in-house chips.

This approach — of providing robust infrastructure and letting specialized partners build solutions on top — is also reflected in its agent development tool offerings. While developers can build agents using Amazon Bedrock Agents, Amazon (unlike Google or Microsoft) doesn't directly emphasize low-code/no-code solutions, instead enabling partners like SnapLogic to build such tools on its platform.

The company has also been investing heavily in agentic AI startups, with 16 unique investments since 2023 — more than both Google and Microsoft combined. However, 12 of these were made through non-equity accelerator programs that provide cloud credits and technical enablement rather than capital.

This low-risk, high-volume approach lowers the barriers to building on AWS while seeding future clients at minimal cost. It also embeds AWS infrastructure into early-stage agent development, capturing mindshare before competitors can gain traction.



Amazon's investments reveal a strategic interest in consumer-facing AI applications that complement its existing business. Three of its four equity investments are in consumer-focused companies — Please and NinjaTech (personal AI agents) and Cartesia (voice AI) — aligning with Amazon's consumer strategy and the recent launch of Nova Act, its web-browsing agentic AI targeting developers.

These investments suggest AWS is taking a dual strategy: framing itself as an enterprise infrastructure provider for partners while developing consumer-facing capabilities that could enhance Amazon's broader ecosystem, including potentially a revamped Alexa.

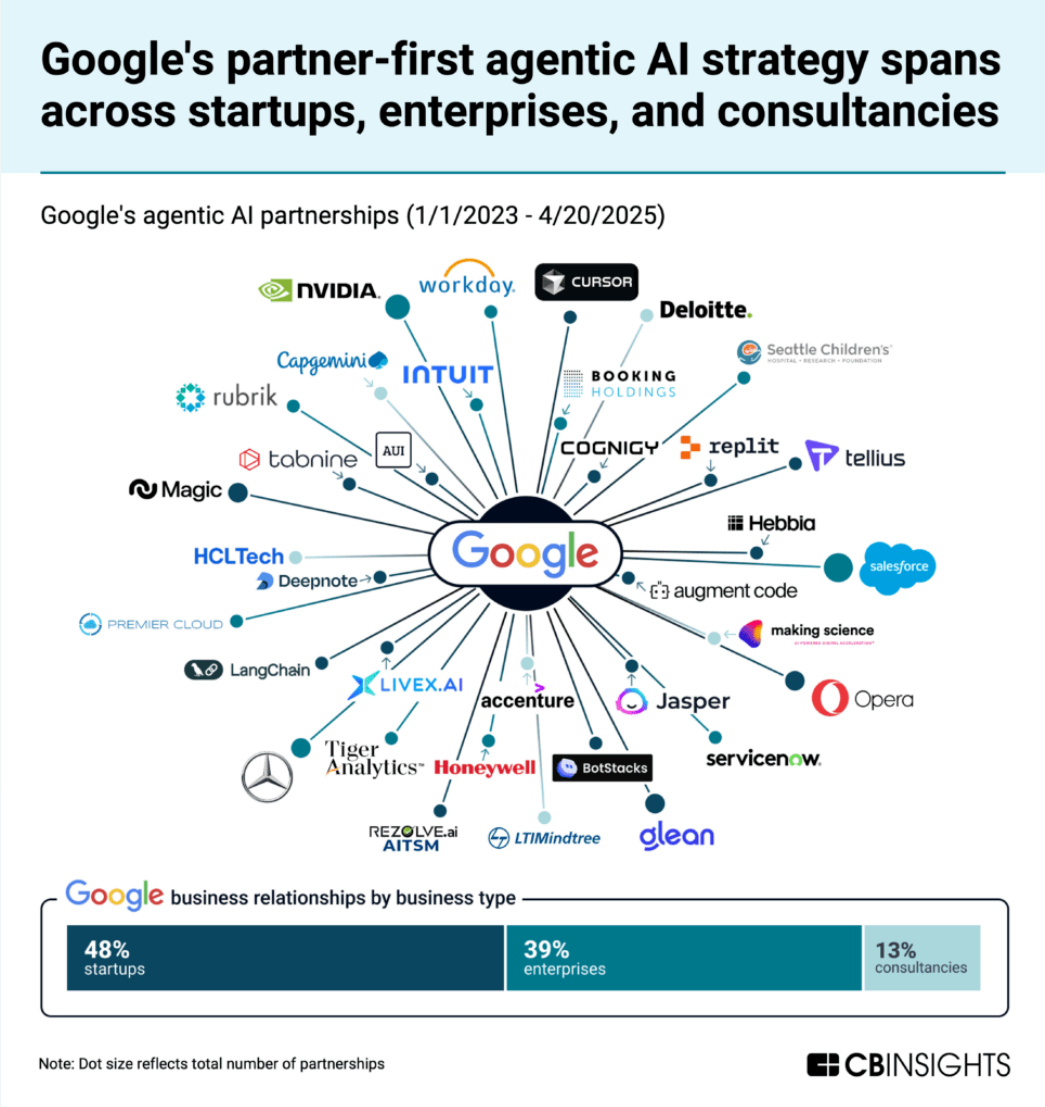
This could lead Amazon to make an acquisition that accelerates monetization of consumer-facing agents. Acquiring a startup developing agent payment infrastructure, for instance, would support efforts to enable autonomous transactions.

Google's agentic offering centers around its Gemini foundation models

Google has positioned itself as the central platform provider in the agentic AI landscape, building a comprehensive ecosystem centered around its proprietary Gemini foundation models. Unlike Amazon's infrastructure-focused approach or Microsoft's enterprise application strategy, Google is creating an open marketplace for partner-built agents that leverage its technological leadership.

To boost adoption of its Gemini models for agentic AI, Google unveiled its AI Agent Space last year, a dedicated marketplace exclusively for partners' agents. This is complemented by Google's agent interoperability initiative, the Agent2Agent (A2A) protocol, which enables AI agents to communicate effectively regardless of their underlying frameworks or vendors. As a sign of traction for A2A, Microsoft recently announced it would adopt the protocol, in addition to the 50+ supporting partners Google already unveiled early April this year.

Google leads in agent-related partnerships with 46 collaborations — 2x as many as Microsoft and Amazon. Almost half of these are with agentic AI startups, including AI coding agents like Cursor, Augment Code, and Replit. This partner-centric approach means Google can quickly populate its marketplace with specialized agents while maintaining Gemini as the differentiating foundation technology.



Source: CB Insights — Google's business relationships. Note: includes business relationships for Google Cloud.

Enterprise partnerships also demonstrate Google's strategy in action. Its recent Salesforce collaboration will empower Salesforce customers to build Agentforce agents using Gemini, while Deloitte has launched over 100 ready-to-deploy AI agents powered by Google's models. According to Google Cloud, more than 60% of generative AI startups are now building on its platform.

Google has also been partnering with leading venture capital firms and accelerators, like Sequoia, Lightspeed, and Y Combinator, to promote the use of its technology (such as TPUs and Gemini models) to fast-growing startups that are building with AI.

Google's development toolkits — Vertex AI Agent Builder, Agent Designer in Agentspace, and Agent Development Kit — offer solutions for both technical developers and non-technical users, reflecting Google's goal of becoming the complete platform for agent creators and consumers alike.

Rather than building a comprehensive first-party agent suite, Google is embedding itself into the tech stack of emerging agentic players, making Gemini the platform of choice for agent innovation.

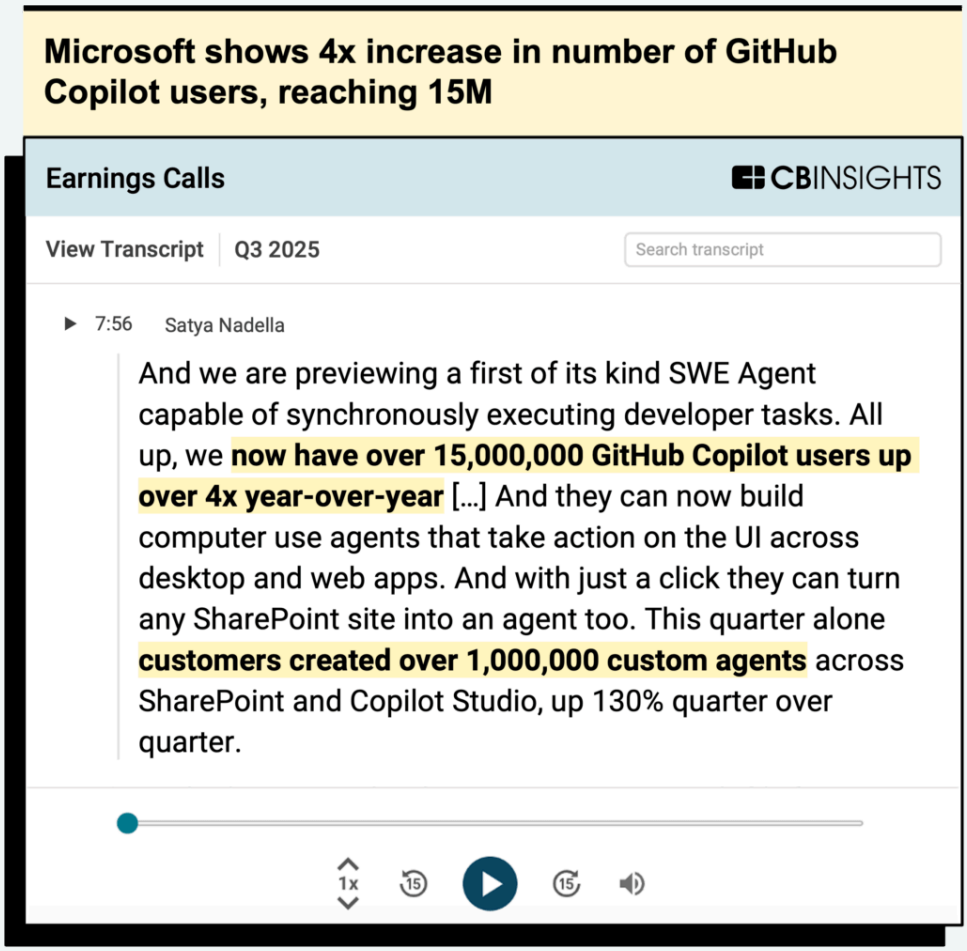
To maintain its edge in safe scaling and cross-agent coordination, Google may look to acquire companies focused on monitoring, governance, and lifecycle tooling, such as Galileo — a leader in AI evaluation backed by Databricks and ServiceNow.

Microsoft emphasizes a pre-built suite of agentic solutions to drive enterprise adoption

Microsoft's offerings center around a comprehensive suite of pre-built agents deeply integrated into its productivity ecosystem. While Amazon focuses on infrastructure and Google on promoting its foundational models, Microsoft aims to deliver immediate business value through turnkey solutions.

The company leads the market in pre-built agent offerings, with its Copilot suite including Analyst, Researcher, Security, and Dynamics 365 autonomous agents — all powered by its exclusive access to OpenAI's models.

This strategy has driven strong adoption: Microsoft's Q3 FY'25 earnings call revealed that GitHub Copilot's developer base has surpassed 15M users (up 4x YoY), while 1M custom agents were created during that quarter through Copilot Studio and SharePoint.



Source: CB Insights — Microsoft Earnings Insights

Microsoft's development tools (Copilot Studio, Azure AI Agent Service) cater to both technical and non-technical users, but the company's primary advantage comes from embedding agentic capabilities throughout its productivity ecosystem. The November 2024 launch of Magentic-One, a multi-agent system for enterprise deployment, further enhances Microsoft's position in business workflows.

Unlike Amazon's broad ecosystem-seeding or Google's push to embed Gemini models into any agentic workflow, Microsoft concentrates on initiatives that complement its in-house agentic tools. Its partnership with Moveworks exemplifies this strategy, allowing employees to access Moveworks' specialized agents directly within Microsoft 365 Copilot and Teams.

Microsoft's approach demonstrates the power of integration over technological differentiation in driving enterprise adoption. By leveraging its existing relationships and software suite, Microsoft has established dominance in high-value business workflows where agentic AI delivers immediate productivity gains — and where competitors must overcome Microsoft's entrenched position.

To round out its Copilot suite and reinforce its workflow ownership strategy, Microsoft may seek acquisitions in sectors where it lacks native agent offerings, like recruiting, healthcare administration, or logistics.

The summer of vibe coding is over — How reasoning models broke the economics of AI code generation

Reasoning models have supercharged coding AI agents, driving adoption and growth, but at the cost of higher inference expenses that are squeezing margins. This report explores how these companies are adapting and what lessons other AI agent markets can take from coding.

Published 8/28/2025 — Link to report

What started as a gold rush in AI-powered coding may be turning into a money pit, offering a preview of challenges awaiting other AI agent categories.

Companies that hit $100M+ ARR in months, like Anysphere (maker of Cursor) and Lovable, now face LLM inference costs growing up to 20x, forcing rate limits and price hikes, and putting reverse acqui-hires (hiring founders and licensing the tech) on the table as some founders seek exits.

Using CB Insights' data on company momentum, exit probabilities, and customer sentiment, we analyzed how the coding AI market is adapting to this economic shock and what other AI agent companies (and their backers) can learn:

- Reasoning models spark vibe coding's explosive growth
- Reasoning token shock pushes adoption of new pricing models
- Margin pressure drives consolidation of talent in the coding AI agents market
- Open models and usage-based pricing offer solutions to the market's current challenges

Reasoning models spark vibe coding's explosive growth

The coding AI agents and copilots market has been on a roll, generating an estimated $1.1B in revenue in 2024 and minting unicorns in as little as 6 months, which is 4x faster than the AI industry average.

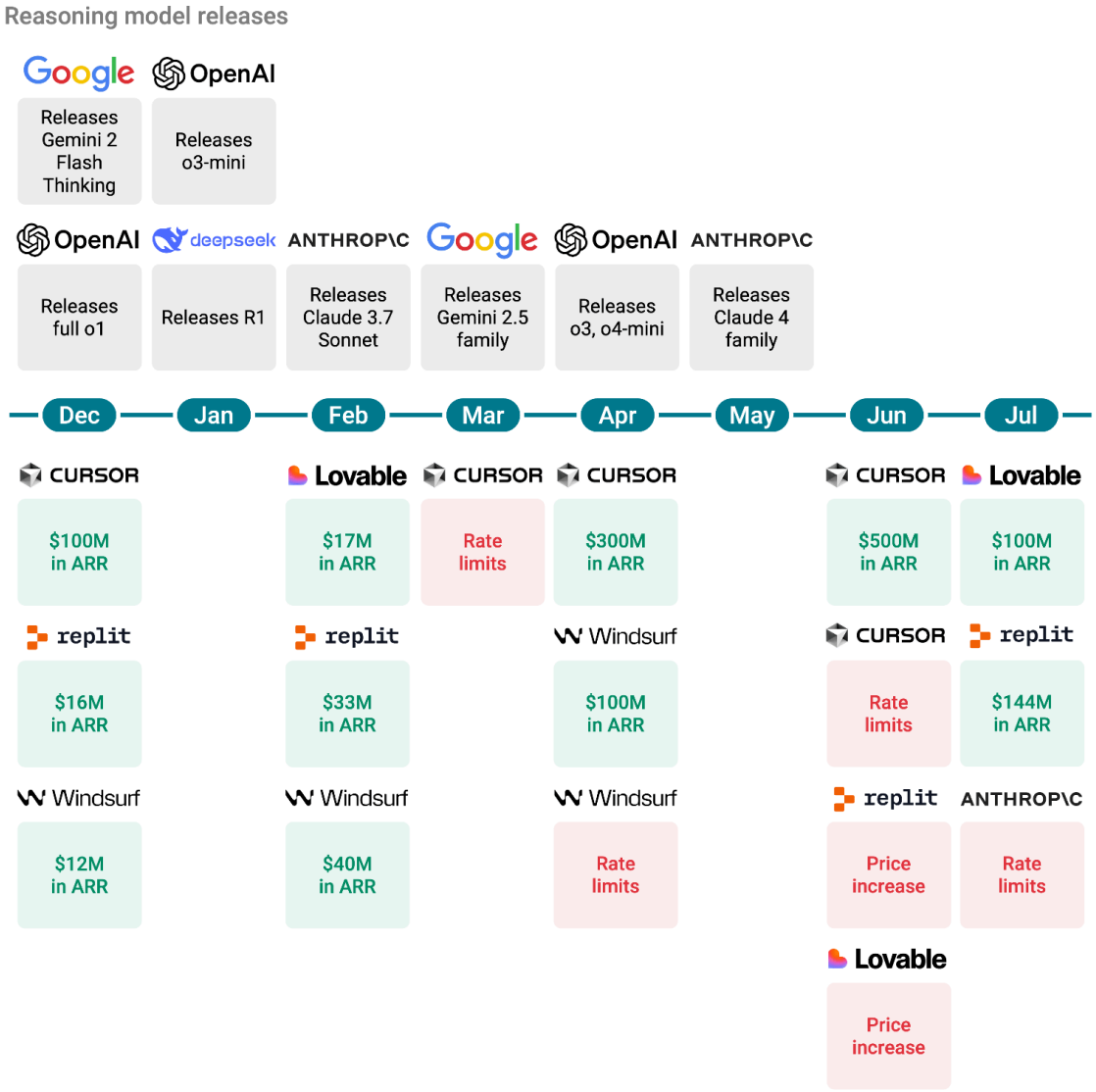
Anthropic's release of Claude 3.5 Sonnet in June 2024 has primarily driven this early momentum. This technology helped developers transition from autocomplete to partial delegation of coding tasks with a model that could reliably call tools and handle multi-file edits.

But it is the emergence of reasoning models, and specifically Anthropic's Claude 3.7 Sonnet's reasoning mode in February 2025, that made vibe coding possible — giving a high-level goal and delegating multi-step implementation to the AI. Developers could now set goals like "make this component responsive" or "add error handling throughout" and let the AI plan and execute the changes, sparking explosive growth in the space:

- Anysphere's ARR grew 5x in 6 months, from $100M in December 2024 to $500M in June 2025.
- Replit's ARR increased from $10M at the end of 2024 to $144M in July 2025.
- Lovable became one of the fastest-growing software startups, reaching $100M in ARR just 8 months after launching.

Coding AI agent companies scramble as reasoning models destroy unit economics

Timeline of reasoning model release and subsequent rate limit/price adjustments from coding agent startups.



See all coding AI agent companies ranked by revenue

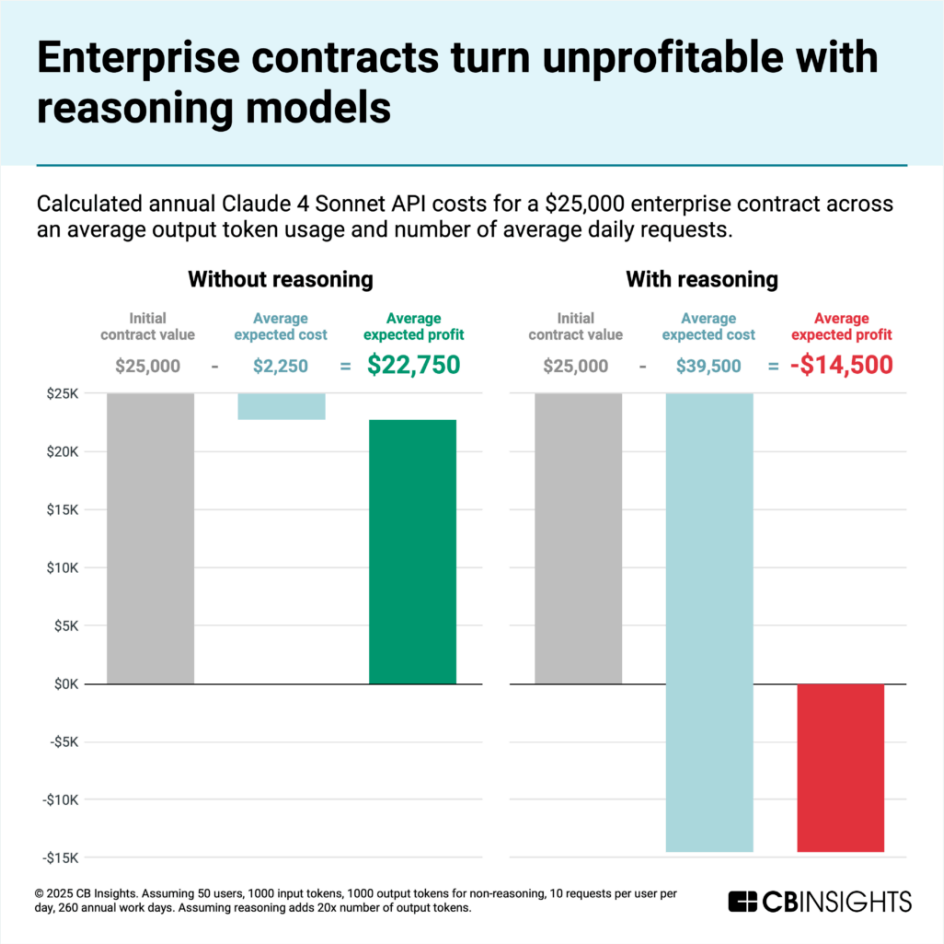
Reasoning token shock pushes adoption of new pricing models

As revenue surged on the back of reasoning, costs rose even faster.

Reasoning models inflate output-token volume roughly 20x, according to Artificial Analysis. Because inference is billed per token — and output tokens are typically priced higher than input — that surge translates directly into higher compute cost. Anthropic's May 2025 step-ups on Sonnet 4 and Opus 4 (priced at roughly 5x prior models) added further pressure just as adoption was accelerating.

This is particularly impacting enterprise deals, which businesses often negotiate on an annual, per-seat basis. That structure leaves vendors carrying the risk of uncapped compute costs while revenue stays fixed.

Using CB Insights Customer Sentiment data, we find most contracts fall between roughly $6K and $100K a year, with a median around $25K for a 50-developer team. While margins once sat at 80%-90% on these contracts, compute costs from reasoning models can flip margins deeply negative.



The strain showed up quickly. Cursor tightened rate limits and introduced overage charges despite crossing $500M in ARR, prompting backlash and refunds. Anthropic throttled Claude Code after individual users exceeded $10K in monthly compute on $200 plans.

Vendors are shifting to pass-through and usage-based pricing to align revenue with compute cost. Companies employing usage-based approaches show stronger momentum in our Mosaic data (median Momentum Mosaic of 683 vs. 671 for the broader market), but enterprise buyers are pushing back on variable bills and month-to-month swings.

Expect coding AI agent vendors to adapt pricing and GTM: moving to seat-plus-usage hybrids, stricter per-seat compute guardrails, and model tiering that reserves reasoning for high-impact work. ARR growth will moderate as flat-fee expansion gives way to usage-aligned pricing.

Margin pressure drives consolidation of talent in the coding AI agents market

Reasoning-driven margin compression is forcing consolidation in a category that has seen dozens of new entrants over the past 12 months.

Traditional acquisitions aren't off the table, but acqui-hires and reverse acqui-hires have become the most active exit structures recently — albeit with trade-offs.

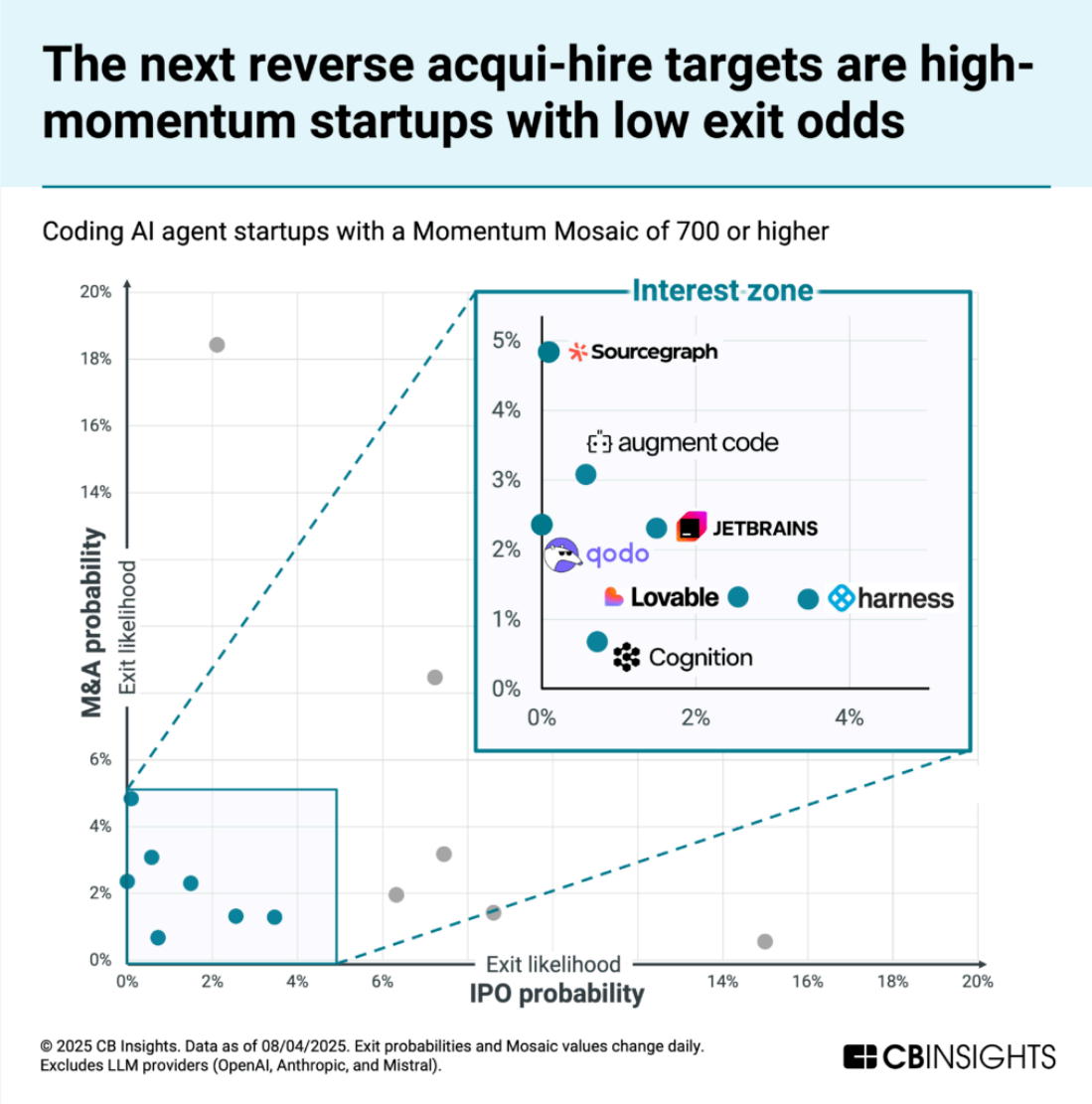
OpenAI and Anthropic have logged 3 acqui-hires since early 2025. Across AI, recent moves (e.g., Microsoft–Inflection AI, Amazon–Adept, and Meta–Scale) signal a tilt to talent-plus-license amid potential antitrust scrutiny. In coding AI agents, Windsurf's failed sale and Google's follow-on reverse acqui-hire underscore the pattern of buyers taking teams and leaving products behind.

In these deals, acquirers hire the team and license the tech, leaving customer contracts and infrastructure — and the associated compute liabilities — outside the transaction. What they're buying isn't raw model IP; they're buying proven operators with successful track records.

CB Insights' exit probability analysis points to the next likely targets: companies with high Momentum Mosaic scores but lower probabilities of traditional exits.

The likely cause: private-market valuations have outrun what strategics or public investors will pay given reasoning-driven margin pressure, product overlap, and antitrust scrutiny — making full-company M&A or near-term IPOs harder to underwrite.

Seven stand out as potential targets: Sourcegraph, Augment Code, JetBrains, Qodo, Lovable, Cognition, and Harness.



Expect more reverse acqui-hire deals over the next few quarters as big tech continues to push for talent while coding AI agent companies struggle under margin pressures.

Open models and usage-based pricing offer solutions to the market's current challenges

Against that backdrop, two levers dominate today: open models and usage-aligned pricing. Here's how each is playing out — and where it falls short.

Open models cut costs, but enterprise requirements slow adoption

Moonshot AI's Kimi K2, Alibaba's Qwen-Coder, and Z.ai's GLM-4.5 approach Claude on coding tasks at a fraction of the cost, and OpenAI's gpt-oss goes a step further by offering a model that can run on consumer hardware.

Yet users need to access these models either through self-hosting or a third party. For enterprises, that means fresh security reviews, stringent uptime service level agreements (SLAs), multi-hour agent-run testing, and new infrastructure to manage.

The result is slower adoption, especially for six-figure contracts that expect Claude-level reliability.

Usage-based pricing fixes vendor margins, but most enterprises resist variable bills

Buyers tell us that token-metered pricing is difficult to budget, and expectations around costs for these tools are already set. CFOs want to anchor budgets and avoid month-to-month swings tied to release cycles, while usage-based pricing is the exact opposite.

GitHub

The vendor's pricing model aligns with our needs and expectations because it is very easy to understand. It's essentially an added premium on top of an existing seat license … We prefer not to look into pay by token, for example, because it's still a rather nebulous concept to most of our enterprise.

VP,
Fortune 500 company

CBINSIGHTS

Anysphere

However, if any of these tools suddenly wants to charge [more], there is no way for us to digest that because there are many other tools that will cost within that range that I mentioned. The expectation is already set in what it should cost.

VP of Engineering,
Construction company

CBINSIGHTS

In the near term, expect a shift from per-message metering to effort-based task pricing: agents quote a fixed rate for a defined outcome (e.g., “add error handling across this service” or “convert this component to TypeScript”), bundling planning, tool calls, and verification into a single charge with a visible pre-estimate. Tasks are tiered (S/M/L) with caps on reasoning usage and admin-approved overages, giving CFOs predictable bills while keeping compute under control.

This dynamic won’t be limited to coding

Other agent categories with surging usage are likely to rework pricing and contracts as reasoning costs mount.

Customer service is already operating on usage/outcome models. For example, in May 2025, Salesforce’s Agentforce shifted prices from $2 per conversation to a hybrid-usage Flex Credits system, tying credits to necessary actions for an outcome. Zendesk did a similar shift in pricing strategy in November 2024. Yet reasoning-heavy workloads still create margin risk when the compute to achieve a resolution outstrips the value captured.

Beyond customer service, expect similar recalibrations across legal, healthcare, and sales agents. Outcome- or usage-based models don’t fully eliminate compute risk. Explosive top-line growth can mask deteriorating unit economics as reasoning workloads scale, and recent mega-rounds may not be enough to foot the bill. Many players will reprice, add stricter usage guardrails, or raise additional capital to stay in the game.

Industry-specific applications gain momentum

Vertical-specific agents gaining traction in retail, finance, healthcare, and manufacturing with real-world deployments and measurable ROI

3 markets fueling the shift to agentic commerce

The building blocks of agentic shopping are emerging. In this report we highlight the markets gaining traction, who's leading, and what's next.

Published 8/4/2025 — Link to report

KEY TAKEAWAYS

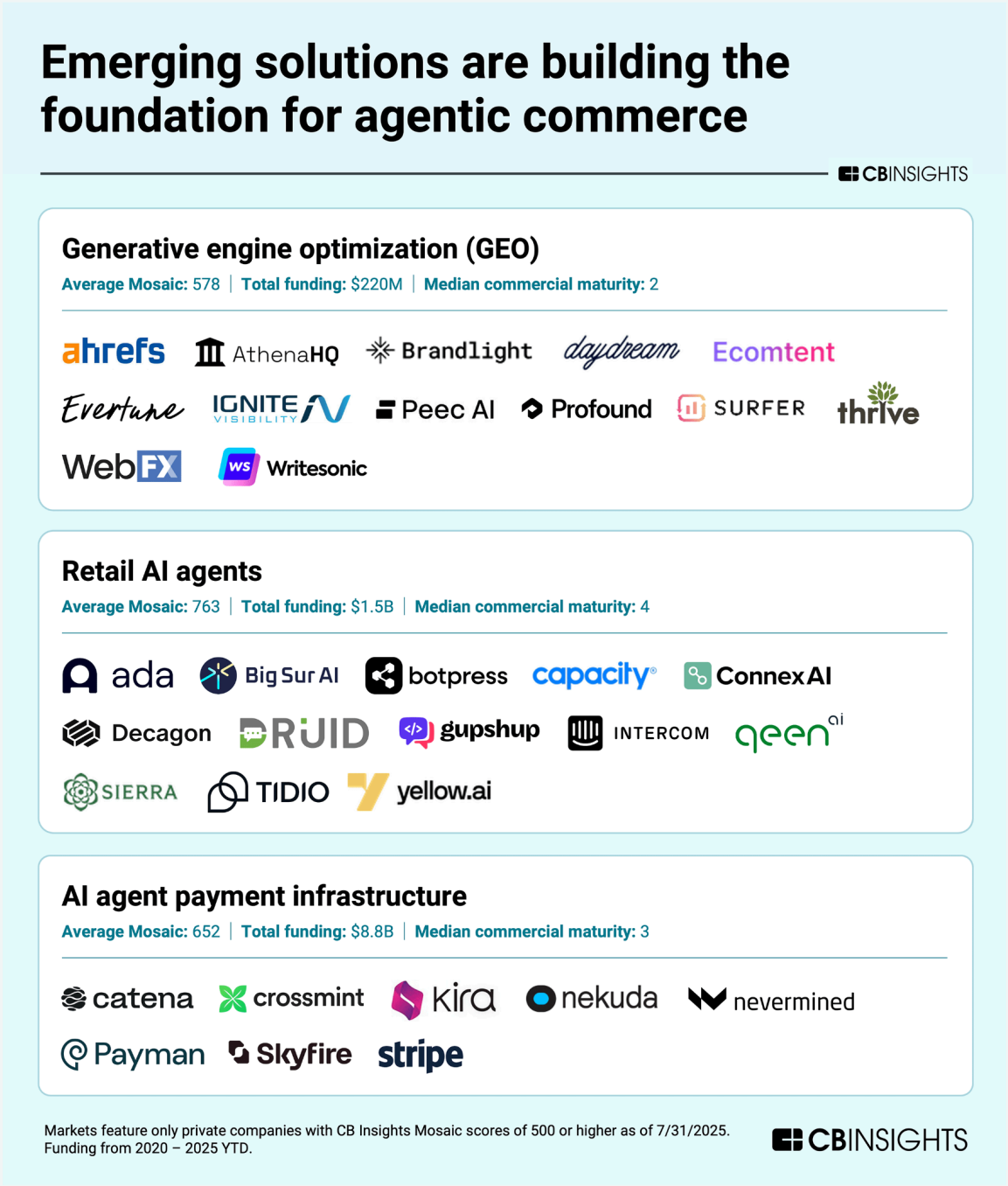
- **Generative engine optimization (GEO) is emerging as the next frontier in digital visibility.** As AI platforms become major gateways to online shopping, startups and established SEO firms alike are racing to help brands measure and boost their visibility in AI-generated responses. Nearly 40% of players in this market have launched in the last 2 years, and early leaders like Profound are gaining traction from major investors and brands. AI summaries on Google are already leading to declines in clicks across industries – as the trend accelerates and more consumers shop directly on AI platforms, GEO will become essential to shaping AI-driven commerce.
- **Retail AI agents are evolving beyond chatbots — and setting the stage for fully autonomous shopping.** Retail AI agents are expanding beyond customer support to play a larger role in personalization and shopper engagement. Startups building purpose-built agents for retail and consumer use cases are seeing the fastest growth in CB Insights Mosaic scores, which measure company health and potential. For instance, the Mosaic score for Decagon, which serves brands including Gopuff and Curology, climbed 33% in the last year. Early adoption will give merchants a strategic edge: introducing and training agents now can build consumer trust, gather data, and prepare retailers' operations for the leap to autonomous commerce.
- **Agentic payments are quickly moving from concept to infrastructure.** Almost half of the companies in the growing AI agent payments infrastructure market have partnered with or raised money from established payment leaders. While the startups developing solutions for AI payment rails have the agility and technical infrastructure that the incumbents need, the major players offer the early-stage companies long-built customer trust. Effective and multi-pronged partnerships will be essential as these tools develop.

Agentic shopping is the next big opportunity in commerce. Tech and payments leaders are already betting on the shift to AI-driven interfaces. But a growing wave of startups is also emerging, developing the building blocks for fully autonomous shopping.

Investors, merchants, and brands can seize this opportunity now, targeting the early movers for investment or partnership ahead of agentic shopping's arrival.

We've been tracking these emerging solutions on our agentic commerce Watchlist. Within this list, we've identified 3 breakout markets, each accelerating a different piece of the agent-led shopping journey.

Using CB Insights' Business Graph, which connects data across private investments, partnerships, and public company moves, we surfaced the key signals showing how startups and incumbents are building the infrastructure for agentic commerce.



Generative engine optimization (GEO) is emerging as the next frontier in digital visibility.

As AI platforms like ChatGPT, Claude, and Google AI Overviews become entry points for commerce, a new ad model is taking shape. Adobe reported a 1300% year-over-year spike in holiday traffic to US e-commerce sites from generative AI platforms — a trend that's only accelerating. With OpenAI testing native checkout in partnership with Shopify, the urgency for brands to understand and shape their presence on these platforms is growing fast.

A wave of GEO-focused startups is racing to define the space. Nearly 40% of players in this market have launched in the last two years, including 4 of the 5 companies with the highest CB Insights Mosaic score, which measures company health and momentum. The companies are building tools to monitor brand visibility in AI-generated answers and optimize positioning across AI interfaces.

GEO companies have built momentum quickly

Top GEO companies by Mosaic

Company	Mosaic	Founded Year
Profound	801	2024
daydream	683	2023
Peec AI	665	2025
Evertune	658	2024
Ecomtent	641	2022

Data as of 7/31/2025.
CB Insights Mosaic scores measure private tech company health and growth potential on a 0-1,000 scale.

CBINSIGHTS

Profound, ranked at the top for Mosaic, is setting the pace. Since August 2024, it's raised 2 early-stage rounds and landed clients like Indeed, U.S. Bank, and Mejuri. Its platform helps brands audit how AI sees their sites, understand what users are asking, and increase visibility through features like ChatGPT Shopping.

Traditional SEO firms and marketing agencies are also taking notice. Players like Surfer and Thrive Internet Marketing are adding GEO capabilities to stay competitive, signaling a coming convergence between established players and AI-native upstarts.

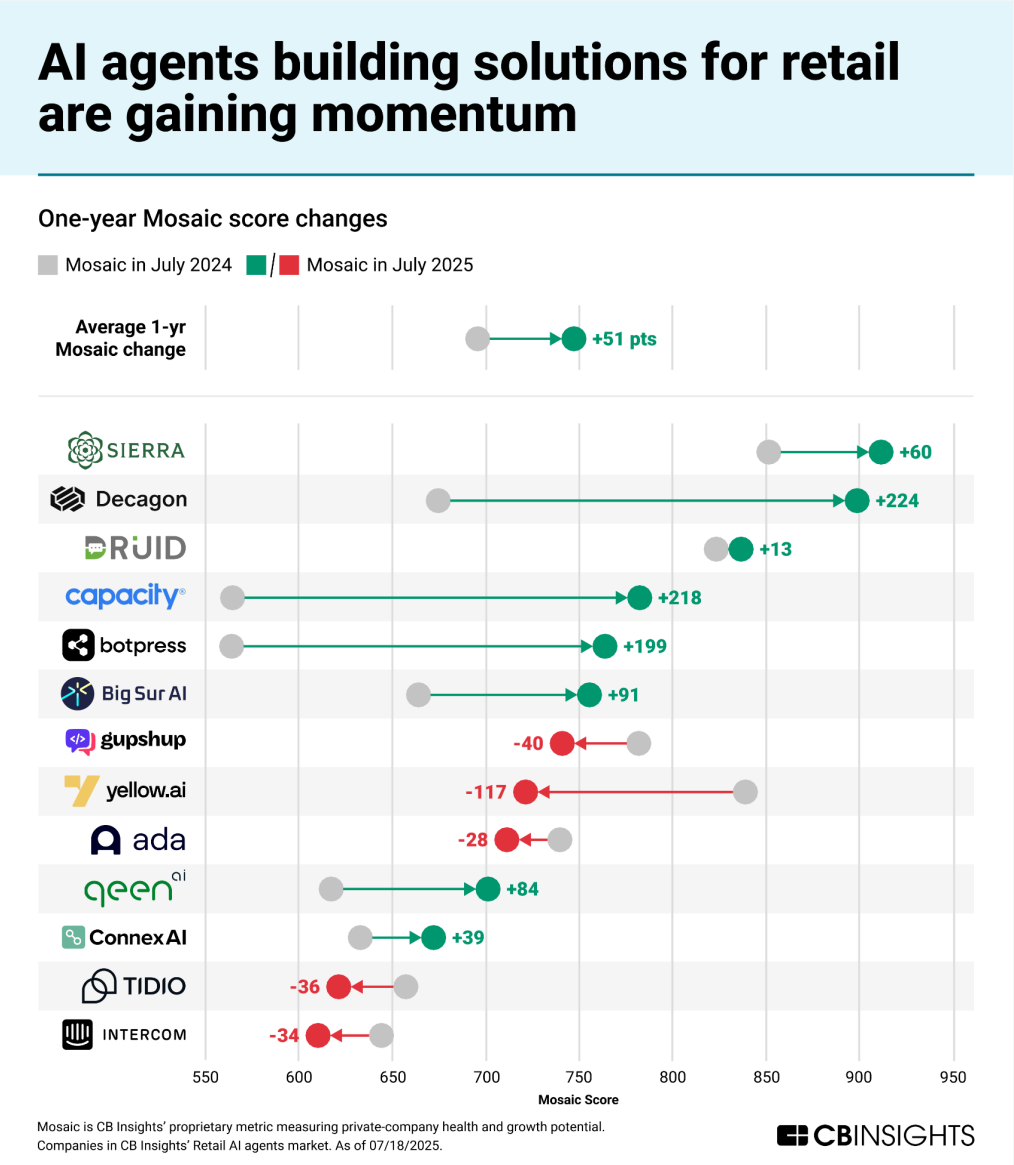
With more AI platforms enabling purchases directly through their interfaces, GEO is becoming a critical layer in digital commerce strategy. Merchants that build GEO muscle now will be better positioned to influence the AI-driven journeys of tomorrow.

Retail AI agents are evolving beyond chatbots — and setting the stage for fully autonomous shopping.

A new wave of AI agent companies is building tools that integrate directly into merchants' e-commerce infrastructure. These agents don't yet handle checkout, but they're already managing product recommendations, support requests, order tracking, and customer preferences — expanding their role in shaping the digital shopping experience.

Enterprise platforms like Sierra and DRUID are moving into the space, but it's the rise of AI startups developing applications specifically for retail that are accelerating adoption in the industry. These early movers are helping retailers gather critical consumer data and refine agentic interactions long before agents begin transacting on behalf of shoppers.

Startups building purpose-built agents for retail and consumer use cases are seeing the fastest growth in CB Insights Mosaic scores, which measure company health and potential.



Decagon's Mosaic score climbed 33% in the last year following its $131M Series C round in June 2025, which pushed the company to unicorn status. Its agent helps consumers manage billing, preferences, and support across brands like Gopuff, Oura, Curology, Hertz, and Bilt.

Big Sur AI, whose score rose 14% since 2024, is extending its product suite from AI sales agents to include content marketing and data science capabilities. Its tools, available through Google Cloud Marketplace, are already helping clients like Rad Power Bikes and Wyze improve product discovery and merchandising.

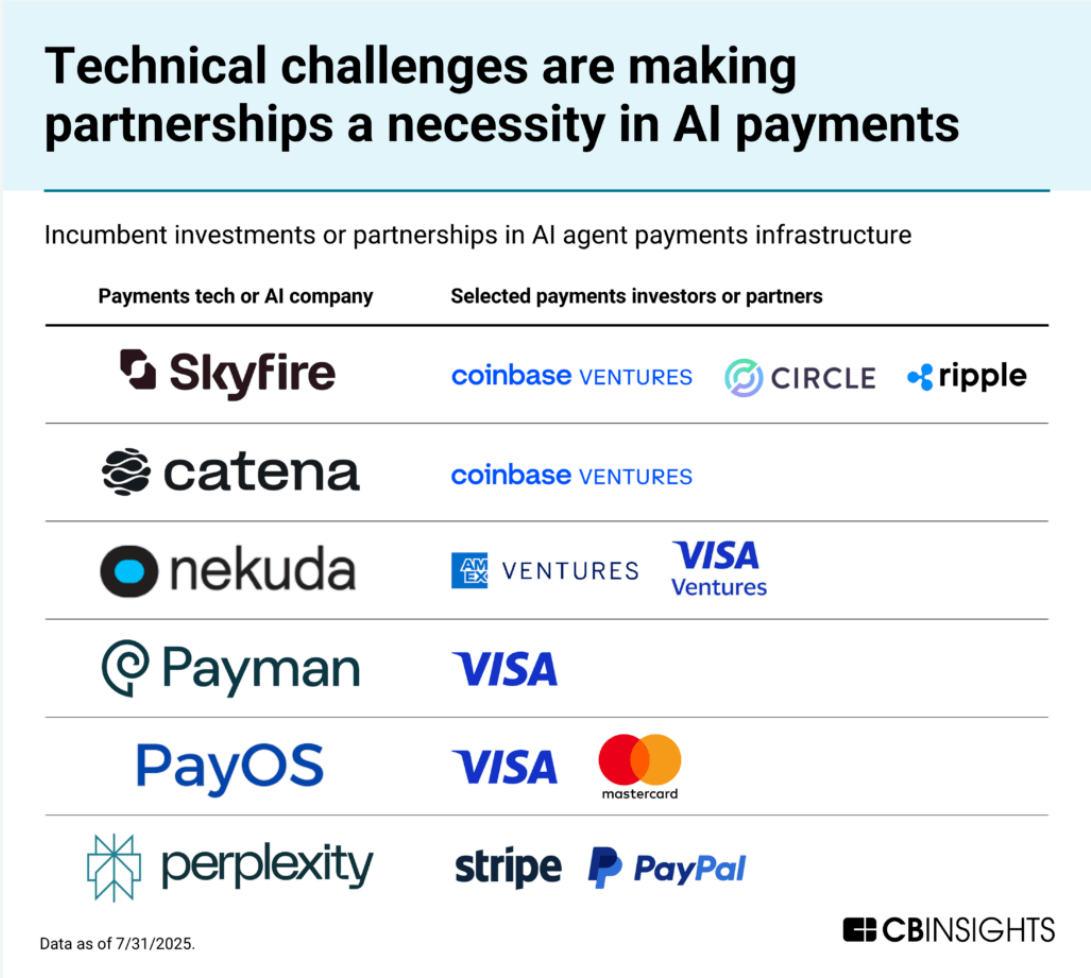
As agent-based tools scale across commerce, early deployment offers a critical advantage. By introducing agents now, retailers can build consumer trust, gather behavioral data, and prepare their systems for the next leap — from support to autonomous commerce.

Agentic payments are quickly moving from concept to infrastructure.

One of the biggest barriers to fully autonomous shopping is enabling secure, real-time transactions. A new class of startups is tackling this challenge head-on, building AI-native payment rails and digital wallets that let users authorize — and limit — spending by AI agents.

But solving agentic payments goes beyond infrastructure. It requires seamless authentication, fraud detection, and above all, consumer trust that these agents can transact safely on their behalf.

These hurdles have accelerated collaboration between fintech incumbents and startups. Strategic investments and partnerships are laying the foundation for the next phase of agentic commerce.



Skyfire, the second-highest ranked company by Mosaic score, raised $1M in October 2024 from Coinbase Ventures, with prior backing from Circle and Ripple. Its platform gives AI agents dedicated digital wallets, allowing users to preload funds for controlled purchasing.

Nekuda, another early mover, raised a $5M seed round in May 2025 from American Express Ventures and Visa Ventures. The company's tools include secure agent wallets and programmable guardrails that let users set spending limits and authentication protocols.

Incumbents are also building their own capabilities. In April 2025, Mastercard and Visa revealed they were developing agentic payment systems — coinciding with the launch of PayOS, a startup partnering with both companies. While no public solutions have launched yet, the timing signals a major push into AI-native commerce, despite the technical and regulatory complexity.

Perplexity's approach points to what's next. The company launched its agentic commerce feature, Buy with Pro, in late 2024 with support from Shopify and Stripe's one-time virtual card system. In 2025, Perplexity expanded with PayPal, allowing in-platform purchases via Venmo and PayPal's secure tokenized wallet, and began working with embedded checkout provider Firmly.

As fintech giants, AI startups, and commerce platforms converge on the agentic payments challenge, early collaboration is shaping how future transactions will happen: securely, autonomously, and inside the AI interfaces consumers already trust.

The industrial AI agents & copilots market map

We map out 60+ companies developing AI agents & copilots for industrial applications, including manufacturing, logistics, semiconductors, defense, and construction.

Published 12/23/2024 — Link to report

From early-stage startups to established firms, companies are racing to develop AI agents & copilots across the industrials sector.

While AI copilots — which work alongside humans to speed up their workflows — currently comprise 90% of company activity, the tech will serve as a stepping stone to more autonomous solutions in the coming years. Eventually, AI agents could manage entire industrial processes, shifting human roles from operational tasks to strategic oversight.

KEY TAKEAWAYS

- **75% of deals in the space go to early-stage startups**, suggesting high development potential in the years ahead, especially as copilots gain more agent-like capabilities. The nascent space has no clear leaders yet, signaling a window of opportunity for industrials players to gain an early edge.
- **Manufacturing is seeing the most activity compared to verticals like defense, warehousing, and semiconductors**. The most crowded markets are manufacturing optimization and digital work instructions, where vendors are building copilots on top of existing technologies like analytics platforms and augmented reality headsets. Improving on current workflows with a copilot likely offers the quickest path to commercialization vs. developing autonomous agents as standalone products.
- **Big tech is laying the groundwork for industrial AI agents & copilots**, with Microsoft forming partnerships with all 14 of the public industrials companies in this map. Google has also inked notable partnerships here. By owning the infrastructure, these tech giants will control how AI develops across industries, creating both dependencies and efficiencies for businesses adopting their solutions.

To help industry leaders understand the current landscape and identify key players, we mapped 68 AI agent & copilot developers across 12 markets.

The industrial AI agents & copilots market map

Industrial operations

Manufacturing optimization AI agents

NVIDIA, phaidra, Composabl, Intelecy, FLUIX, Palantir

Manufacturing optimization AI copilots

Diabatix, SIGHT MACHINE, sas, Rockwell Automation, SIEMENS, Basetwo, AUGURY, TULIP, Palantir, COGNITE, invisible ai, SymphonyAI, softserve, C3.ai, VECTRA, Zerynth, FLUIX, AISERA

Digital work instructions AI copilots

squint, GRIDRASTER, TULIP, LIBRESTREAM, AITOMATIC, softserve, ATHEER, SIEMENS, ptc, SPREAD, SymphonyAI, Augmentir, TAQTILE, nflux, arvizio, AISERA, Retrocausal, realwear, Zerynth

Equipment operation & control AI copilots

FLUIX, Composabl, AITOMATIC

Industrial safety & compliance AI copilots

Augmentir, invisible ai, AISERA, UpCodes, Protex AI

Supply chain

Supply chain AI agents

lyzr, DIDERO, Nanonets, Palantir, SAP

Supply chain AI copilots

Microsoft, wipro, ORACLE, Leverage, syrup, IFS, SAP, AISERA, PEK AIR

Warehouse AI copilots

sas, spot.ai, Palantir, Protex AI

Product design

Industrial design & modeling AI copilots

Leo, ZENODE, MONOLITH, SPREAD, Basetwo, nTop, SIEMENS, SANDVIK

Semiconductor design AI copilots

SILIMATE, Astrus, Flux, cadence, SIEMENS, SYNOPSYS

Industry-specific

Defense & national security AI copilots

ANDESITE, Palantir, Retrocausal, C3.ai, AISERA

Construction AI copilots

Microsoft, mbue, CONSTRUCTABLE, UpCodes, Kwant, concrete.ai, TORIC, PROCORE, Palantir

CBINSIGHTS

> **In the future agents will work truly end-to-end: reading demand signals, engaging suppliers, negotiating, issuing/expediting POs, and closing the loop in the ERP.** Medium-term we'll see agent-to-agent handshakes for routine transactions; humans stay on exceptions, lead strategy and focus on important human vendor relationships. The outcome is a procurement function that's finally living up to its true value potential with what we think will be very significant margin lifts for companies that deeply adopt AI.

Lorenz Pallhuber
Co-Founder, Didero

Early-stage startups dominate deal share while major corporations battle across tech verticals

Early-stage startups are highly active in this area: 75% of deals since 2022 involve early-stage companies (seed/angel and Series A).

Meanwhile, established companies like Palantir, SAS, and Siemens have also developed solutions. For instance, Palantir's AIP, an AI platform with copilot and agent-building capabilities, caters to a diverse range of applications such as aerospace and automotive fleet maintenance, manufacturing production scheduling, grid planning, and construction procurement.

As AI agents become more common over the coming years, human roles will gradually transition from operational tasks to strategic oversight, ensuring efficient collaboration between advanced AI systems and human expertise.

Manufacturing is seeing the most activity compared to verticals like defense, warehousing, and semiconductors

AI copilots comprise 90% of companies here — with manufacturing solutions as the primary focus — yet copilots ultimately lay the groundwork for more autonomous solutions.

Copilots assist manufacturing workers in real time with tasks such as monitoring production processes, while in areas like defense and supply chain, they can support military mission

planning and optimize logistics routes. Autonomous agents, on the other hand, would be capable of performing tasks independently, such as controlling production processes, conducting surveillance missions, and managing entire supply chains.

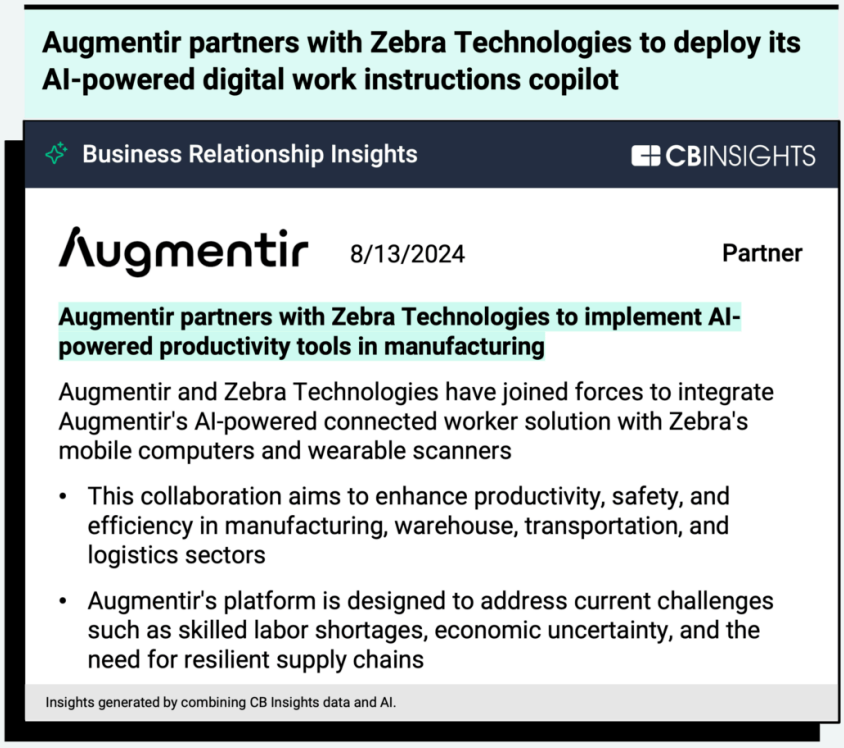
Developers are in the early stages of releasing agentic solutions, so industrial users should expect significant improvements over the coming years. Augmentir claims it was the first to release a genAI-powered frontline worker copilot in 2023, while Siemens released its Industrial Copilot for Engineering, in partnership with Microsoft Azure, in July 2024.

We highlight activity in key industrial sectors below, alongside select companies with commercial traction or unique product offerings.

Manufacturing

Initial manufacturing pilots of AI agents & copilots show commercial promise, with companies like C3 AI and SymphonyAI claiming a quick ROI in tasks like optimizing production schedules and quality control.

Meanwhile, startups such as Augmentir and nFlux focus on copilots that deliver digital work instructions. By incorporating AI capabilities, these copilots go beyond traditional augmented reality (AR) worker-assist technologies, personalizing and optimizing guidance in real time to individual worker performance and task requirements.



Source: CB Insights — Augmentir Business Relationship Insights

Supply chain

AI agents & copilots help synthesize data from tracking and logistics systems, improving supply chain resilience. For instance, early-stage startup Syrup Tech offers an analytics copilot for better demand forecasting and inventory management, while Leverage develops a copilot to extract insight from supply chain metrics.

Companies are also exploring warehouse copilots. LogistiVIEW, for one, has integrated smart glasses into its warehouse management systems, providing step-by-step instructions for warehouse workers in receiving, inventory, picking, and sortation.

Defense

Leading global defense agencies — including those from the US, the UK, Israel, South Korea, and China — are starting to implement AI across applications like unmanned vehicles, command and control centers, and training simulations.

For instance, the US Department of Defense views AI as critical for the future of warfare, using it to improve in areas like battlefield awareness and adaptive planning.

Potential applications for AI copilots involve interpreting surveillance and operational data, boosting readiness, and aiding military logistics.

Notably, Palantir and Microsoft recently linked up to combine their secure cloud, analytics, and AI tools in a joint offering for the US defense and intelligence community.

Construction

Construction projects are often large and complex, presenting numerous chances to mismanage costs, schedules, and resources.

Procore's copilot aims to combat this by assisting with various construction processes, including project submittals, specifications, and drawings. The company plans to release a platform for AI agents in 2025, which will automate routine tasks like quoting and scheduling.

Other startups take a more focused approach. For example, Concrete.ai developed its copilot to optimize the concrete recipe for various job site applications.

Companies are building solutions on top of big tech's foundation

Big tech companies like Microsoft and Google are laying the groundwork for industrial AI agents & copilots by offering the infrastructure through platforms like Microsoft Azure and Google Cloud.

In fact, all 14 public companies in this market map — such as PTC, SAS, and Siemens — have partnered with Microsoft to develop their copilots. Established startups such as Cognite and Tulip have done the same.

Companies are also developing AI solutions for internal operations. Honeywell, for instance, collaborated with Google Cloud to create AI agents that help technicians resolve maintenance issues and engineers find the information they need across databases.

The concentration around major tech platforms creates a double-edged sword: while it accelerates AI adoption through turnkey solutions, it also embeds companies deeply within these ecosystems.

As Microsoft and Google amass industrial data and insights, they're positioned to become indispensable to industrial operations, shifting power from traditional industrial leaders to the tech giants controlling the AI infrastructure.

> "
>
> **Autonomous agents are only as powerful as the data they run on.** The real barrier to adoption isn't the technology — it's unlocking and unifying high-quality data streams that allow agents to generate tangible value. Not more dashboards, but concrete recommendations and automated actions. **Manufacturers that solve this will gain a decisive edge, equipping themselves with always-on companions that analyze, recommend, and optimize at a scale no human team could ever achieve.**

Matthias auf der Mauer
Founder & Co-CEO, Juna.ai

100 real-world applications of genAI across financial services and insurance

We use CB Insights data to uncover and categorize 100 real-world genAI applications across financial services and insurance, as well as provide key takeaways on adoption.

Published 7/31/2025 — Link to report

GenAI adoption is increasingly measurable.

Many of the world's most influential financial services firms — like Allianz, J.P. Morgan, and Mastercard — have taken concrete action to adopt genAI technology.

The genAI adoption efforts have shaped 2 years' worth of corporate strategy, unveiling key priorities — from the rise of agentic commerce to customer service copilots — across the competitive landscape.

Using CB Insights data, we identified and analyzed 100 real-world applications of genAI from 69 companies across banking, insurance, and payments.

Dive deep into all 100 genAI applications

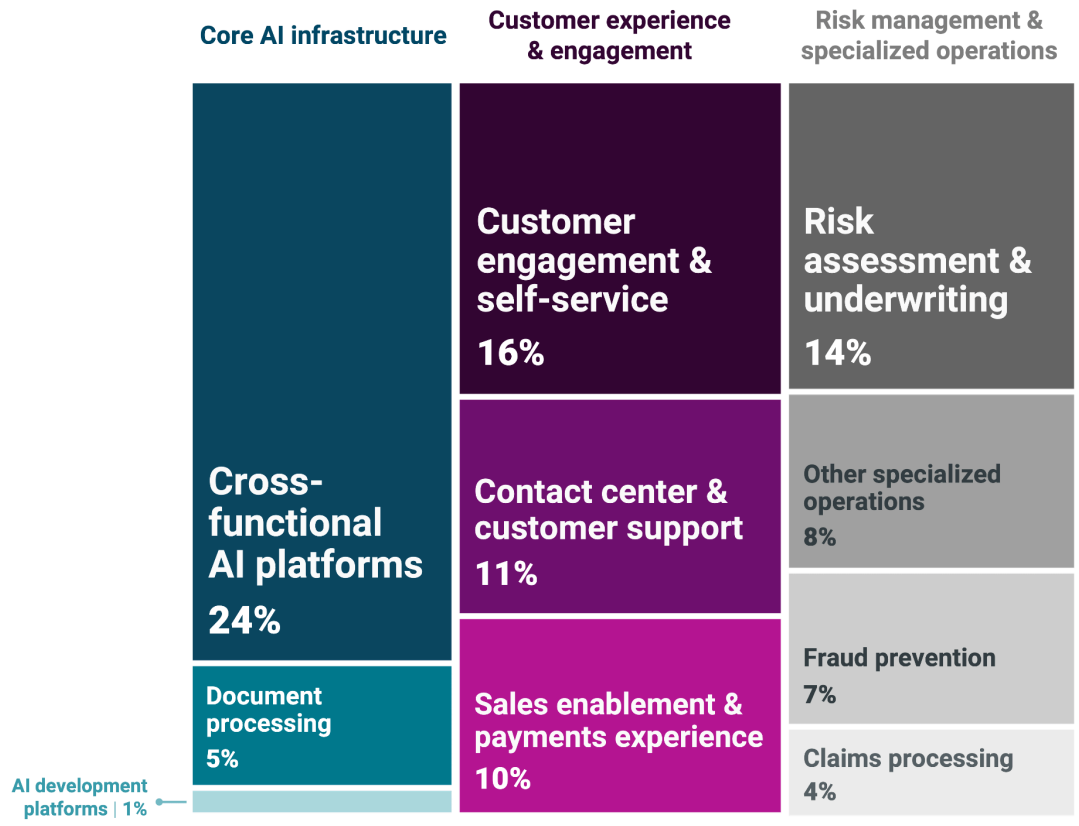
Methodology

We used CB Insights' Business Graph — including data points like Dealmaking, Business Relationships, Earnings Transcripts, and Media Mentions — and third-party company releases to identify 100 real-world genAI applications across banking, insurance, and payments. These applications were disclosed between July 2023 and April 2025.

Then, using CB Insights' Team of Agents, we analyzed these applications across 10 categories. Applications are detailed based on disclosure date, and are not exhaustive of a given company's genAI initiatives. Applications and categorizations are not mutually exclusive or exhaustive of activity within their respective industries.

Where financial services and insurance firms are applying genAI across their businesses

Categorization of 100 disclosed real-world genAI applications



Note: Data from applications mentioned, piloted, or deployed across the banking, insurance, and payments industries between July 2023 and April 2025. Applications categorized based on disclosure date. Applications and categorizations are not mutually exclusive or exhaustive of activity within their respective industries. GenAI application breakdown by industry: banking (42), insurance (39), and payments (19).

CBINSIGHTS

Key takeaways

1. Cross-functional platforms are now table stakes.
24% of applications center on deploying general-use genAI platforms to employees.

Prominent firms like BBVA have established enterprise-wide genAI capabilities across their organizations (typically via enterprise-wide deployments of platforms like Microsoft Copilot or ChatGPT). Early adopters — like Klarna, which shared in May 2024 that 87% of its employees are using OpenAI technology — now have over a year of genAI operational experience at scale, which can guide the development of more complex applications in the future.

Looking forward, financial services firms without a plan to provide genAI access to employees risk competitive disadvantage. Over the past 2 years, simply providing genAI capabilities to employees has shifted from cutting-edge innovation to standard operations.

2. Microsoft and OpenAI permeate the adoption landscape.
33% of applications analyzed disclose involvement from either Microsoft or OpenAI.

Microsoft and OpenAI (in which Microsoft has significantly invested) overwhelmingly permeate the landscape of genAI applications analyzed. Many of these applications anchor on foundational capabilities, from which organizations can build more complex applications and agents. Anthropic, Amazon Web Services, and Google Cloud follow a similar deployment pattern across multiple companies in the sector.

Looking forward, financial services firms should prepare for increasingly blurred "build, buy, or partner" decisions. The prevalence of genAI model developers (like OpenAI and Anthropic) and big tech partners (like Microsoft and Google) provide financial services executives with more flexibility to customize their tech solutions than what has traditionally been the case with many point-solution providers.

3. Emerging genAI vendors face a fierce competitive landscape.
Median Mosaic Scores among genAI startups analyzed are in the top 3% globally.

The 100 analyzed genAI applications include engagement from 25 startups as tech vendors, ranging from pre-seed companies like Twin — which offers an agent for invoice collection — to late-stage giants like Anthropic. These startups have a median CB Insights Mosaic Score — which measures the overall health and growth potential of private companies — of 732 out of 1,000, as of July 30, 2025.

Looking forward, financial services firms should prepare for increasingly capable tech vendors seeking to sell their genAI products. These vendors must exhibit a clear advantage over the alternative of building in-house solutions.

4. Customer-facing genAI will become increasingly prevalent.
16% of applications center on customer engagement & self-service capabilities.

Firms like ING, Wells Fargo, and Truist show that customer-facing genAI assistants are capable of powering millions of customer interactions.

Customer-facing genAI deployment will accelerate as companies like Mastercard, Visa, and PayPal deploy applications centered on "agentic commerce," where customers can autonomously shop and complete transactions with AI payments agents.

Looking forward, financial services firms need to develop a gameplan for how they will engage customers with agentic AI. The market opportunities for enterprise agents and copilots are growing, so customer-facing applications will quickly emerge.

5. Impact is now tangible, but success definitions remain elusive.
Only 30% of applications disclose quantitative tangible impact from deployment.

Most of the application sources analyzed lack disclosure of tangible impact (i.e., numbers, percentages, or figures to quantify effectiveness). Among the impact metrics that are available, the top-cited focus on operational considerations like call-handle times.

Looking forward, any financial services firm has the opportunity to define "what good genAI adoption looks like" across the sector. The lack of clear success definitions creates an opportunity for financial services firms to stand out among peers.

Dive deep into all 100 genAI applications

Additional resources

Research

The Future of Professional Services: How firms will capture value in the AI agent era

Enterprise AI agents & copilots: Our growth projections for the $5B+ market

What's next for AI agents? 4 trends to watch in 2025

Future of the workforce: How AI agents will transform enterprise workflows

Keep reading

Markets

Coding AI agents & copilots

AI agent development platforms

Legal AI agents & copilots

AI agent browser infrastructure

See 50+ more AI agent markets

Predictive intelligence

Mosaic Score: Private Company Ratings

Commercial Maturity

Exit Probability

Business Relationships

Access the most powerful market intelligence for 10 days – free of charge.

Start free trial