INFORMATION MATTERS
ANALYST RESEARCH // AGENTIC AI

The Agentic AI Market 2026

Market Sizing, Competitive Landscape & Growth Forecast

First Quarter 2026

Fiona (AI Analyst) | Technology Analyst
Martin De Saulles | Principal Analyst
Information Matters | April 2026

Headline findings

- Central 2026 TAM estimate revised to **$40B** (range $33–$48B) — up $2B on the v2 midpoint, driven by the April 2026 revision of Anthropic's run-rate revenue to $30B and a narrower attribution-rate uncertainty.
- Anthropic passed OpenAI on run-rate revenue in April 2026 ($30B vs ~$24–25B), with >1,000 customers at $1M+ ARR — a number that has more than doubled since February 2026.
- Agentforce + Data 360 ARR crossed $2.9B (up >200% YoY); Agentforce alone reached $800M (+169% YoY) across 29,000 deals — but the implied $28K average contract value still signals pilot-scale deployment, not production-scale rollout.
- Coding agents remain the most credible production signal: Cursor $2B ARR, GitHub Copilot 4.7M paid seats, Claude Code ~$1B ARR within six months of launch.
- Our four-scenario 2030 forecast ranges $70B (bear) to $215B (bull), with a $140B base case. The scenario is set not by CAGR assumption but by three named, observable triggers (see §6.2).

Executive Summary

The agentic AI market is real, growing fast, and substantially larger than the application-layer-only estimates suggest — but still substantially smaller than the headline figures from content-factory research firms would have you believe. This third edition revises our 2026 central estimate to $40B (range $33–$48B), up $2B on the v2 midpoint. The upward revision is not a change in method; it is a change in an input. Anthropic's April 2026 disclosure of $30B annualised run-rate revenue (up from $9B at end-2025 and $19B in March) makes the single most consequential assumption in our market-sizing — the agentic share of foundation model revenue — a larger number in absolute terms.

We read the current state of the market as three stacked narratives running in parallel. First, a foundation-model layer consolidating faster than any enterprise software category in history, with Anthropic's enterprise-heavy mix now overtaking OpenAI's consumer-heavy mix on run-rate revenue. Second, an application layer splitting along a clear fault line — coding agents in genuine production, every other category still converting pilots. Third, an infrastructure layer (evaluation, observability, orchestration, voice) in which three or four winners are being chosen right now and in which the capital deployed is not yet matched to the strategic importance.

The most important pattern in the April 2026 primary data is the divergence between token consumption and deal economics. Salesforce disclosed 19 trillion tokens processed to date (up 5× year-on-year) and 2.4B "Agentic Work Units" delivered. That is not pilot-scale usage — something real is happening inside Agentforce deployments. But the $28K average contract value across 29,000 deals still says the buyers are pricing this as experimentation, not as replacement of headcount. Both things are true at once: heavy usage, light commitment. Which side wins by 2027 is the single most important question for the application-layer market.

The So What, sharpened from v2. The bottom-up evidence from developer tools remains the most reliable signal in this market — 85% of developers globally use AI tools regularly, 74% had adopted specialised developer AI tools by January 2026 (JetBrains, 24,534 respondents). But the top-down evidence from foundation model revenues is now equally credible as a signal, because the disclosures are primary, frequent, and specific. The market is being built bottom-up at the developer layer and top-down at the model layer simultaneously. The missing middle — durable enterprise production deployments with six- and seven-figure annual contracts at scale — is where the 2027 narrative gets decided.

1. Scope and Methodology

This report estimates the size of the global agentic AI market in 2026 and projects growth through 2030. All figures cited are drawn from primary sources: company earnings calls and press releases, SEC filings, investor letters, disclosed funding-round metrics, and peer-reviewed

developer surveys. We do not cite market research firms whose methodology is opaque. Where we rely on estimates rather than disclosures, we say so explicitly and explain our reasoning.

1.1 Defining the market

Agentic AI is defined here as AI systems that take multi-step actions in the world, use tools or APIs to interact with external systems, and operate with meaningful autonomy between human checkpoints. This definition excludes pure text generation, standard chatbots without tool use, and AI-assisted features with no autonomous action capability. It includes coding agents that write and execute code, customer operations agents that resolve tickets, orchestration systems that plan and execute multi-step workflows, and foundation model API usage specifically attributable to agentic applications.

This definition is narrower than the "generative AI" market and broader than the "enterprise agent platforms" slice. The former conflates agentic use with all LLM usage; the latter misses the foundation model and developer tooling layers where most economic value is currently being generated. The boundary we police most carefully is the one between "assisted" and "agentic" — a feature that auto-completes one line of code is not agentic; a feature that reads a codebase, opens a pull request, and responds to CI failures is.

1.2 The foundation-model attribution rate — our key methodological assumption

The most important and most challengeable decision in this report is the treatment of foundation-model revenues. Anthropic and OpenAI together have approximately $54–55B in combined annualised run-rate revenue as of April 2026 ($30B and $24–25B respectively). Including this in full would produce a market-size figure that is largely a measure of two companies. Excluding it entirely — as many market analyses do — understates the economic value actually being generated by agentic AI.

We apply an attribution rate of 25–35% to foundation-model revenues, representing the estimated share of API calls and product usage that involves genuinely agentic behaviour. We arrive at this range from four sources:

- **Anthropic's own product breakdown.** Claude Code, launched publicly in mid-2025 and now at approximately $1B ARR, is entirely agentic. The April 2026 Bloomberg confirmation that >1,000 customers now spend $1M+ on Claude annually — more than double the February 2026 figure — is strong evidence that the mix is shifting toward workflow and tool-use patterns, not chat.
- **a16z CIO Survey (100 Global 2000 executives, October 2025).** 44% of enterprises are using Anthropic in production; 37% use five or more AI models. The survey distinguishes model usage for content generation versus workflow automation — the latter is consistently reported as the primary driver of budget increases.

- **OpenAI product mix.** CFO Sarah Friar has disclosed that enterprise is accelerating faster than consumer but remains a minority of revenue. Operator, Deep Research, and the Assistants API with tool use represent a growing but still minority share of OpenAI's $24–25B run rate.
- **Salesforce agentic-work-unit telemetry.** New in FY26 Q4: Salesforce reports 2.4B AWUs delivered (+57% QoQ) against 19T tokens processed (5× YoY). This ratio — ~125 tokens per agentic work unit — is a useful triangulation on how much token consumption is agentic versus generative. Applied to Anthropic's API traffic, it supports a 30%+ agentic share.

We treat 30% as our central-case attribution, with 25% as the low case and 35% as the high case. This is a conservative range. The agentic share of foundation-model revenue is growing faster than the overall business; by 2028 we expect this attribution rate to exceed 50%.

Source: Anthropic announcement, 6 April 2026 (anthropic.com/news); Bloomberg via Yahoo Finance, 7 April 2026; a16z Enterprise AI Report, October 2025; Salesforce Q4 FY26 earnings release, 26 February 2026 (s205.q4cdn.com).

1.3 Private-company sizing methodology

For private companies with no disclosed revenue, we use LinkedIn headcount as a proxy multiplied by revenue-per-employee benchmarks derived from disclosed-revenue companies in each segment. For infrastructure companies: $180,000–$250,000 revenue per employee. For application-layer companies: $120,000–$160,000. These are conservative relative to top-quartile SaaS benchmarks and deliberately do not reflect the anomalous revenue-per-employee ratios at the very top of the market (Cursor reportedly operates with fewer than 100 employees at $2B ARR — a $20M-plus per-employee ratio that is not generalisable).

Public-company data is taken directly from earnings press releases and investor-call transcripts filed with the SEC or published on investor-relations pages.

1.4 Reconciliation with other published figures

Our 2026 central estimate of $40B sits above the headline TAMs typically reported by application-layer analyses and below the largest content-factory numbers. The reconciliation:

Published figure	2026 value	Why it differs from our $40B central
Information Matters IM#1512 (Feb 2026)	$52B by 2030	Scope: application + infrastructure only, no foundation-model attribution. IM#1512 measures the stack; this report adds the model layer that drives it. The two figures are consistent, not contradictory.
Application-layer analyses (typical)	$12–18B	Scope narrower still: excludes foundation models and often excludes coding agents. Useful as a lower bound on the enterprise-software line; not a total market figure.

Published figure	2026 value	Why it differs from our $40B central
Aggregator-firm figures (typical)	$9–47B	Methodology opaque; typically a CAGR applied to a base. Wide range reflects the arbitrary nature of the base. We do not cite these figures.
Broad "generative AI" market figures	$60–120B	Scope far broader: includes all LLM usage (consumer chat, content generation, image/video), not just agentic use. Not comparable to our scope.

2. Market Size: 2026

Our central estimate is $40B for the global agentic AI market in 2026. The range of $33–$48B reflects uncertainty in the foundation-model attribution rate and, to a lesser degree, in private-company sizing. Here is the segment-by-segment build.

2.1 Foundation model layer

As of April 2026, the three largest foundation-model providers have disclosed or credibly reported the following:

Company	Revenue / ARR	Source	Agentic attribution
Anthropic	$30B ARR (Apr 2026)	Company announcement, 6 Apr 2026; Bloomberg	25–35% = $7.5–$10.5B
OpenAI	~$24–25B ARR (Apr 2026)	Multiple reports citing CFO disclosure	25–35% = $6.0–$8.75B
Google (Gemini)	Est. $3–5B	Not separately disclosed; analyst estimates	25–35% = $0.75–$1.75B
xAI, Mistral, Z.ai, others	Est. $1–2B combined	Public reporting; funding disclosures	25–35% = $0.25–$0.7B

Total foundation-model agentic attribution: $14.5–$21.7B (central case: $17B). This is the range we carry forward.

The Anthropic trajectory warrants specific scrutiny. The sequence is now documented in near-monthly increments: $9B at end-2025; $14B in February 2026; $19B in March; $30B in April. That is $16B of run-rate growth in roughly eight weeks. Meritech's Alex Clayton has publicly said no public software company's IPO trajectory in the last two decades shows growth like this. The infrastructure behind it is visible in two complementary data points: NVIDIA data-centre revenue of $193.7B for FY2026 (+68% YoY); and Anthropic's April 2026 Broadcom and expanded Google partnership, explicitly framed as a response to capacity constraints. This is a company being pulled by demand, not a company pushing a sales motion.

A methodological caveat worth stating plainly: the number of Anthropic customers spending $1M+ per year roughly doubled between February and April 2026. That is the tell on the attribution rate. Customers who commit $1M+ on Claude annually are not using it for ad-hoc

summarisation. They are building agentic workflows — coding pipelines, research automation, customer-operations agents — where the token volume scales with autonomous action. If the $1M+ cohort continues to double at the current cadence, our 30% central attribution will be too low by mid-2027.

Source: Anthropic announcement, 6 April 2026 (anthropic.com/news); Bloomberg via Yahoo Finance, 7 April 2026; NVIDIA Q4 FY2026 earnings, 26 Feb 2026 (nvidianews.nvidia.com); OpenAI CFO blog post, Q1 2026.

2.2 Enterprise agent platforms

The enterprise segment has the most transparent primary data of any segment, because all the major players are publicly listed and have begun reporting AI-specific metrics.

Company	Product	Disclosed metric	Source
Salesforce	Agentforce + Data 360	ARR >$2.9B (incl. $1.1B Informatica); Agentforce $800M ARR (+169% YoY); 29,000 deals (+50% QoQ); 2.4B AWUs; 19T tokens processed	Q4 FY26 earnings, 26 Feb 2026
ServiceNow	Now Assist	$1.5B raised FY26 guidance; deals >$1M ACV +130% YoY	Q1 2026 earnings, Apr 2026
Microsoft	M365 Copilot + agents	~15M M365 Copilot seats (+160% YoY); 90% Fortune 100 coverage	Q2 FY2026 earnings, Jan 2026
Palantir	AIP	$4.4B FY2025 revenue; US commercial +121% YoY	Q3 2025 SEC filing
IBM	watsonx	$2B software book of business	Q4 2025 earnings, Jan 2026
SAP / Oracle	AI agents	Est. $400–600M each	Analyst estimates; not disclosed

The Salesforce data point continues to warrant scrutiny, but in a more nuanced way than in v2. Two numbers move in opposite directions. The average contract value implied by $800M across 29,000 deals — roughly $28,000 — is a pilot metric, not a transformational enterprise commitment. But the consumption metrics (2.4B agentic work units delivered; +57% QoQ; 19T tokens processed, 5× YoY) are consistent with heavy production usage inside deployed accounts. Reconciling the two means the Salesforce Agentforce book splits into two distinct populations: a long tail of low-commitment pilots and a short head of genuine production deployments. The FY27 disclosure to watch is not ARR; it is the median deal size within the top quartile of accounts.

ServiceNow is the most credible large-scale enterprise agentic deployment story on primary data. Deals over $1M ACV growing 130%+ YoY is exactly the production-scale signal that Agentforce's aggregate deal-size data lacks. ServiceNow has workflow integration, enterprise relationships, and the data access that agentic AI needs in production. The $1.5B raised guidance implies management has visibility into delivery, not pipeline.

Total enterprise agent platform revenue (agentic-attributable): $8–10B for 2026 (central $9B), up from $7–9B in v2, reflecting the Salesforce and ServiceNow Q4 disclosures.

Source: Salesforce Q4 FY26 earnings transcript and press release (s205.q4cdn.com); ServiceNow Q1 2026 earnings call (newsroom.servicenow.com); Microsoft Q2 FY2026 earnings; Palantir Q3 2025 SEC filing (sec.gov); IBM Q4 2025 earnings (newsroom.ibm.com).

2.3 AI coding agents

The coding agent market remains the most transparent segment in the stack because several key players have voluntarily disclosed revenue milestones in the context of funding rounds and CEO interviews.

Company	ARR	Source	Notes
Cursor (Anysphere)	$2B	CNBC Nov 2025; Fortune Mar 2026	Fastest SaaS to $500M ARR in history
GitHub Copilot	>$1B est.	4.7M paid subs; Microsoft earnings	Analyst estimate from disclosed sub count
Claude Code (Anthropic)	~$1B	Press reports citing Anthropic data	$1B within ~6 months of public launch
Amazon Q Developer	n/d	Bundled with AWS; not separately disclosed	—
JetBrains AI / Codeium / others	~$0.5B est.	Headcount methodology	—

Total coding agent market: $5–7B ARR (central $6B), unchanged from v2. This is the segment with the clearest evidence of production-scale adoption. The JetBrains Developer Ecosystem Survey (24,534 respondents, April–June 2025) found 85% of developers use AI tools regularly and 62% rely on at least one AI coding agent. These are not survey respondents reporting aspirations — they are practitioners describing daily workflows.

The April 2026 JetBrains workplace AI tools survey added an important nuance: organisational policy lags practice by approximately a year. Seventy-four percent of developers had adopted specialised AI developer tools by January 2026, but only 41% reported that their employer had a formal AI tools policy. This gap is a near-term headwind to paid enterprise seat growth — buyers who want to standardise have to create the policy first — and a medium-term tailwind, because the procurement conversation is now about which vendor, not whether.

Source: CNBC (Cursor funding, Nov 2025); Fortune (Cursor CEO interview, Mar 2026); Microsoft Q2 FY2026 earnings; JetBrains Developer Ecosystem Survey 2025 (devecosystem-2025.jetbrains.com); JetBrains workplace AI tools survey, April 2026 (blog.jetbrains.com).

2.4 Infrastructure, orchestration and vertical agents

These segments are dominated by private companies. We use the headcount methodology described in §1.3, cross-referenced against the 564 companies in the Information Matters database.

Segment	Companies	2026 estimate	Methodology / signal
Infrastructure & enablement	184	$3–4B	Headcount × $180–250K/employee. Eval & observability (68 cos) is the strategic sub-segment.
Orchestration layer	38	$1–2B	Headcount × $180–250K/employee. Consolidating — see §5.4.
Vertical & specialised	126	$2–3B	Headcount × $120–160K/employee. Coding (82 cos) dominates.
Capability-specific agents	46	$0.5–1B	Headcount × $120–160K/employee. Voice most developed.

2.5 Unit economics and margin analysis

The single biggest question hanging over the foundation-model sub-segment is not growth; it is margin. We consolidate the primary-source data that bears on this.

On training cost, the April 2026 Anthropic announcement — combined with the Broadcom capacity deal and existing Google partnership — surfaces a cost structure that differs materially from OpenAI's. Public reporting suggests OpenAI projects ~$121B in compute spend in 2028 alone, with projected losses of ~$85B that year; Anthropic's training-cost peak is projected at roughly $30B in the same period, approximately four times less, with profitability projected in 2028 or 2029. The two figures are projections, not disclosures, but the relative scale matches the disclosed hardware choices (Anthropic's use of Google TPU plus Amazon Trainium plus — newly — Broadcom custom silicon versus OpenAI's NVIDIA-dominated footprint).

On inference cost, the public price curve continues to fall. Anthropic and OpenAI have each disclosed price cuts of 60–80% on flagship models over the last 18 months. This compresses the revenue per token materially, which is why the correct framing is: token volume is growing faster than revenue, and agentic workloads are growing faster than total token volume. Both dynamics show up in the Salesforce disclosure — 5× token growth year-on-year is disproportionate to the 10% Salesforce total revenue growth, and consistent with an agentic-workload mix shift.

On the application layer, gross margins are under pressure at the coding-agent tier specifically. Cursor's reported $2B ARR with a reportedly small headcount implies revenue per employee well above $10M — but the cost of goods sold for an agent that makes heavy inference calls to a frontier model is meaningful and rising with usage. The market has not yet seen a disclosed gross margin for a major coding-agent vendor. When one is disclosed (either via an IPO filing or a secondary), it will be the most important data point in the sector. Our working estimate is 50–

60% gross margins for coding agents in 2026, well below the 75–85% range of conventional SaaS.

2.6 Total market summary

Segment	2026 Low	2026 Central	2026 High	Primary source
Foundation models (agentic portion)	$14.5B	$17B	$21.7B	Anthropic / OpenAI disclosures
Enterprise agent platforms	$8B	$9B	$10B	Earnings calls (public filings)
AI coding agents	$5B	$6B	$7B	Funding disclosures + sub counts
Infrastructure & orchestration	$4B	$5B	$6B	Headcount methodology
Vertical & capability-specific	$2.5B	$3B	$4B	Headcount methodology
TOTAL	$34B	$40B	$48.7B	

Central estimate: $40B. The range of $34–$48.7B is driven almost entirely by uncertainty in the foundation-model attribution rate. At 20% attribution the total falls to $32B; at 40% attribution the total reaches $51B. The narrower stated range reflects our view that the correct attribution range as of April 2026 is 25–35%, not 20–40%, given the April Anthropic disclosure on $1M+ customers.

3. Enterprise Adoption: What the Primary Data Shows

Beyond the revenue figures, the primary data on enterprise AI spending and adoption provides important context for understanding where the market actually is — and where the gap between signal and procurement reality shows up.

3.1 Enterprise spending patterns

The a16z annual enterprise AI survey (100 verified VPs and C-level executives in Global 2000 companies; October 2025) provides the most granular primary data on enterprise AI spend:

- Average enterprise LLM spend has risen from ~$4.5M two years ago to ~$7M today, with respondents expecting 75% growth in the coming year to ~$11.6M.
- Innovation budgets have dropped from 25% to 7% of LLM spending — AI is no longer experimental; it is being funded from operational IT and business-unit budgets.
- 44% of enterprises are using Anthropic in production. 37% are using five or more AI models. OpenAI, Google, and Anthropic hold dominant combined market share.

- Enterprises expected to spend ~$3.9M on AI applications but actually spent ~$6M — a significant underspend-to-overspend shift that suggests organic adoption is outpacing budgeted plans.

Source: a16z Enterprise AI Report, October 2025 (a16z.com/ai-enterprise-2025).

3.2 Developer adoption as a leading indicator

Developer tool adoption consistently leads enterprise software adoption by 12–24 months. The coding-agent numbers are therefore a forward indicator for the broader enterprise market, not just a segment of it.

- 85% of developers use AI tools regularly (JetBrains, 24,534-respondent survey, 2025).
- 74% had adopted specialised AI tools for developers by January 2026.
- 29% use GitHub Copilot at work; 18% use Cursor; 18% use Claude Code.
- GitHub Copilot is deployed at 90% of Fortune 100 companies (Microsoft earnings disclosure).
- Only 41% of developers report a formal employer AI tools policy (JetBrains April 2026) — the policy-to-practice gap is a near-term procurement headwind and a medium-term tailwind.

When 85% of the people who build and maintain enterprise software are already using AI agents daily, the question for enterprise adoption is not "if" but "when procurement catches up with practice."

Source: JetBrains Developer Ecosystem Survey 2025; JetBrains workplace AI tools survey, April 2026 (blog.jetbrains.com).

3.3 Q1 and early Q2 2026 momentum signals

The first four months of 2026 produced an unusually dense cluster of primary-source disclosures. We flag the ones that materially change the read on the market.

- **Anthropic's sequence, January–April 2026.** $14B (Feb) → $19B (Mar) → $30B (Apr). The $1M+ customer count more than doubled in the February-to-April window. This is the most consequential run-rate trajectory in the enterprise software history of record and the single largest input to our upward v3 revision.
- **Salesforce FY26 close, February 2026.** $41.5B total revenue (+10% YoY); Agentforce + Data 360 ARR $2.9B (+200% YoY); Agentforce rebranded across all five revenue lines; Informatica acquisition closed adding $1.1B Cloud ARR. The rebrand matters: it is a statement that Salesforce will not report agentic revenue as a separable line going forward, which reduces future transparency.
- **Anthropic–Broadcom compute deal, April 2026.** Confirms that the Anthropic revenue trajectory is constrained by compute supply, not demand. The material read-through is that 2026 Anthropic revenue is a function of how fast Broadcom, Google, and Amazon can deliver capacity, not how fast Anthropic can sell.

- **OpenAI 1.9 GW compute footprint disclosure.** CFO Friar's statement that OpenAI compute capacity has grown nearly 3× year-on-year to ~1.9 GW establishes compute as the binding constraint on OpenAI revenue. Same pattern as Anthropic: demand is not the bottleneck.
- **ServiceNow Q1 2026, April 2026.** Raised FY26 Now Assist guidance to $1.5B. The ≥$1M ACV cohort grew 130%+ YoY — the cleanest production-scale enterprise signal in the data.

4. Market Structure: Five Segments

The Information Matters database tracks 564 companies across five primary segments. The distribution shows a market weighted heavily toward infrastructure and horizontal enterprise applications — the picks-and-shovels layer — relative to vertical specialisations.

Segment	Companies	2026 revenue est.	Maturity signal
Infrastructure & enablement	184 (33%)	$3–4B	Rapid: 98 dev platforms, 68 eval/obs tools
Horizontal enterprise agents	170 (30%)	$8–10B	Growing: CX and Sales furthest ahead
Vertical & specialised	126 (22%)	$2–3B	Accelerating: Coding dominant (82 cos)
Capability-specific	46 (8%)	$0.5–1B	Early: Voice most developed
Orchestration layer	38 (7%)	$1–2B	Consolidating — see §5.4

4.1 Infrastructure & enablement (184 companies)

The evaluation and observability sub-segment (68 companies) is the one we remain most focused on. Production AI agents behave unpredictably; you cannot run them in regulated or customer-facing contexts without observability tooling. This category is underfunded relative to its strategic importance and will consolidate into 3–5 dominant platforms. Braintrust, Arize AI, and LangSmith are building real businesses. The winner here will be very large.

AI Agent Development Platforms (98 companies) is more contested. The incumbents are encroaching from both sides — foundation model providers building orchestration in, enterprise platforms embedding development tooling. The independent players need a defensible specialisation (domain, compliance, integration depth) or they will not survive consolidation.

4.2 Horizontal enterprise agents (170 companies)

Customer Operations (48 companies) is the most mature use case. The ROI is measurable (ticket deflection rates, resolution times), the technology is ready for production, and enterprise budgets are being committed. Salesforce's own Agentforce deployment on help.salesforce.com

handled over 750,000 requests in Q1 FY26 and cut case volume 7% YoY — a rare inside-the-tent data point that supports the category's production-readiness.

Revenue/Sales agents (41 companies) are growing but face a structural headwind: buyers are becoming sophisticated at filtering AI-generated outreach, eroding the engagement rates that justified early adoption. The companies that survive 2027 will pivot from volume outreach to research depth and account intelligence.

4.3 Vertical & specialised agents (126 companies)

Coding dominates with 82 companies — and the revenue data confirms it is the furthest-developed vertical. Supply Chain (34 companies) is the most interesting adjacent opportunity: complex enough to need specialist AI, clear enough ROI to justify procurement, less regulated than Legal or Health, and underserved by the horizontal platforms. Legal (5 companies) and Health (4 companies) are early but will be large — the regulatory and liability constraints that slow adoption also create durable competitive moats for companies that solve them first.

4.4 Orchestration layer (38 companies)

This segment will shrink fastest. Workflow orchestration (33 companies) is being absorbed into foundation-model APIs and enterprise platform suites. The independent orchestration layer will consolidate to a handful of specialists. The remaining value is in Human-in-the-Loop systems for high-stakes deployments where autonomous action requires human sign-off — a genuine and growing requirement in regulated industries.

4.5 Capability-specific agents (46 companies)

Voice is the most developed sub-segment here and, on our view, the single category most likely to produce a new very-large business in 2026–2027. The primary indicators: ElevenLabs' disclosed ARR milestones, Deepgram's enterprise traction, and the growth in voice-agent infrastructure primitives (Hume, Cartesia, Inworld). Customer service phone calls represent an enormous market. The first company to solve naturalness, latency, and multilingual at scale will own a category. Worth specific attention: the latency benchmarks published by these vendors are approaching human-turn-taking thresholds, which is a step-change, not a gradual improvement.

5. Competitive Landscape

5.1 The ones winning now

- **Anthropic.** $30B ARR in April 2026, up from $9B four months earlier, spending approximately 4× less than OpenAI on model training. This is developer- and API-driven growth at a speed that enterprise software has never seen. Training-cost advantage is now a documented pattern, not a single-quarter artefact. IPO preparation (Wilson Sonsini

retained, per multiple reports) will eventually force disclosure of unit economics; our expectation is that margin disclosure will be the positive surprise when it arrives.

- **Cursor.** $2B ARR, fewer than 100 employees, used by more than half the Fortune 500. The most compelling bottom-up adoption story in the stack. CEO Michael Truell told Fortune (March 2026) the company is focused on long-term mission, not IPO. Funding-round disclosures imply a $6B+ ARR trajectory by end-2026.
- **ServiceNow.** The $1.5B raised Now Assist guidance for FY26 is the most credible large-scale enterprise agentic deployment story. ServiceNow has workflow integration, enterprise relationships, and the data access that agentic AI needs in production. Considerably underrated versus Salesforce in analyst coverage.
- **Microsoft.** ~15M M365 Copilot seats and 90% Fortune 100 coverage is a distribution advantage that no independent vendor can replicate in its addressable market. The question for Microsoft is whether seat growth translates to agent-specific revenue or gets absorbed into bundled pricing that obscures the signal.

5.2 The ones to watch

- **Palantir.** Full-year 2025 revenue of $4.4B, US commercial growing 121% YoY, entirely AIP-driven. The company's intensive deployment model — five-day AIP bootcamps with deep workflow integration — is well suited to high-stakes agentic deployments where you cannot afford hallucinations. Not a household name in AI coverage; should be.
- **Evaluation & observability.** Braintrust, LangSmith, Arize AI, and three or four peers are building the quality-assurance layer for production agentic AI. This category will produce one very large business. Currently fragmented; currently underfunded relative to its strategic importance.
- **Voice-agent infrastructure.** ElevenLabs, Hume, Deepgram, and Cartesia are building the voice layer. The first company to combine sub-500ms latency, natural prosody, and enterprise-grade multilingual support at scale will own a category worth tens of billions.
- **Vertical coding specialists.** The 82 coding companies will not all survive horizontal incumbent pressure, but a subset — those that solve a specific pain (legacy migration, security review, infrastructure-as-code, data pipelines) — will. Watch for the ones whose customer list overlaps less with Cursor's than it does with traditional devops vendors.

5.3 The ones we would question

- **AI SDR platforms.** Engagement rates on AI-generated outreach are eroding as buyers adapt. The value proposition needs to shift from volume to research quality and personalisation. Companies that do not make this shift will have a difficult 2027.
- **Generic horizontal agent builders.** Too many companies selling the same "build your own agent" platform without defensible specialisation. Salesforce, ServiceNow, and Microsoft will absorb this functionality. The window to establish a specialised position is closing — we estimate 9–12 months.

- **Orchestration-only vendors.** The orchestration layer is being squeezed from above by foundation-model APIs adding tool use and from below by enterprise platforms embedding workflow. Independent orchestration is a feature, not a category.

5.4 Consolidation predictions through 2027

We expect four consolidation patterns to play out by year-end 2027. Each is stated as a falsifiable prediction with an observable trigger.

Pattern	Prediction	Observable trigger
Orchestration absorption	Independent workflow-orchestration vendors collapse to 3–5 specialist survivors; the rest are acquired or exit.	LangChain, CrewAI, or a peer announces acquisition by a model provider or enterprise platform.
Eval/obs consolidation	The 68 eval/observability vendors consolidate to 3–4 platforms; one exceeds $500M ARR.	A Series D at $3B+ valuation in the category, or a platform acquisition by a hyperscaler.
Voice-agent platform emergence	At least one voice-agent infra vendor crosses $500M ARR and announces an enterprise platform play.	A voice vendor signs a named $50M+ annual contract with a top-10 global bank or telco.
Coding-agent margin reset	First major coding-agent vendor discloses gross margin below 60% in an IPO filing or earnings release, repricing the sector.	Cursor S-1, GitHub Copilot segment disclosure, or a Codeium/JetBrains AI earnings break-out.

6. Growth Forecast: 2026–2030

6.1 Base case

Our base case projects a CAGR of 35–40% through 2030, decelerating from the current 45–50% as the market matures. Enterprise AI spending is currently growing at 75% annually (a16z) — our CAGR assumption is deliberately conservative relative to current momentum, on the grounds that current growth rates will not compound for five years uninterrupted. The base case runs below the momentum because momentum tends to break on capacity constraints, not demand constraints; compute supply is the binding constraint through 2027 on our view.

Year	Bear case	Base case	Bull case	Key inflection
2026	$34B	$40B	$49B	Current: developer adoption + foundation-model growth

Year	Bear case	Base case	Bull case	Key inflection
2027	$44B	$55B	$70B	Enterprise pilots convert (or do not); eval/obs consolidates
2028	$58B	$77B	$102B	Vertical specialisation; margin disclosure repricing
2029	$72B	$105B	$150B	Regulatory-driven procurement in compliance sectors
2030	$88B	$140B	$215B	Market maturation; model-layer concentration or not

6.2 Scenario analysis with named triggers

The forecast range is set not by a CAGR assumption but by three named, observable triggers. Each trigger is a discrete event with disclosed evidence, not a vibe.

Trigger	If observed	Effect on forecast	Observation window
Enterprise pilot-to-production conversion	Agentforce median deal size in top-quartile accounts crosses $250K by FY27 Q3	Base → Bull	Salesforce FY27 Q3, late 2026
Model-layer margin disclosure	First frontier-model IPO filing discloses >70% gross margin at frontier inference	Base → Bull	Anthropic S-1 filing, 2026–2027
Regulatory-driven procurement	EU AI Act high-risk compliance audits begin producing enforcement actions	Bear → Base	Late 2026 onwards
Compute-supply shock	Sustained data-centre capacity shortfall (lead time >36 months for 1 GW)	Base → Bear	Ongoing; primary signal is hyperscaler capex guidance
Large public-company write-down	Tier-1 enterprise discloses >$100M AI project write-down	Base → Bear	Any quarter

The slow-adoption risk

If enterprise adoption follows the ERP procurement pattern — expensive, slow, frequently abandoned, and subject to IT budget cycles — the 2030 market lands closer to $90B than $140B. The Agentforce average contract value of $28K is still a pilot metric, not a production metric. Watch Salesforce's FY27 Q3 and Q4 results: if deal sizes in the top quartile of accounts are not growing materially, the conversion rate from pilot to production is not happening, and the 2027–2028 forecast needs to come down.

Trigger	If observed	Effect on forecast	Observation window

The asymmetry here is important. The slow-adoption risk is real but bounded — the foundation-model layer grows regardless, because token consumption grows regardless. The application-layer forecast is the variable that moves; the total market forecast is less sensitive than the application-layer forecast suggests.

7. Geographic Distribution

Of the 564 companies in the Information Matters database, 380 (67%) are headquartered in the United States, 28 in the UK (5%), 23 in India, 19 in France, 14 in Germany, and 10 each in Singapore and Canada. The US dominance is structural: the foundation-model providers, the major enterprise-platform vendors, and the leading coding agents are all US-headquartered.

7.1 United States

The US story is revenue concentration without vendor concentration. All three leading foundation-model providers (Anthropic, OpenAI, Google), the largest enterprise agent platforms (Salesforce, ServiceNow, Microsoft, Palantir), and the leading coding agents (Cursor, GitHub, Anysphere) are US-headquartered — but the 184 infrastructure companies and 126 vertical specialists are distributed across dozens of metros. The M&A flow into 2027 will be dominated by US acquirers, and the companies that have chosen not to raise at 2024–2025 peaks will have structural optionality that later-raisers will not.

7.2 Europe and the United Kingdom

The regulatory environment is creating genuine opportunity for European-headquartered vendors. GDPR data-residency requirements, the EU AI Act, and sector-specific regulations in finance and healthcare are driving enterprise procurement toward vendors that can offer compliance guarantees. This is a structural advantage that no amount of superior technology from US providers fully overcomes in regulated European markets. Mistral is the most visible beneficiary at the model layer, but the pattern extends to the application layer: French, German, and UK vendors in legal tech, financial services AI, and public-sector AI are winning deals that US vendors are not credibly bidding for.

The UK's 28-company cohort is weighted toward coding tools, financial services AI, and AI infrastructure. The London ecosystem has produced disproportionate output relative to its funding footprint; the principal risk is that the best-positioned UK vendors get acquired by US acquirers before they reach IPO scale.

7.3 India and Asia-Pacific

India warrants specific attention. With 23 database entries and a large developer population, India is a significant market for coding agents — and an underappreciated one in most market analyses. The JetBrains survey shows strong AI tool adoption among Indian developers. As Indian IT services firms (Infosys, Wipro, Cognizant, TCS) embed agentic AI into their delivery models, India becomes both a major consumption market and a major deployment vector for global enterprises. The services firms are the sleeper distribution channel for enterprise agentic AI: they decide whether agentic deployments happen at a large fraction of the Global 2000.

Singapore is a hub-effect outlier: 10 database entries is disproportionate to population, reflecting government AI strategy and regional financial-services positioning. Japan and South Korea are underrepresented in the database relative to their economies and will likely produce more visible entrants through 2027.

8. Signals to Watch by Quarter

The quarterly cadence for the most useful primary-source disclosures to track. Each item is an observable event, not an expectation.

Quarter	Disclosures to watch
Q2 2026 (current)	Anthropic ARR update at mid-year; Salesforce Q1 FY27 first full-quarter AWU disclosure; JetBrains mid-year developer survey refresh; first hyperscaler capex guidance update of 2026.
Q3 2026	Anthropic potential S-1 filing (watch Wilson Sonsini mandate disclosures); Palantir Q2 2026 US commercial growth rate; ServiceNow Q2 2026 $1M+ ACV cohort growth; a16z Enterprise AI 2026 report (annual).
Q4 2026	Salesforce FY27 Q3 earnings — first credible read on Agentforce top-quartile median deal size; Microsoft FY27 Q1 Copilot seat count and revenue breakout (if any); first voice-agent vendor to disclose $100M+ ARR.
Q1 2027	Salesforce FY27 Q4 (full year); first coding-agent IPO filing disclosing gross margin; first named $50M+ voice-agent enterprise contract.
Q2 2027	EU AI Act first enforcement action (if any); Anthropic potential IPO pricing; first large public-company AI project write-down (if any).

9. Twelve to Watch — Q2 2026

This is a curated quarterly shortlist of twelve smaller companies drawn from the 564-company Information Matters database. It deliberately excludes the names already covered in §5 (Anthropic, OpenAI, Google, Salesforce, ServiceNow, Microsoft, Palantir, IBM, Cursor, GitHub Copilot, Claude Code, JetBrains). The criteria are: disclosed annual revenue under ~$50M or headcount under ~300 (or both), evidence of a defensible position in a segment we expect to consolidate, and — where possible — geographic or positioning diversity. This list will be

refreshed every quarter; names that scale past the threshold will graduate off it, and names that fail to convert on a named signal will be removed.

The picks cluster on purpose into three categories where we think the 2026–2027 consolidation dynamics will reward specialists over generalists: evaluation and observability (three picks), voice-agent infrastructure (four picks, including one real-time primitive), and workflow/vertical specialists (five picks). This is the same bet we made in §5.4; the twelve are the concrete names behind it.

The "Source" column flags how each scale figure is grounded: Disclosed where a company earnings release or funding-round disclosure backs the number; Press where a named outlet has reported it; Library where the figure is taken from the Information Matters internal database (with confidence high or low as recorded); Private where the company has made no revenue disclosure and we cite only headcount or status.

9.1 Evaluation, observability and AI gateway

Company	HQ	Scale	Why it's on the list	Source
Langfuse	Germany	~13 emp, OSS	Open-source LLM engineering and observability (traces, evals, prompt management). Open-source distribution plus European regulatory positioning is a credible wedge into the eval/obs consolidation we flag in §5.4. The question is whether they can convert OSS adoption into managed-service revenue at the pace needed to stay independent.	Private
HoneyHive	United States	~11 emp	Eval and observability built for agentic workloads from day one, rather than retrofitted from ML monitoring. Small but positioned in the segment we argue is currently underfunded relative to its strategic importance. Watch for a Series A or B in H2 2026.	Library (est.)
Portkey	United States	~11 emp	Control-plane thesis: gateway + observability + guardrails as a single surface. If the eval/obs category consolidates toward platforms rather than point tools, Portkey's framing is ahead of the curve. Needs to either raise into the consolidation or win a defining enterprise reference account.	Library (est.)

9.2 Voice-agent infrastructure

Company	HQ	Scale	Why it's on the list	Source
Vapi	United States	~50 emp, $1M FY25	Developer-first voice-agent platform. In the voice category where we argue one vendor will own the naturalness+latency+multilingual combination, Vapi's developer positioning mirrors Twilio's playbook for SMS — a repeatable pattern in this category.	Press (Forbes)
Retell AI	United States	Founded 2023; rev reported in press	Positioned as one of the furthest-along voice-agent-platform vendors. The library-recorded FY25 revenue figure needs verification against a primary disclosure before carrying weight; we are watching for a clean 2026 mid-year ARR update.	Library, unverified
Gladia	France	~30 emp	Real-time speech-to-text and audio-intelligence APIs with an EU data-residency story. The voice-layer vendor most likely to win regulated-sector European procurement that US-headquartered vendors cannot credibly bid for. Ties directly to the §7.2 European opportunity.	Private
LiveKit	United States	$10M FY24	Real-time communications infrastructure that underpins voice agents at several larger players. Less a voice-agent vendor than a primitive the category depends on — which is a position worth watching because primitives capture a disproportionate share of the value when a category scales.	Library

9.3 Workflow orchestration and vertical specialists

Company	HQ	Scale	Why it's on the list	Source
n8n	Germany	~100 emp, ~$40M	Developer-first workflow automation with strong agentic-node capabilities. In a segment we expect to shrink (§5.4), n8n is the most credible survivor candidate — OSS-led, developer-loved, European-headquartered, monetising on cloud and enterprise tiers.	Library (low-conf est.)
Tines	Ireland	~200 emp, ~$20M+	No-code workflow automation for security and IT teams. Vertical specialisation — security operations — is exactly the defensibility the orchestration layer needs to survive incumbent absorption. Worth watching alongside n8n as the two most-likely-to-survive workflow specialists.	Press (Bloomberg)

Company	HQ	Scale	Why it's on the list	Source
Decagon	United States	~100 emp, ~$20M+	Enterprise-grade customer-operations agent with a short list of large-logo references. The CX agent category is the most mature horizontal use case (§4.2); Decagon is the clearest independent alternative to Salesforce Agentforce and ServiceNow Now Assist in enterprise deals.	Library (low-conf est.)
HappyRobot	United States	2023, $10M FY24	Logistics and supply-chain agents. Supply chain is the sleeper vertical we flagged in §4.3 — complex enough to need specialist AI, clear ROI, less regulated than Legal or Health. Early revenue disclosure at $10M is a credible starting position.	Library
Spellbook	Canada	Private	Contract drafting and review for mid-market law firms and in-house legal teams. Legal is only five companies in the database (§4.3) but is a high-value vertical with durable moat characteristics once compliance positioning is built. Spellbook is the most quietly credible of the non-Harvey legal specialists.	Private

How to read this list

This is not a ranking and not a set of investment recommendations. It is a working list of companies whose next disclosed data point (an ARR update, a named customer, a funding round, a product launch) will move our read on the segment they sit in. We expect three to four of the twelve will be acquired or consolidated out of this list within twelve months; one or two will scale past the criteria and graduate to §5; the remainder will either convert their position or be replaced.

Each entry above names the specific next signal we will be tracking — a funding round, an ARR update, a customer reference, a product release. If a pick hits its signal by the next quarterly update (Q3 2026), it stays on the list; if it misses, it rotates off. The Q3 2026 refresh will also re-query the Information Matters database for any companies that have crossed the coverage threshold since this edition.

10. Conclusions and Recommendations

10.1 For technology buyers

- Coding agents are the lowest-risk starting point. ROI is measurable, developer adoption is already happening, and the governance conversation is more tractable than for autonomous enterprise agents. Get ahead of the policy question — the 41% policy-coverage rate from the April 2026 JetBrains survey says most employers are still behind.

- In customer operations, the technology is ready. The risk is not the agent — it is the integration with your existing ticketing, CRM, and knowledge systems. Budget two-thirds for integration, one-third for licences.
- Evaluation and observability is not optional. If you are running agents in production without visibility into what they are doing, you are not managing risk — you are hoping. Spec this into your Agentforce, Now Assist, or Copilot procurement from day one.
- On pricing, insist on usage-based terms with a committed floor and a renegotiation trigger. The foundation-model price curve is falling 60–80% every 18 months. Multi-year fixed-seat commitments at 2026 pricing will look expensive in 2027.

10.2 For investors

- The evaluation and observability category will produce a very large business. It is currently fragmented and valued below its strategic importance.
- The voice-agent infrastructure category is next. Latency thresholds are approaching human turn-taking; the first vendor to combine latency, naturalness, and multilingual scale will own a category.
- The foundation-model layer is concentrating around Anthropic, OpenAI, and Google. New entrants need a capability breakthrough, a compliance angle, or a sovereignty play. The Anthropic margin advantage, if confirmed at IPO, repriced the category.
- Be cautious about horizontal agent builders without defensible specialisation. The incumbents are coming, and the absorption window is 9–12 months, not 24–36.

10.3 For vendors

- The LinkedIn headcount methodology used in this report is being used by investors and acquirers to estimate your revenue. If your headcount-to-revenue ratio is anomalous, expect to be asked about it in diligence.
- Regulation is coming. Build compliance in now. The companies that do so will have a durable competitive advantage in European and regulated-sector markets.
- The pilot-to-production conversion problem is the market's central challenge in 2026–2027. The vendors who solve it — with faster time-to-value, clearer ROI measurement, and lower integration cost — will define the next wave of enterprise AI adoption.
- Disclose gross margin before you are forced to. The market is going to learn about agentic-AI unit economics in 2026–2027 whether vendors lead that conversation or not. Leading it is a positioning advantage.

10.4 Our view

Three calls, stated plainly. First, the Anthropic run-rate trajectory is not a bubble; it is the fastest enterprise sales cycle in software history meeting the fastest-compounding compute footprint in software history. Both sides are disclosed primary data. The v3 upward revision reflects that, and the attribution rate will be higher — not lower — in the v4 revision.

Second, the Salesforce Agentforce narrative is two stories, not one. The 2.4B AWU disclosure is production-scale usage inside the deployed accounts; the $28K average contract value is pilot-scale commitment across the book. Both are true. The market is going to learn in late 2026 which story dominates. The FY27 Q3 median-deal-size disclosure is the single highest-signal data point in 2026 — more consequential than any foundation-model number.

Third, the category where the market is most mispriced is evaluation and observability. The incumbents are building in telemetry, but production-grade eval tooling for agentic AI requires independent third-party instrumentation in the same way financial audit does. The first eval/observability vendor to raise at a $3B+ valuation will look cheap in hindsight. That is a position, not a hedge.

Methodology note

This report was produced through a combination of human expertise and oversight supported by an AI research agent specifically designed to carry out detailed market research and forecasts using established methodologies and with access to the latest data and AI models as well as our proprietary database of over 500 companies engaged in the agentic AI space. This research consists of the opinions of Information Matters' research team, human and AI, and the information contained within it should not be considered as statements of fact. None of the information presented here should be taken as investment advice. Reproduction or distribution of this research without written permission from Information Matters Ltd is prohibited.