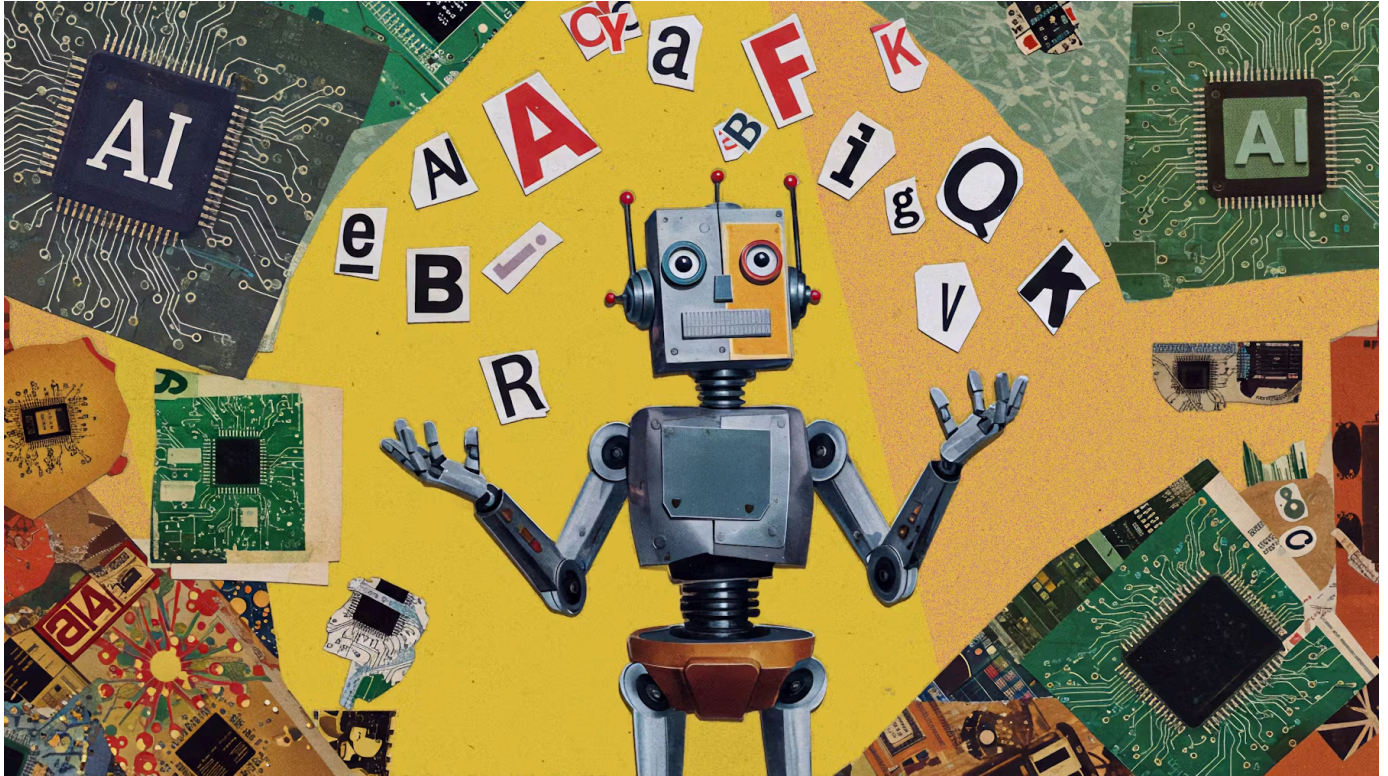


The Seven Kinds of AI Agents



Art via Clark Miller



By Aaron Holmes

Jul 7, 2025, 10:00am PDT

The term “agent” or “agentic” has become so ubiquitous in the business world that it can now seem to mean everything or nothing at all.

In its simplest form, an agent is artificial intelligence that handles multistep tasks without requiring a human to steer it the whole time. In other words, an agent goes beyond what the basic version of ChatGPT can do.

In the past year, AI firms have launched agents that appear in stand-alone apps, as sidebars or add-ons in existing apps, or as dialog boxes on webpages. As the table below shows, you're now likely to encounter seven types of AI agents in the wild as a consumer, where they may aim to resolve your customer support complaints without requiring a human employee, or in the workplace, where they handle tasks like processing human resources forms or resolving IT tickets.

Customers of AI agents for work tasks typically pay a premium. For instance, OpenAI's Codex coding agent, which can handle tasks like debugging an application without much human oversight, costs \$6 per 1 million output tokens, nearly four times more than paying to access a large language model like the kind that powers ChatGPT. (Tokens are words or parts of words.)

Similarly, Microsoft charges \$40 per month per person for access to similar coding agent features through GitHub, or four times more than the basic version that only does more rudimentary actions, like suggesting ways for developers to complete a line of code they're writing.



Agent Archetypes

Category	Definition	Products	Common Uses
Business-task agents	Can take actions in multiple enterprise software apps.	UiPath, Microsoft Power Automate, Zapier + AI	Invoice processing, data entry, document classification, scheduling
Conversational agents	Text or voice chatbots that resolve customer support tickets or employees' IT or HR questions through a back-and-forth exchange.	Google Dialogflow, Microsoft Dynamics 365, Amazon Lex, Salesforce Agentforce, ServiceNow NowAssist, Sierra, Decagon	Customer service, IT service tickets, HR tasks
Research agents	Retrieve, analyze, and validate information from trusted sources, including academic literature or web content.	OpenAI Deep Research, Perplexity Pro, Scite Assistant, Elicit, AlphaSense	Academic-style research, sourcing citations, answering hard technical questions
Analytics agents	Analyze structured data to produce graphics/charts or reports.	Power BI Copilot, Tellius, ThoughtSpot, Glean	Data querying, creating dashboards, business insight summaries
Developer agents	Assist software engineers by handling complex coding tasks	Cursor, Windsurf, GitHub Copilot, Q Developer, Claude Code, Cognition's Devin, Cline	Code completion, documentation, debugging, refactoring, site reliability engineering
Domain-specific agents	Specialized agents built with domain knowledge for high-stakes or regulated fields such as law, medicine, or finance	Harvey (legal), Hippocratic AI (healthcare), Vic.ai, Rogo and Hebbia (finance)	Contract analysis, medical triage, financial analysis

AI firms hope agents will fuel another revenue boom. OpenAI, which has considered charging customers as much as \$20,000 per month for its most advanced doctorate-level research agents, has forecast that its revenue from agents will balloon to \$29 billion in 2029, up from \$3 billion this year. And Microsoft recently reorganized its business units that sell AI tools to developers to maximize sales of agents.

What sets AI agents apart from more rudimentary chatbots is their ability to take multiple actions, sometimes by connecting to various applications like Salesforce CRM, coding software Jira or repositories such as GitHub to complete tasks.

A new technology is making it easier for agents to access multiple applications to solve problems. For example, customers of Cursor, which has a coding agent feature, might want to reference information about software bugs from Jira, which tracks common bugs. Anthropic developed open source software called Model Context Protocol to enable AI models to access this information.

Article continues below

The Information Events

AI Agenda Live NYC: The Next Wave

Tue., September 30, 2025

New York, NY

[Learn more](#)

AI agents have evolved quickly in the past two years, though they can still make mistakes and require some human oversight. Agents often combine LLMs with reasoning models or other types of planning components that figure out what steps to take, and in what order, to complete a task. They usually include memory systems so the agent can better remember prior interactions and a customer's preferences. And they usually require so-called orchestrator software that coordinates all of these technologies.

The agents can now use MCP or similar open-source protocols to connect to and use other applications or to tap information from thousands of apps like Slack or Jira.

Carvana's Coder

For example, Carvana, an online car retailer, developed an agent that employees can instruct to build new webpages by writing in plain English what they want to see, said Alex Devkar, a senior vice president.

The agent is based on software from Cline, an AI code generator, and incorporates OpenAI's o3 model and Anthropic's Claude Sonnet 4 model to make plans and generate code, respectively. The agent also uses MCP to connect to other applications, including Jira and GitHub, to draw on the company's code base and ensure the webpage it's creating has code and design styles consistent with existing pages.

The agent then uses MCP to connect to design app Figma to create a new webpage, Devkar said.

Engineering teams at Carvana started working on the tool late last year and have been tweaking it on a weekly basis by swapping out different models and connected applications to see what works best, Devkar said. He declined to comment on the cost of running the tool but said it is "significantly" less than paying people to do the same thing.

"I view this as a sea change, and our goal as an organization is to keep expanding these capabilities," he said.



Lavingia said the tools have helped the firm reduce the number of software engineers on staff, saving hundreds of thousands of dollars annually

Agents with more advanced capabilities are at a nascent stage. Microsoft has let customers test a site reliability engineer agent powered by its GitHub Copilot coding software. The agent can keep tabs on a customer's webpages around the clock, notice when a website goes down or is acting buggy, and attempt to figure out what may be causing the outage and apply different patches to restore the site, according to a demonstration from Microsoft. The company has not disclosed how much the agent will cost.

Goldman Sachs has tested the SRE agent tool, according to someone familiar with the company's tests.

Microsoft and Salesforce are trying to convince businesses they can develop their own agents, similar to Carvana's. Salesforce's Agentforce and Microsoft's Copilot Studio, for example, let companies create an agent to, say, resolve tech support issues on their own, such as helping an employee reset their password. Workers can also configure the agent so employees can access it through chat apps like Slack or Teams.

Besides holding a conversation, many agents take actions like searching for information in a company's database or sending an email on a customer's behalf. For example, Gmail users can now schedule a meeting with colleagues by giving Gemini a time and list of attendees, and the AI will automatically create a Google Calendar event and email the attendees.

A major part of those efforts involves creating agents that can work for a long time on a single task. Given a simple description of a task, OpenAI's Codex agent can work for up to 30 minutes without human supervision on a coding problem. Anthropic said its latest model, Claude 4, was able to work on a coding problem for up to seven hours.

Developers say the results aren't perfect, but reviewing the agent's work and making small changes to it still saves time compared to coding from scratch.

High Costs, High Savings?

Reasoning models like OpenAI's o3 and o4-mini have gotten cheaper in recent months, helping companies improve the quality of agents they develop.

Businesses say the higher costs of using agents can be worth it.

For instance, online marketplace startup Gumroad has been spending several thousand dollars a month on coding agents including Cursor and Devin that can program applications with minimal oversight, according to CEO Sahil Lavingia—a considerable increase from the several hundred dollars per month the 25-person company had been spending on basic subscriptions to ChatGPT. But Lavingia said the tools have helped the firm reduce the number of software engineers on staff, saving hundreds of thousands of dollars annually.

Similarly, professional services giant EY developed an agent to automate compliance audits of customers' IT systems, according to principal Sameer Gupta. The agent, powered by models from OpenAI and tools from Microsoft's Azure, connects to various data sources from a client's IT systems, as well as documents such as contracts, to compile the data in a single document with suggestions for how the customer can reduce risk.

The cost of running the agent is “a significant step up” compared to rudimentary chatbots, Gupta said, but he estimates it has saved EY between 5% to 20% compared to paying people to handle the risk analysis manually.

Jon Victor also contributed to this article.

Aaron Holmes is a reporter covering tech with a focus on enterprise and cybersecurity. You can reach him at aaron@theinformation.com or on Signal at 706-347-1880.