

■ A GUIDE FOR AI FOUNDERS AND GTM LEADERS

Rethinking enterprise growth for agentic products

How AI agent companies get through enterprise review, price contracts that finance clears, and grow accounts that renew.



Contents

START HERE

Before you start 3

Why this ebook exists 3

About the data 4

PART 1

The new enterprise buyer 6

01 Who buys agents, and from which budget 7

02 From pilot to production 11

PART 2

Winning the enterprise deal 15

03 Clearing the trust review 16

04 Pricing and contracting enterprise agent deals 20

PART 3

Growing inside the account 24

05 How agent accounts expand 25

06 Choosing the go-to-market motion 28

07 Proving value after the signature 31

LOOKING AHEAD

What changes next 34

08 What comes next for enterprise agents 34

BACK MATTER

Reference 36

Benchmarks at a glance 36

About Clocknext 38

Why this ebook exists

Enterprise buyers want AI agents. Getting from a first pilot to a large account that renews still takes most agent companies longer than they expect.

The demand is real. Menlo Ventures estimates that enterprises spent \$37 billion on generative AI in 2025, more than three times the year before, and most of the organizations McKinsey surveyed in 2025 were at least experimenting with agents.

Far fewer have scaled them. In the same McKinsey survey, 23% of organizations were scaling agents anywhere in the business. Gartner predicts that more than 40% of agentic AI projects will be canceled by the end of 2027, citing rising costs, unclear value, and weak risk controls.

Enterprise growth for agent companies is decided in the distance between those numbers. The vendors that grow inside large accounts tend to have three things: an agent the buyer's security and legal teams will let act on their systems, a contract whose cost finance can predict, and value reporting that holds up at renewal.

This ebook is for founders and go-to-market leaders at AI agent companies that sell to enterprises. It follows an account from the first budget conversation to renewal.

Six things to take away

1. **Agents are bought from core budgets now.** Innovation budgets fund a small share of AI spending, so agent deals go through full procurement, security, and legal review.
2. **Pilots convert more often than the headlines suggest.** The harder step is a result the buyer can see in its P&L. Plan for that step from the first meeting.
3. **Trust is part of the product.** Permissions, audit records, and evaluation evidence decide how much an agent is allowed to do, and so how much it can earn.
4. **Enterprise contracts combine a commitment with usage.** A committed floor, metered or outcome-based parts, overage rules, and spend controls give buyers predictability and give you room to grow.

5. **Accounts expand along four lines.** More volume, more workflows, more departments, and more autonomy. Each needs its own contract terms and its own margin check.
6. **Value reporting decides renewal.** A monthly report that your sponsor can forward to their own leadership is one of the most useful things you can build.

How to read it

Part 1 covers the buyer and the path from pilot to production. Part 2 covers the trust review and the contract. Part 3 covers growth after signature, and Chapter 8 looks ahead. If you have a large deal in review, start with Part 2. If your enterprise customers are not expanding, start with Part 3.

About the data

This is a young market and the reliable data is thin. Investor research from Menlo Ventures, a16z, Bessemer, ICONIQ, and Emergence is useful, and its authors invest in some of the companies they describe. Gartner figures are often predictions, and we label them. Benchmark results depend on who is in the sample, so we show the sample next to the number. Company figures other than filings are self-reported, so we use company examples mainly for how they price.

NOTE

Data in this ebook is current as of October 2026. Benchmarks for these metrics may have changed slightly by the time you read this guide.

The vocabulary we use

TERM	MEANING
Commitment	The amount a customer agrees to spend over a term, drawn down as they use the product.
Drawdown	How usage or outcomes are subtracted from a commitment, often from prepaid credits held in a wallet.
True-up	A charge at the end of a period when usage falls short of the commitment.
Outcome	A completed business result you charge for, such as a resolved ticket.
Signal	A metered event from an agent, such as a model call or a completed step, tagged with its customer and its provider cost.
Provider cost	What the model or tool provider charged you for the work behind a signal.
Margin	What you bill minus provider cost, measured per customer, agent, or model.
NRR and GRR	Net and gross revenue retention, with and without expansion.

The new enterprise buyer

Agent purchases now come from core budgets and pass through full review. This part covers who buys, where the money comes from, and how pilots turn into production.

01 Who buys agents, and from which budget

02 From pilot to production



Who buys agents, and from which budget

Enterprise AI spending has moved into core budgets. That brings more money, more process, and more people who can stop a deal.

Enterprise spending on generative AI has grown quickly. Menlo Ventures estimates it at \$1.7 billion in 2023, \$11.5 billion in 2024, and \$37 billion in 2025, with about \$19 billion of the 2025 total going to applications. Where the money comes from has changed as much as the total. In a16z's 2025 survey of about 100 enterprise CIOs, innovation budgets had funded about a quarter of spending on large language models the year before. By 2025 they funded 7%. The rest came from central IT and business-unit budgets, and CIOs expected their AI budgets to grow by about 75% over the following year.

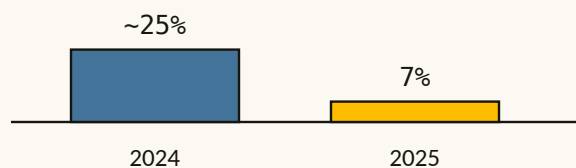
Money from core budgets arrives with core-budget process. a16z reports that enterprise AI purchases now go through the same procurement discipline as other software: security questionnaires, legal review, vendor onboarding, and a budget owner who has to defend the line item again next year.

FIGURE 1.1

AI spending has moved into core budgets, and enterprises buy more than they build

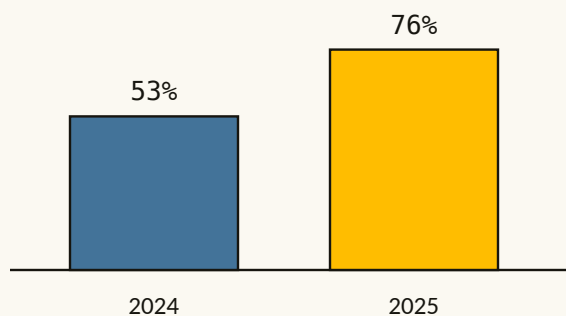
Share of LLM spend from
innovation budgets

A16Z CIO SURVEY



Share of AI use cases
bought rather than built

MENLO VENTURES



Sources: a16z, How 100 Enterprise CIOs Are Building and Buying Gen AI in 2025 (June 2025), about 100 CIOs; Menlo Ventures, 2025: The State of Generative AI in the Enterprise (December 2025), about 500 US decision-makers. Both firms invest in companies in this market.

Enterprises also buy more than they build. Menlo found that 76% of enterprise AI use cases in 2025 were purchased, up from 53% in 2024. Many purchases start small: Menlo estimates that 27% of AI application spending comes through product-led motions, against 7% for software overall. The central contract that follows still has to be sold to a committee.

The budget you compete with

Agents often do work that the enterprise pays people or outsourcing firms to do today, so the budget they draw on, and the comparison the buyer makes, can move from software to labor or services. No independent dataset yet shows how far this has gone, so ask each buyer where the money will come from.

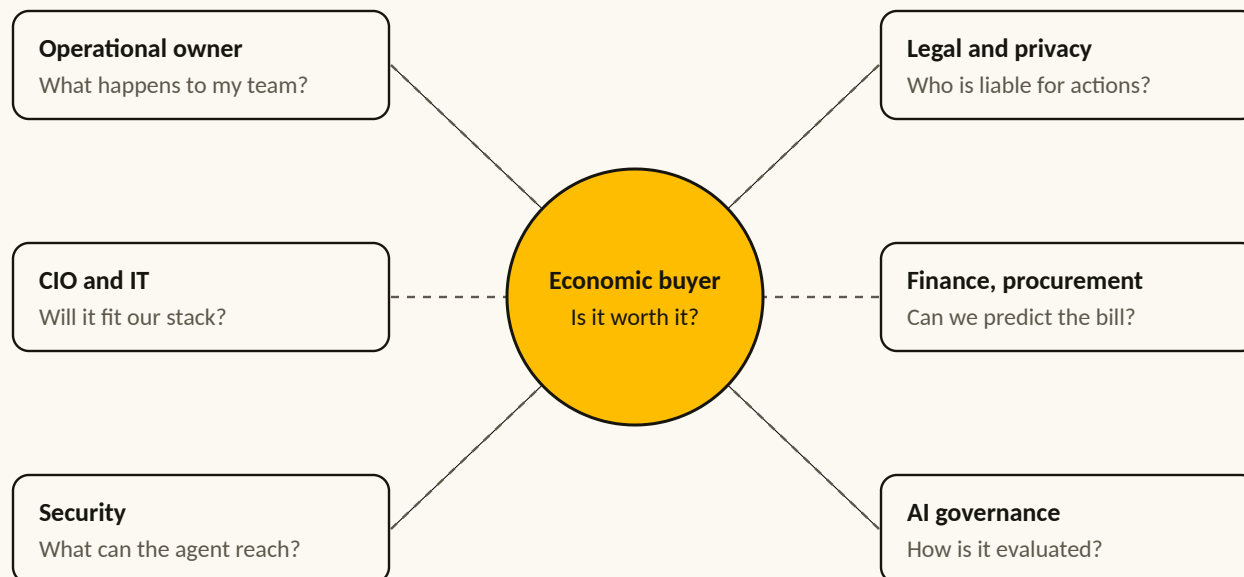
BUDGET SOURCE	WHAT THE BUYER COMPARES YOU WITH	PRICING IT TENDS TO SUPPORT
Software	Other tools, usually priced per user	Seats or a platform fee, with usage on top
Headcount	Fully loaded cost per task or per person	Per task or per outcome, against a cost baseline
Outsourced services	Per-ticket or per-transaction vendor rates	Outcome pricing, which this buyer already understands
Cloud commitment	Money already committed to a cloud provider	A marketplace purchase that draws down that commitment (Chapter 6)

More people can say no

An agent reads data and takes actions, so more functions have a say in the purchase than in a typical software deal. The economic buyer decides whether it is worth the money. Six other groups can slow the deal or stop it.

FIGURE 1.2

The agent buying committee



The first question each group tends to ask. Your champion is usually the economic buyer or the operational owner. The other groups need their own material, usually earlier than you expect.

■ IN PRACTICE

In the first meeting, ask who else will review the purchase and what each of them will need. Send the security and legal material in week one, before anyone asks for it.

SOURCES FOR THIS CHAPTER

1. Menlo Ventures, 2025: *The State of Generative AI in the Enterprise* (December 2025), menlovc.com
2. a16z, *How 100 Enterprise CIOs Are Building and Buying Gen AI in 2025* (June 2025), a16z.com

From pilot to production

Buyers do move agents into production. The harder step is a result their finance team can see.

Few topics in enterprise AI produce more conflicting statistics than pilot conversion. Four widely quoted numbers point in different directions, mostly because each one measures something different.

FIGURE 2.1

Four numbers, four questions

MENLO VENTURES • DEC 2025

47%

of AI deals reach production, against 25% for traditional software.

Measures: deals where a company has already chosen to explore a specific purchased product.

MCKINSEY • NOV 2025

23%

of organizations are scaling agents somewhere in the business.

Measures: organization-wide adoption. 62% are at least experimenting.

MIT NANDA • AUG 2025

95%

of generative AI pilots showed no measurable P&L impact.

Measures: all initiatives, internal builds included. Small sample and a narrow definition of impact, widely debated.

GARTNER • JUN 2025

40%+

of agentic AI projects will be canceled by the end of 2027.

A prediction, supported by a poll of webinar attendees.

Sources: Menlo Ventures (December 2025); McKinsey, The State of AI in 2025 (November 2025), 1,993 respondents; MIT NANDA, The GenAI Divide (2025); Gartner (25 June 2025).

Read together, they tell a consistent story. Enterprises that have chosen a product usually get it into production, and purchased products do better than internal builds. In a Fortune interview about the MIT study, its lead author said that buying from specialized vendors and building partnerships succeeded about 67% of the time, while internal builds succeeded about a third as often.

The step that fails more often comes after deployment. In McKinsey's survey, 39% of organizations reported any effect on EBIT from AI, and most of those put it below 5%. A pilot that works technically but never shows up in the buyer's numbers is at risk at the first budget review.

Why agent projects stall

Gartner names three causes behind its cancellation prediction: escalating costs, unclear business value, and inadequate risk controls. Each maps to something the vendor controls: a bill the buyer can forecast and cap (Chapter 4), a baseline with regular reporting (Chapter 7), and permissions with audit records (Chapter 3).

Buyers will also test reliability. In ICONIQ's July 2026 survey of AI builders, nearly half said their own internal agents need a person to step in on at least 30% of tasks, and the most common failure was a multi-step workflow breaking partway through. Expect enterprise buyers to look for the same failure in your pilot.

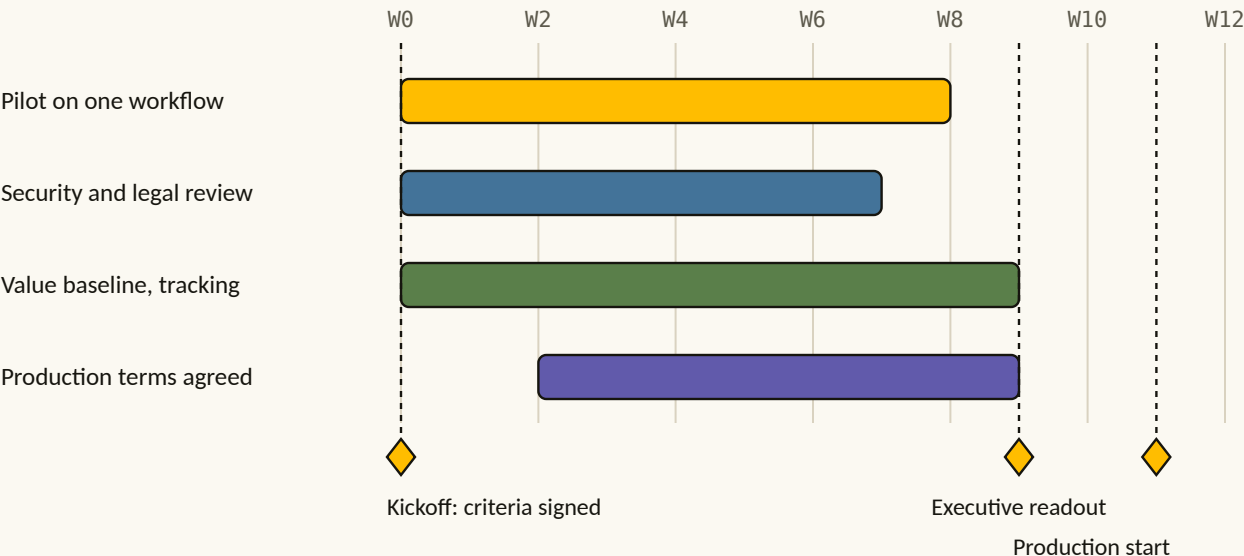
A pilot designed to convert

SEVEN PARTS OF A PILOT THAT CONVERTS

- 01 **One workflow, one owner, one metric.** Record the current cost and quality of the unit of work before you start.
- 02 **Success criteria in writing.** Thresholds, method, and who measures, in the pilot agreement.
- 03 **Conversion terms agreed.** The production price, commitment, and start date if the criteria are met.
- 04 **Security review in parallel.** Start the trust review in week one (Chapter 3).
- 05 **Your cost measured from day one.** Know your provider cost per unit before you quote production pricing.
- 06 **A defined escalation path.** What the agent hands to people, and whether that work is billable.
- 07 **A readout on the calendar.** Booked at kickoff, with the economic buyer in the room.

FIGURE 2.2

A pilot plan with parallel tracks



Conceptual, in weeks. The trust review, commercial terms, and value baseline run alongside the pilot, so a successful result can move straight to production.

SOURCES FOR THIS CHAPTER

1. Menlo Ventures, *2025: The State of Generative AI in the Enterprise* (December 2025), menlovc.com
2. McKinsey, *The State of AI in 2025* (November 2025), mckinsey.com
3. MIT NANDA, *The GenAI Divide: State of AI in Business 2025* (2025), and Fortune coverage of the report (18 August 2025), fortune.com
4. Gartner, *Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027* (25 June 2025), gartner.com
5. ICONIQ, *State of AI: The Builder's Economy* (July 2026), iconiq.com

Winning the enterprise deal

Two reviews sit in the middle of every agent deal: whether the buyer can trust the agent, and whether finance can approve the contract. This part covers both.

03 Clearing the trust review

04 Pricing and contracting enterprise agent deals



Clearing the trust review

An agent acts on the buyer's systems. Security, legal, and governance teams will review what it is allowed to do, as well as what data it can see.

Security is the barrier enterprise leaders mention most. In KPMG's AI Quarterly Pulse for the fourth quarter of 2025, 80% named cybersecurity as the biggest barrier to their AI strategy, up from 68% at the start of the year. By the first quarter of 2026, 63% of organizations required people to verify agent outputs, up from 22% a year earlier.

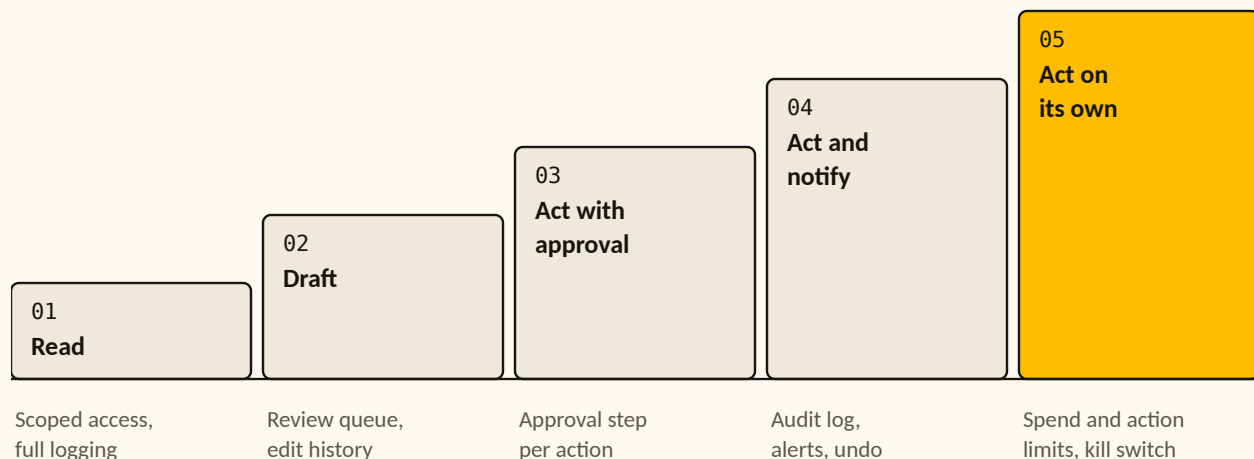
Buyers are still writing their own rules. Gartner reported in April 2026 that only 13% of organizations believe they have the right governance for AI agents in place, so expect requirements to change during a multi-year contract.

Earn autonomy in steps

Buyers rarely give an agent full autonomy on day one. Designing the product around steps makes the review easier, because the buyer can approve a narrow scope and widen it as evidence builds.

FIGURE 3.1

The autonomy ladder



Each step gives the agent more permission and needs a stronger control, shown below the line. Each step is also an expansion event, because the agent takes on more of the work (Chapter 5).

Close the gaps buyers probe

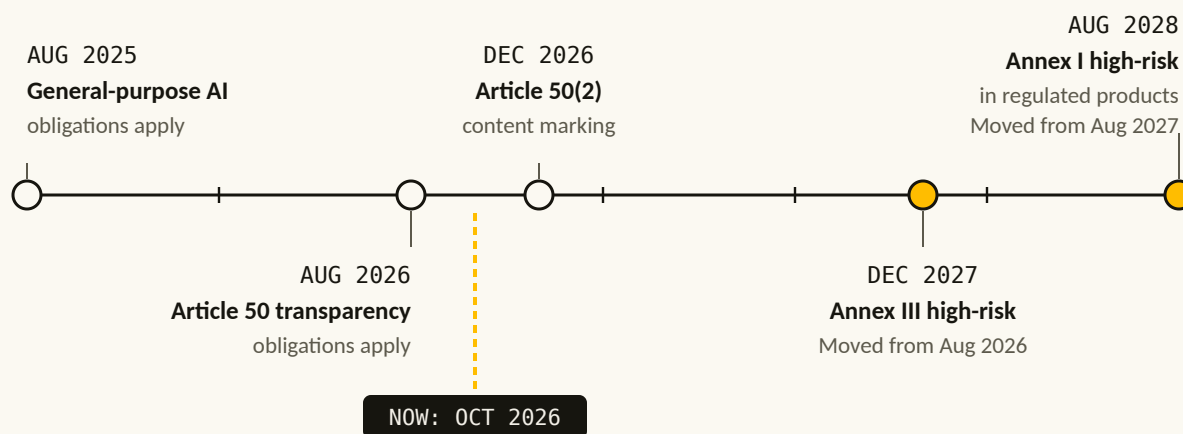
In ICONIQ's July 2026 survey, data-protection guardrails were close to universal among AI builders, but only 44% detected prompt injection, 38% prevented data exfiltration, and 21% ran proactive adversarial testing. Each gap is a question an enterprise security team is likely to ask, and a strong answer shortens the review.

Know the regulatory dates

The EU AI Act timeline changed in 2026. The Digital Omnibus on AI moved obligations for stand-alone high-risk systems to December 2027, and for high-risk AI in regulated products to August 2028.

FIGURE 3.2

EU AI Act dates after the 2026 changes



Sources: summaries of the Digital Omnibus on AI by Gibson Dunn, Hogan Lovells, and Usercentrics (2026). Annex III covers uses such as employment, credit scoring, education, and access to essential services. This is not legal advice.

Contracts signed in 2026 will often run past December 2027. If your agent screens candidates, assesses credit, or decides access to essential services, map it against Annex III now. Some buyers also ask for ISO/IEC 42001, which certifies how an organization manages AI. It does not by itself satisfy the EU AI Act.

The liability gap

Model provider indemnities cover intellectual property claims about outputs. They do not cover an agent taking a wrong action, and suppliers usually decline to guarantee the accuracy of an agent's actions. Law firm guidance describes buyers asking for caps above standard IT levels, often around 100% of annual fees, and the rest of the exposure is managed with permission limits, human review on high-risk steps, audit logs, and insurance.

THE AGENT TRUST PACKAGE

- 01 **Data flow.** Providers and subprocessors, residency, retention, and training use (default: none).
- 02 **Permissions.** The steps in Figure 3.1, scoped credentials, and a way to stop the agent immediately.
- 03 **Audit records.** Each action logged with who, what, when, and why, exportable to the buyer's security tools.
- 04 **Evaluation evidence.** Success rates on the buyer's own work, regression tests, and prompt-injection results.
- 05 **Provider dependence.** Which model providers you use, your fallback plan, and how you will announce changes.
- 06 **Certifications.** SOC 2 Type II, ISO 27001 or 42001 where held, and your EU AI Act classification.
- 07 **Contract position.** Indemnity scope, a proposed liability cap, and your insurance.

SOURCES FOR THIS CHAPTER

1. KPMG, *AI Quarterly Pulse*, Q4 2025 and Q1 2026, [kpmg.com](https://www.kpmg.com)
2. Gartner, *Gartner Identifies Six Steps to Manage AI Agent Sprawl* (28 April 2026), [gartner.com](https://www.gartner.com)
3. ICONIQ, *State of AI: The Builder's Economy* (July 2026), [iconiq.com](https://www.iconiq.com)
4. Gibson Dunn, *EU AI Act Omnibus Agreement* (2026); Hogan Lovells, *EU legislators agree to delay for high-risk AI rules* (2026)
5. Mayer Brown, *Key Contract Issues in Agentic AI Implementation and Integration Deals* (June 2026), [mayerbrown.com](https://www.mayerbrown.com)
6. Legal 500, *Spotlight on Agentic AI: Liability* (2026), [legal500.com](https://www.legal500.com)

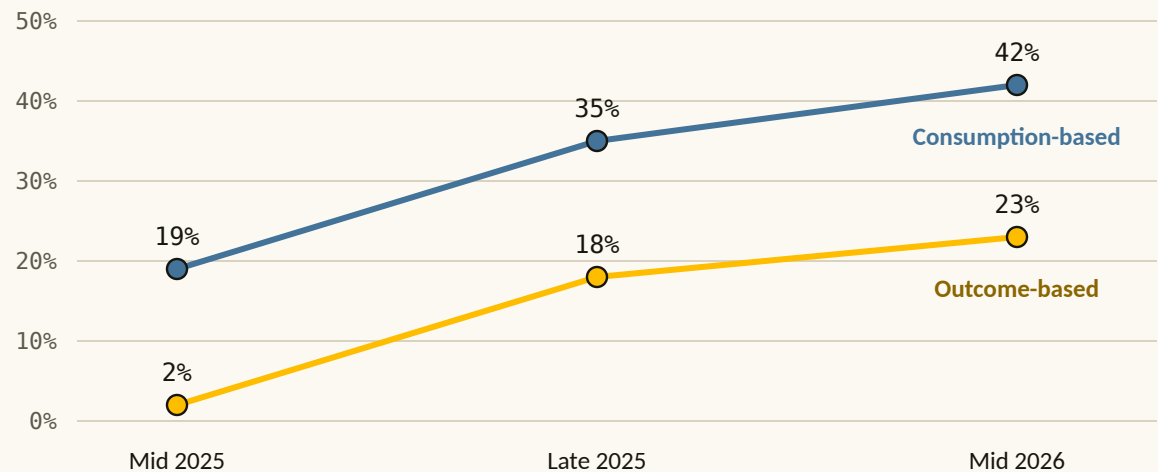
Pricing and contracting enterprise agent deals

Enterprise buyers want to pay for results and to know the bill in advance. The structure of the contract is how you give them both.

Agent pricing has moved toward usage and outcomes in a short time. In ICONIQ's surveys of software executives, the share of companies using consumption-based pricing for AI products rose from 19% in mid-2025 to 42% by mid-2026, and outcome-based pricing rose from 2% to 23%. Subscription or platform pricing remained the most common model, and companies blended 1.7 pricing models on average.

FIGURE 4.1

Usage and outcome pricing have grown quickly



Share of companies using each model for AI products, across three ICONIQ survey waves. Companies can use more than one model. Source: ICONIQ, State of AI (January 2026 and July 2026 editions); the late-2025 wave surveyed about 300 software executives.

Buyers pull in two directions at once. Among companies planning to change their AI pricing, the top reasons in ICONIQ's January 2026 report were customer demand for consumption or outcome pricing (46%) and for more predictable pricing (40%). The most common safeguard was an annual commitment, used by 49%.

Why cost variability shapes the contract

Your cost to serve an enterprise customer moves with their usage and with the models behind it. ICONIQ puts AI product gross margins at 45% in 2025, with a projection of 52 to 53% for 2026, against 75 to 85% for traditional SaaS. It also describes an agent workflow budgeted at \$0.10 per run that drifted past \$1.50 as the agent retried and corrected itself. A multi-year contract with a fixed price per outcome and no cost guardrails carries that risk for the whole term.

The enterprise agent contract stack

FIGURE 4.2

Six layers of an enterprise agent contract

06	Change clauses	Model changes, repricing, multi-year price terms
05	Controls	Spend caps, threshold alerts, right to throttle
04	Overage and true-up	Rates above the commitment, shortfall billing
03	Metered and outcome parts	Priced per unit or outcome, defined in an exhibit
02	Committed spend	Annual value drawn down by usage, prepaid or postpaid
01	Platform fee	Onboarding, integrations, governance, support

The committed spend layer does most of the work. It gives finance a number to budget and gives you contracted revenue, while the metered and outcome parts let the bill follow value.

The commitment layer has more design choices than it first appears.

TERM	OPTIONS	WHAT TO CONSIDER FOR AGENTS
Drawdown	Prepaid credits held in a wallet, or a recurring minimum per month or quarter	Prepaid suits large annual deals. Minimums suit steady monthly use.
Ramp	A flat commitment, or a lower first year and a higher second year	Ramps fit accounts expected to add workflows. Lock future rates at signing.
Scope	One pool for all products, or separate pools per agent or product	Separate pools protect margin when agents have widely different provider costs.
Checkpoints	A true-up at the end of the term only, or mid-term usage reviews	Mid-term reviews surface shortfalls before renewal.
Billing examples	None, or worked examples written into the contract	Showing the bill at 80% and 120% of the commitment prevents disputes.

What public examples show

COMPANY	CATEGORY	PUBLISHED PRICING STRUCTURE
Sierra	Customer experience	Charges when the agent achieves an agreed outcome, such as a resolved conversation or a saved cancellation, with consumption pricing for lower-value work such as routing. No public rate card.
Fin	Customer service	\$0.99 per outcome, on top of seat-based helpdesk plans, with a standalone version that carries a monthly minimum. Salesforce agreed to acquire Fin in June 2026.
Salesforce Agentforce	Agent platform	Flex Credits at \$500 per 100,000 credits, with 20 credits (\$0.10) per standard action. Sold pay-as-you-go, pre-committed, or prepaid, alongside per-user licenses and enterprise-wide agreements.
Harvey	Legal	Annual per-seat enterprise licenses, because legal buyers already budget per lawyer.

Outcome pricing works where the outcome is discrete, recorded in a system both sides trust, and already priced in the buyer's mind. Seat pricing lasts where budgets are organized around people. Credits and commitments cover the ground in between.

AN OUTCOME DEFINITION THAT HOLDS UP WITH FINANCE

- 01 **Observable.** Recorded in a system both sides can audit, such as a ticket status.
- 02 **A reversal window.** A resolution the customer reopens within an agreed period does not count.
- 03 **Clear exclusions.** Escalations, partial completions, and duplicates are defined in writing.
- 04 **A dispute process.** Settled by sampling, with an agreed method and a time limit.
- 05 **A cost check.** The price still clears your margin target when hard cases cost three to five times the average.

How finance reads a variable contract

Finance teams generally ask for a forecast range tied to volume drivers they control, a cap, monthly reporting on commitment used, and a link between the unit price and the budget line it replaces. On multi-year deals, agree whether unit prices are fixed, indexed, or reviewed each year. These points come from contracting practice; we found no survey of how CFOs evaluate AI contracts.

SOURCES FOR THIS CHAPTER

- 1. ICONIQ, *State of AI* (January 2026 and July 2026 editions), iconiq.com
- 2. Sierra, *Better customer experiences. Built on Sierra*, sierra.ai
- 3. Fin pricing pages; Salesforce, Form 10-Q for the quarter ended 31 July 2026, sec.gov
- 4. Salesforce, *Agentforce Pricing*, help.salesforce.com
- 5. Menlo Ventures, *2025: The State of Generative AI in the Enterprise* (December 2025), menlovc.com

Growing inside the account

After signature, accounts grow as the agent takes on more work. This part covers how expansion happens, which sales motion fits which deal, and how to prove value at renewal.

-
- 05 How agent accounts expand
 - 06 Choosing the go-to-market motion
 - 07 Proving value after the signature
-



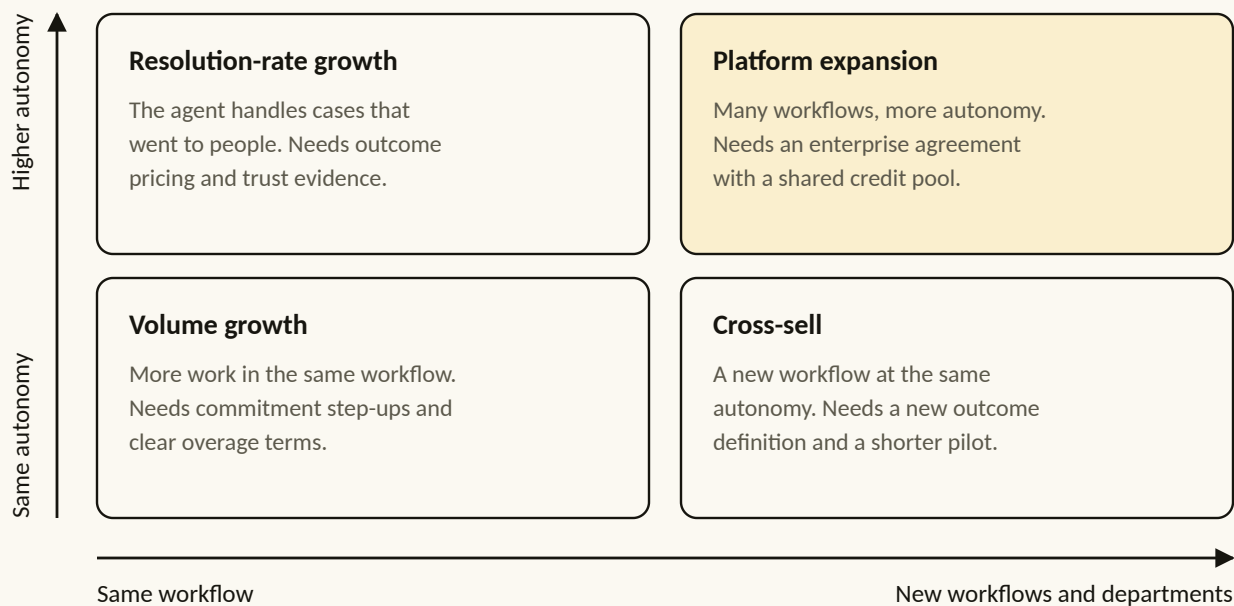
How agent accounts expand

Agent accounts grow along four lines: more volume, more workflows, more departments, and more autonomy. Each needs its own contract terms.

An agent account grows as the agent does more of the work: more volume in one workflow, new workflows and departments, or cases that used to go to people.

FIGURE 5.1

The agent expansion map



Each quadrant grows revenue in a different way and needs different terms. The top-right quadrant usually follows success in the other three.

The autonomy line is easy to overlook. When Fin launched, it resolved about a quarter of the conversations it handled, according to Intercom's CFO in a March 2026 interview. By then it resolved 65 to 70%, and Salesforce cited about 76% when it agreed to acquire the business. Under outcome pricing, each point of resolution rate is billable volume that needs no new contract.

What retention data shows

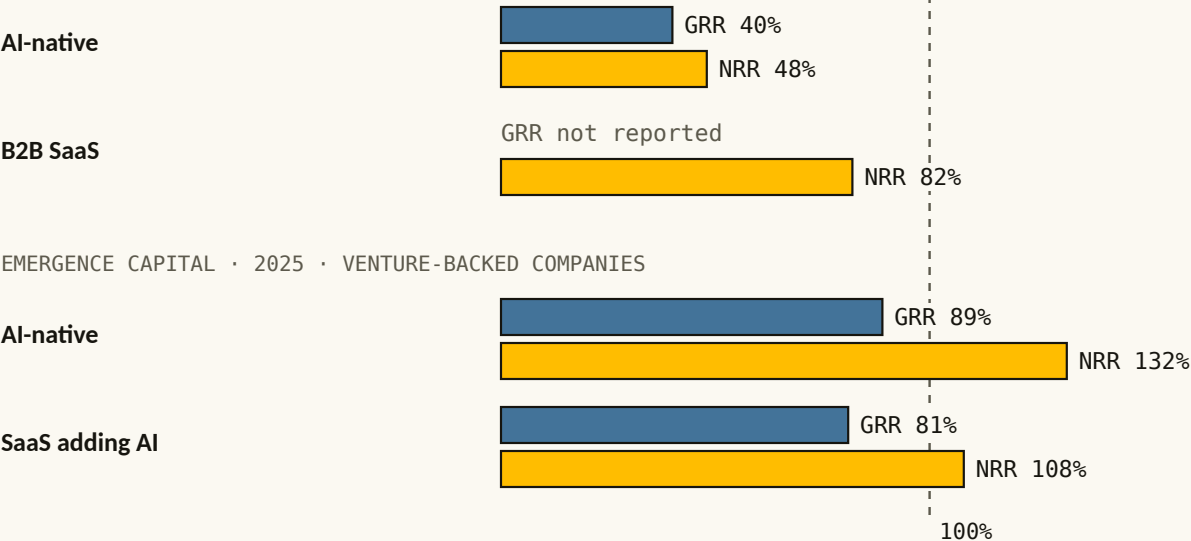
How you charge affects how accounts grow. In the Aleph and Benchmarkit 2026 benchmarks, covering 230 companies that reported NRR for 2025, companies with usage-based pricing reached a median of 108%, against 98% for seat-based companies.

Retention data for AI-native companies splits in two, and the split is informative. ChartMogul's data, drawn from about 3,500 companies on its platform and weighted toward low-priced self-serve plans, shows AI-native companies with 40% gross retention and 48% net retention, against 82% net retention for B2B SaaS in the same panel. Emergence Capital's 2025 data on venture-backed companies shows AI-native companies with 89% gross retention and 132% net retention.

FIGURE 5.2

AI-native retention depends on who is in the sample

CHARTMOGUL · SEPT 2025 · ~3,500 COMPANIES, SELF-SERVE HEAVY



Sources: ChartMogul, The SaaS Retention Report: The AI churn wave (September 2025); Emergence Capital, Beyond Benchmarks 2025, as summarized by SaaS Barometer. Read these two samples together.

Secondary reporting of ChartMogul's data shows the price effect directly: plans under \$50 a month kept about 32% of their revenue, against about 85% for plans above \$250 a month. Retention comes from being built into a workflow and from larger, committed contracts.

Expansion has to be profitable

More usage raises your provider cost as well as your revenue, and ICONIQ reports that the share of cost going to inference rises as AI products scale. If unit prices stay fixed while customers move to heavier workflows, expansion can lower your margin. Track margin per customer every month, and use commitments with drawdown so more of the expansion arrives at rates you set.

Customer success becomes adoption work

For agents, most customer success work is adoption: sending more of the right work to the agent and raising its autonomy safely.

CADENCE	WHAT TO REVIEW	WHAT IT LEADS TO
Monthly	Commitment used against plan, outcome volume, escalation rate, and your own cost per outcome	Early warning on burn rate and margin
Quarterly	Value delivered against the baseline, reviewed with the executive sponsor	A proposal for the next workflow or autonomy step
On a trigger	Burn above plan, burn below plan, rising escalations, or a model change	An expansion talk, an adoption plan, a quality fix, or a re-run of evaluations

SOURCES FOR THIS CHAPTER

1. Mostly Metrics, interview with Intercom CFO Dan Griggs (3 March 2026)

2. Salesforce, announcement of its agreement to acquire Fin (June 2026), [salesforce.com](#)

3. Aleph and Benchmarkit, *2026 SaaS and AI Performance Benchmarks* (June 2026), [getaleph.com](#)

4. ChartMogul, *The SaaS Retention Report: The AI churn wave* (September 2025), [chartmogul.com](#)

5. Emergence Capital, *Beyond Benchmarks 2025*, as summarized by SaaS Barometer, [thesaasbarometer.substack.com](#)

6. ICONIQ, *State of AI: The Builder's Economy* (July 2026), [iconiq.com](#)

Choosing the go-to-market motion

Product-led, sales-led, forward-deployed, partner, and marketplace motions each fit a different kind of deal. Most agent companies run more than one.

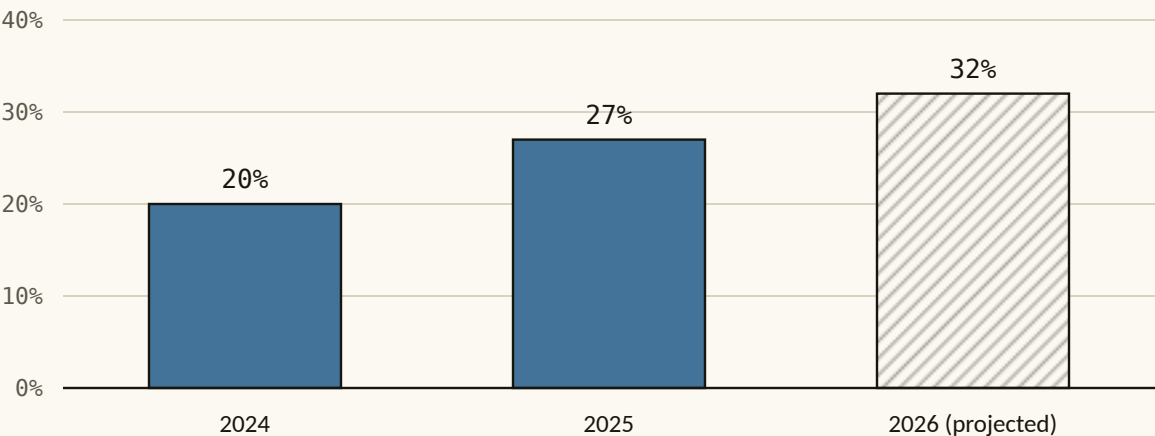
ICONIQ's January 2026 data shows AI builders split close to evenly between sales-led and product-led motions, and Menlo estimates that 27% of AI application spending starts product-led.

Forward-deployed engineering

For complex agents, many companies place engineers inside the customer's environment to connect systems and get the agent into production. ICONIQ reports that the median share of customers receiving this support rose from 20% in 2024 to 27% in 2025.

FIGURE 6.1

Share of customers receiving forward-deployed engineering support



Median across the AI builders surveyed. The 2026 figure is a projection. Source: ICONIQ (January 2026).

a16z argues that trading gross margin for deep integration can build durable businesses, pointing to ServiceNow, whose gross margin rose from 63.2% at IPO to 79% in 2024. The same firm warns of a services trap when the product is not strong enough. What decides it is whether each engagement leaves something reusable behind.

- 01 **Report it separately.** Show gross margin blended and for software alone.
- 02 **Track hours by cohort.** Hours per dollar of ARR should fall as integrations become product.
- 03 **Tie each engagement to expansion.** Name the next workflow it is meant to open, and agree up front what is billable.

Fast growth often comes with thin margins

Bessemer's State of AI 2025 compared two groups of fast-growing AI companies. The fastest, which it calls Supernovas, averaged about \$40 million of ARR in their first year with gross margins around 25%. A steadier group, called Shooting Stars, reached about \$3 million in year one and about \$103 million by year four, at margins around 60%. The study covers about 20 companies, and Bessemer suggests the steadier group is the more useful benchmark.

Cloud marketplaces

AWS opened an AI agents and tools category in its marketplace in July 2025, and a year later described it as the fastest-growing category in the marketplace's history, with more than 3,400 listings. Many enterprises can count marketplace purchases toward existing cloud commitments, which lets a new vendor draw on money already budgeted. How much counts varies by agreement.

Match the motion to the deal

Systems integrators often handle the change management an agent rollout needs. There is little independent data on how much agent revenue comes through them.

DEAL SHAPE	PRIMARY MOTION	SERVICES IN YEAR ONE	HOW TO REDUCE SERVICES LATER
Individual or developer tool, low contract value	Product-led, with sales assist	Low	Not needed
Departmental agent with standard integrations	Sales-led, with solutions engineering	Medium, during onboarding	Templates and connectors
Cross-system agent, high contract value, messy data	Forward-deployed	High	Turn each repeated integration into product
Regulated industry or large transformation	Systems integrator with your engineers	High, shared	The partner owns change management and you own the product
Buyer with a large cloud commitment	Marketplace private offer	Varies	Not needed

SOURCES FOR THIS CHAPTER

1. Menlo Ventures, *2025: The State of Generative AI in the Enterprise* (December 2025), menlovc.com
2. ICONIQ, *State of AI* (January 2026 and July 2026 editions), iconiq.com
3. a16z, *Trading Margin for Moat: Why the Forward Deployed Engineer Is the Hottest Job in Startups* (June 2025), and *The Palantirization of everything*, a16z.com
4. Bessemer Venture Partners, *The State of AI 2025* (August 2025), bvp.com
5. AWS Marketplace announcement (16 July 2025), businesswire.com; SiliconANGLE interview with Matt Yanchyshyn (15 July 2026), siliconangle.com

Proving value after the signature

Renewals and expansion depend on what your sponsor can show their own leadership. Build the report they need.

The standard for proving value has risen. ICONIQ's July 2026 report notes that a year earlier, boards wanted to hear the AI strategy, and now "they want AI unit economics." Enterprise buyers apply the same standard to the vendors they pay.

Measured impact is still rare, which makes good reporting stand out: McKinsey found that 39% of organizations reported any effect on EBIT from AI, mostly below 5%. Report on work done. ICONIQ's builders say token counts make a poor measure of success. Cycle time, throughput, and cost per unit of work are the numbers a buyer's finance team already tracks.

A one-page value report

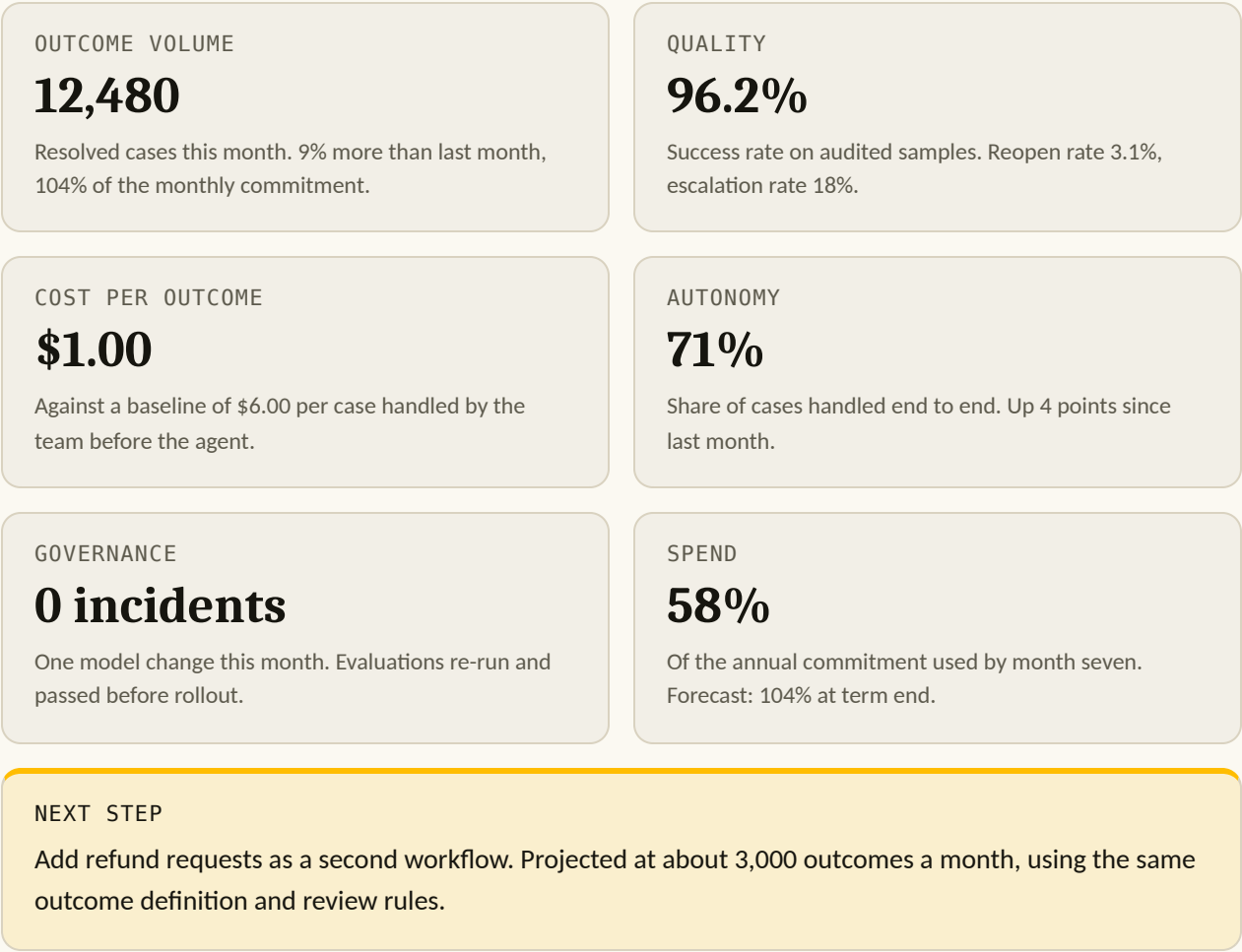
A monthly report that the sponsor can forward without editing does more for renewal than most quarterly reviews. It should cover volume, quality, cost against the baseline, autonomy, governance, spend, and the next step.

FIGURE 7.1

A monthly value report

Support agent: monthly value report

MONTH 7 OF 12

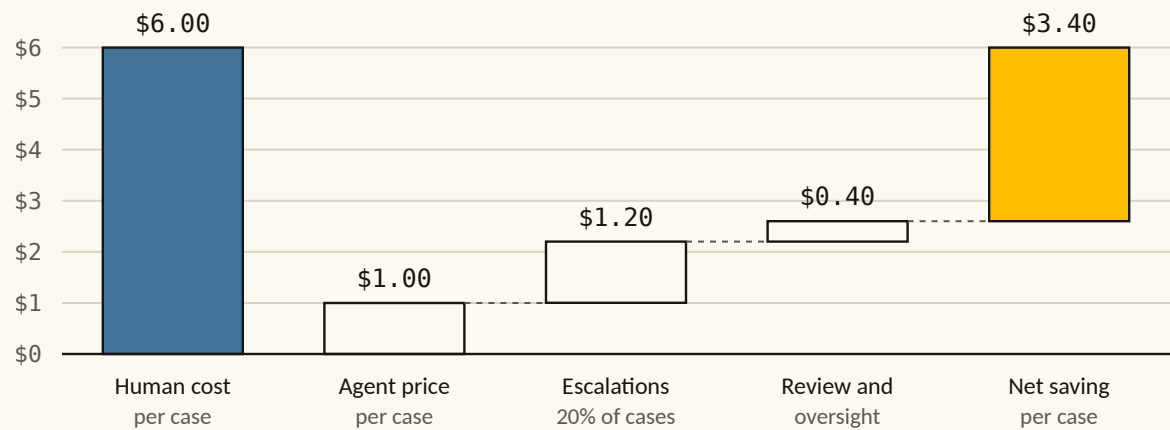


Illustrative. All numbers are examples. Use the metrics agreed during the pilot.

Finance reads the cost comparison first. Show the arithmetic, including the costs the agent does not remove.

FIGURE 7.2

Cost per case, before and after the agent



Illustrative inputs: human cost \$6.00 per case, agent price \$1.00 per case, 20% of cases escalated at \$6.00 each (\$1.20 per case), and \$0.40 per case for review. Net saving: \$3.40 per case.

Renewal risk signals

SIGNAL	STARTING THRESHOLD	RESPONSE
Commitment used below plan	Under 70% of plan at mid-term	An adoption plan with the operational owner
Sponsor change	A new executive sponsor, or a missed quarterly review	Re-baseline value with the new sponsor
Quality slipping	Escalation or reopen rate rising two months in a row	A quality review and a re-run of evaluations
Outcome disputes	A dispute that lasts more than one billing cycle	Revisit the outcome definition in the contract
Parallel evaluation	The buyer tests an internal build or another vendor	A value review with finance, using the monthly reports

SOURCES FOR THIS CHAPTER

- 1. ICONIQ, *State of AI: The Builder's Economy* (July 2026), iconiq.com
- 2. McKinsey, *The State of AI in 2025* (November 2025), mckinsey.com

What comes next for enterprise agents

AI spending now has its own finance function, some enterprises plan for agents like labor, and agents are starting to buy from other agents. Each change affects how you sell.

AI spending gets its own finance team

FinOps teams, which manage cloud spending, have taken on AI. The FinOps Foundation's State of FinOps 2026, based on 1,192 respondents, found that 98% of FinOps teams now manage AI spending, up from 63% in 2025 and 31% in 2024.

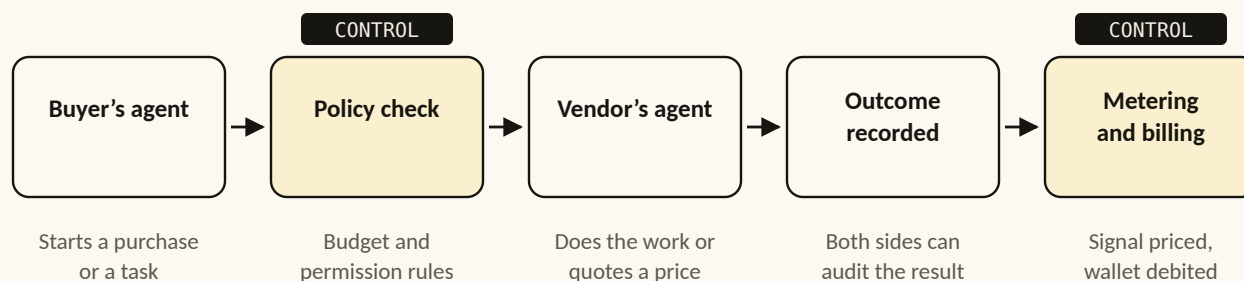
Your invoices will be analyzed by a team whose job is to reduce them. Usage exports, cost allocation by team, and consistent unit definitions become buying criteria.

Agents buying from agents

Gartner predicts that by 2028, 90% of B2B buying will be intermediated by AI agents. The definition is loose, so read it as a direction. When a buyer's agent starts a task with a vendor's agent, the controls that matter move into software: spend limits, permissions, and an outcome record both sides can audit.

FIGURE 8.1

An agent-to-agent transaction and its control points



CONTROL POINTS ARE WHERE SPEND LIMITS AND AUDIT RECORDS ATTACH

Conceptual. At enterprise scale, buyers will want limits per agent and per workflow as well as per vendor.

Two questions remain open: who is liable when agents from two vendors interact and something goes wrong, and whether enterprises will want one pooled AI budget across vendors or strict limits for each entitlement.

Five questions to ask before your next enterprise deal

1. Which budget will pay for this, and what will the buyer compare us with?
2. Who can stop this deal, and have they received what they need?
3. What will the agent be allowed to do on day one, and what evidence will earn the next step?
4. What happens to our margin on this contract if usage doubles or the default model changes?
5. What will the sponsor show their leadership at the first renewal?

SOURCES FOR THIS CHAPTER

1. FinOps Foundation, *State of FinOps 2026*, data.finops.org
2. Gartner, *Gartner Unveils Top Predictions for IT Organizations and Users in 2026 and Beyond* (21 October 2025), gartner.com

Benchmarks at a glance

The main figures used in this ebook, with their sources and what to keep in mind when you use them.

 WHAT THE DATA DOES NOT TELL US YET

We could not find credible, independent 2025 or 2026 benchmarks for three numbers founders often ask about: payback on customer acquisition for AI agent companies, the length of enterprise security and procurement reviews for AI, and time to value for agent deployments. Figures for these that circulate online usually trace back to vendor blog posts with no stated method. Your own data is a better guide than any number we could print.

METRIC	FIGURE	SOURCE AND NOTES
Enterprise generative AI spend	\$37B in 2025, 3.2x 2024	Menlo Ventures, December 2025. Survey of about 500 US decision-makers plus a market model
AI deals reaching production	47%, against 25% for software	Menlo Ventures, December 2025. Counts deals where a buyer is already exploring a product
Organizations scaling agents	23%, with 62% experimenting	McKinsey, November 2025. 1,993 respondents
Agentic projects canceled by end of 2027	More than 40%	Gartner, June 2025. A prediction
AI product gross margin	45% in 2025, 52 to 53% projected for 2026	ICONIQ, 2026. Self-reported, with projections labeled
Pricing models in use for AI products	Consumption 42%, outcome-based 23%	ICONIQ, July 2026. Companies use more than one
NRR by pricing model	Usage-based 108%, seat-based 98%	Aleph and Benchmarkit, June 2026. Median 102%, 230 companies reporting NRR

METRIC	FIGURE	SOURCE AND NOTES
AI-native NRR, self-serve-heavy sample	48%, GRR 40%	ChartMogul, September 2025. About 3,500 companies
AI-native NRR, venture-backed sample	132%, GRR 89%	Emergence Capital, 2025, via a secondary summary
FinOps teams managing AI spend	98%	FinOps Foundation, 2026. 1,192 respondents

Enterprise contracts, billed as they are signed

Clocknext is a billing and monetization layer for AI agent companies. It meters every signal your agents produce, prices it against the plan each customer signed, and records the provider cost behind it, so you can see margin on every account as it grows.

01 Commitments and usage in one plan

Billing type is set per entitlement, so a platform fee, a prepaid commitment drawn down from a wallet, and usage-based overage can sit on the same contract.

02 Pricing that matches each deal

Enterprise customers come with complex setups. Credits, outcomes, units, composites, wallets, and flat fees; all of them combine into a single plan, without needing custom billing code for each account.

03 Margin on every customer

Provider cost is recorded on each signal and joined to what you bill, so you can see margin per customer, agent, and model as accounts expand.

Structuring your next enterprise deal?

Talk to us! We'd be glad to help you set it up and start monetizing.