David Soden

MARKET ASSESSMENT | AUGUST 2026

12,000 Agents Built, 300 Ever Shared

What Salesforce, Workday, Microsoft and ServiceNow actually count when they report agent adoption, why none of it is filed, and how far the numerators sit from a working system

David Soden

Independent analyst and practitioner, enterprise automation and applied AI

davidsoden.com/market-assessment

REPORT DS-2026-977F · 2026 · VERSION 1.0 · PUBLIC

About this report

Every substantive claim in this report carries an evidence classification and, where applicable, a numbered source. Facts are separated from vendor claims, third-party estimates, the author's assessments, and untested hypotheses. Forward-looking sections use scenarios with observable tripwires rather than point forecasts. Stated assumptions are listed before the analysis begins.

Market Assessment reports are published and commissioned at **davidsoden.com/market-assessment**. This document is analysis, not investment, legal, or procurement advice.

Scope and evidentiary standard

This assessment covers the operating metrics that enterprise software vendors used to describe agentic AI adoption in calendar 2026 earnings materials, together with the platform operators publishing agent creation counts. Eight disclosures are examined in detail: Salesforce, Workday, ServiceNow, Microsoft, Okta, nCino, IFS, and Singapore's national healthtech agency Synapxe. The question is narrow and mechanical. Not whether agents work, but whether the published numbers describing their spread mean the same thing across vendors, and whether they support the peer comparisons that CIOs, investors and procurement teams are currently building on top of them.

Where a definition exists, it is quoted as the vendor printed it. Where none exists, that absence is stated rather than filled in. Metric definitions were checked against the filings themselves using the SEC EDGAR full-text search index, so a claim that a term appears in no periodic report is a search result rather than an impression.

Two discounts apply throughout. First, agent counts are unaudited operating metrics with no accounting standard behind them, so period-to-period comparability rests entirely on the vendor choosing not to change the definition. Second, several of the survey figures used to establish how much autonomy enterprises actually permit were commissioned by companies selling agent software. Those are labeled as third-party estimates and their sponsors are named in the text, not only in the source list.

Label	Definition	How to treat it
FACT	Verifiable in the cited primary source: a filing, press release, official documentation, or standards body announcement.	Reliable as stated. Check the source if the exact wording matters.

Label	Definition	How to treat it
VENDOR CLAIM	Reported by a vendor about its own business or product. Self-measured and not audited by any third party.	The vendor said it. Directionally useful, not a measurement. Definitions of the metric are usually undisclosed.
THIRD-PARTY ESTIMATE	A research firm forecast or survey result. Methodology is proprietary and generally not reproducible.	Use for direction and order of magnitude only. Do not build a model on it.
ASSESSMENT	The author's reasoned judgment from the cited evidence. Falsifiable, and the reasoning is shown so it can be argued with.	This is analysis, not reporting. Disagree with it where your evidence differs.
HYPOTHESIS	A proposition offered for testing. Not currently supported by sufficient evidence to assert.	Treat as a research question, not a conclusion. Each one names what would confirm or refute it.
SCENARIO	A structured future state with a stated subjective probability. The probability is the author's judgment, not a calculated figure.	Use the leading indicators attached to each, not the probability. The indicators are observable; the probability is not.

Assumptions this analysis rests on

Four premises are assumed rather than demonstrated. Each is followed by what fails if it turns out to be wrong.

That the published figures are arithmetically accurate within their own definitions. Nothing here alleges misstatement, and the internal reconciliation checks performed on Salesforce's task counts held. If a vendor's underlying count is simply wrong, the disclosure-quality argument survives but the specific comparisons in the analysis do not.

That vendors who publish no definition genuinely have no published definition, rather than one buried in a document not reached by full-text search or by direct retrieval. Definitions were sought in press releases, 8-K exhibits, and the EDGAR full-text index. Investor day decks and prepared remarks not filed with the Commission were not exhaustively swept. If a definition exists there, the disclosure grades understate several vendors, though the argument that the definitions are not comparable across vendors is unaffected.

That the Singapore AgentSea creation and sharing figures are representative of the general pattern on self-service agent-building platforms, rather than an artifact of a healthcare workforce experimenting under an explicit mandate to explore. This is the weakest assumption in the document. If AgentSea is unrepresentative, the specific ratio in the title does not generalise, and the second finding weakens to a single well-documented case rather than a rule.

That the trust surveys measuring permitted autonomy describe roughly the same population as the vendors' customer bases. Both surveys skew toward larger enterprises, which is where the vendors sell, so the overlap is probably reasonable. If enterprises granting full autonomy are systematically absent from survey panels, the third finding overstates the constraint.

The call

ASSESSMENT The agent adoption numbers in this year's earnings materials are not comparable, and most were not built to be. Each vendor counts a different object: a registry entry, a customer with one thing switched on, a discrete task, or contract value tagged to a bundle the customer did not separately buy. Where a platform publishes both what was created and what was reused, the reuse figure is a small fraction of the headline. Treat every agent count as marketing until the vendor publishes a definition, a denominator, and a stable composition. What would move me is a vendor footnoting its agent metric in a filing and holding that definition through a quarter of decelerating growth.

My judgment on the evidence assembled here, nothing more. There is data I can't see and data I may have missed. New evidence can move me, including to the opposite position, and I reserve the right to move.

Executive summary

Agentic AI reporting matured fast in 2026. Vendors moved from anecdote to number, and several now publish figures with more precision than the category deserves. What did not mature is the definitional layer underneath. The result is a set of headline numbers that look like each other, are quoted next to each other, and measure different things.

FINDING 1 Not one of these agent metrics appears in a filed periodic report.

> FACT A full-text search of the SEC EDGAR index returns Salesforce's term "Agentic Work Unit" in exactly three documents, all Exhibit 99.1 earnings releases attached to Form 8-K. "Agentforce and Data 360 annual recurring revenue" returns six, four of them the same 8-K exhibits plus two proxy documents. Workday's phrase "organic agents" appears in its Q2 8-K exhibit. None appears in a 10-Q or a 10-K. The Salesforce 8-K cover states that the information furnished under Item 2.02 "shall not be deemed 'filed' for purposes of Section 18 of the Securities Exchange Act of 1934, as amended, or otherwise subject to the liabilities of that section." [3] [4] [6] [1]

FINDING 2 Where a platform publishes creation and reuse side by side, reuse is 2.5 percent of creation.

> FACT Synapxe, Singapore's national healthtech agency, reported that more than 12,000 AI agents were created on its AgentSea platform between the end of May and late August 2026, and that more than 300 of those were made available to other users for reuse. The eligible population is roughly 80,000 public healthcare professionals. This is the only disclosure in the study that publishes a numerator, a reuse figure, and a population together. [8]

FINDING 3 Microsoft's registry count sits two orders of magnitude above any measured deployment figure.

ASSESSMENT Satya Nadella told the FY26 Q4 call that "nearly 40 million agents registered across tens of thousands of companies within two months of launch." Taking the midpoint of "tens of thousands" as roughly 50,000 companies implies about 800 registered agents per company. Salesforce's own production telemetry puts the figure at 13 activated agents per business as of April 2026. Registration and activation are not the same event, and Microsoft does not say which it counts. [7] [16]

FINDING 4 Salesforce changed what Agentforce ARR contains while reporting its growth rate.

FACT The Q2 FY27 release states: "Beginning in Q2 FY27, Agentforce ARR includes our AI offerings Slackbot and Headless 360." The same release reports Agentforce ARR exceeding $1.5 billion, up over 240 percent year over year. A growth rate spanning a composition change is not a like-for-like comparison, and the release does not quantify the contribution of the newly included products. [3]

FINDING 5 ServiceNow's AI ACV became a packaging outcome in April 2026.

ASSESSMENT On 9 April 2026 ServiceNow replaced five legacy tiers with three, and Now Assist, the Moveworks layer, Workflow Data Fabric and AI Control Tower were bundled into every tier rather than sold as add-ons. Once AI ships inside the base SKU, contract value tagged to AI reflects a pricing decision by the vendor rather than a purchasing decision by the customer. The $1 billion AI ACV milestone reported for Q2 2026 spans that change. [19] [20]

FINDING 6 Workday admits a cull and publishes no count that could reflect it.

FACT CEO Aneel Bhusri told diginomica: "We had a lot more agents when I came back. We killed a lot or we rolled them into bigger agents." Workday's published metric is customers using one or more agents, which is structurally incapable of falling when agents are retired or merged. The company discloses neither the starting agent count nor the number removed. [2] [1]

FINDING 7 The operational metrics describe a world in which roughly one enterprise in twenty permits full autonomy.

THIRD-PARTY ESTIMATE Harvard Business Review Analytic Services, in a survey of 603 leaders sponsored by Workato and AWS, found 6 percent fully trust AI agents to run core business processes autonomously, with 43 percent limiting them to routine tasks and 39 percent to supervised or noncore use. Research by The Futurum Group commissioned by IFS put the equivalent figure at 5.7 percent. Both sponsors sell agent software, which makes the low number harder to dismiss rather than easier. [9] [10] [11]

FINDING 8 Two vendors in the set disclose no agent metric at all, and that is the more defensible position.

FACT Okta's Q2 commentary described "several AI-related deals in the quarter with contract values above the corporate average" without an AI-specific ARR, customer count or transaction volume. nCino reported four US enterprise customers with over $900 billion in combined assets renewing early with "expanded AI commitments" and quantified none of it. Neither published a number that could be misread, because neither published a number. [13] [12]

What the headline numbers actually count

The brief's first question was what is being counted. The answer is that at least five different objects are in circulation, and the vendors reporting them rarely say which one they have chosen.

FACT Salesforce is the only vendor in the set that publishes a definition of its unit of agent work. The Q2 FY27 release states: "The Company defines an Agentic Work Unit ('AWU') as a measure of discrete tasks executed by AI agents in production across the Salesforce platform, including Agentforce and Slack." The Q1 release adds that AWUs "represent the conversion of generative AI capabilities into measurable business outputs, such as resolving customer cases, updating records or triggering automated workflows." [3] [5]

ASSESSMENT That definition is the best in the category and it is still doing less work than it appears to. Resolving a customer case and updating a record are separated by two or three orders of magnitude in commercial value, and both count as one AWU. A vendor could double its AWU total by routing more record updates through an agent without a single additional case being resolved. The metric is honest about what it measures and silent about what a unit is worth.

FACT Workday counts customers, not agents. Its Q2 FY27 8-K exhibit reads: "More than 5,500 customers now use one or more of Workday's organic agents, up more than 35% from last quarter." The word "organic" is load-bearing, since it appears to exclude agents from acquisitions and partners, and it is nowhere defined in the release. There is no footnote attached to the metric; the only footnote marker in that section points at the non-GAAP financial measures discussion. [1]

ASSESSMENT Workday reports more than 11,000 organizations as customers, so 5,500 customers with at least one agent is roughly half the installed base. The vendor did not publish that denominator alongside the numerator, and the ratio is the reader's arithmetic rather than the company's disclosure. A metric whose threshold is one is close to a measure of feature availability. It rises when a customer switches something on and cannot fall when the customer stops using it. [1] [21]

FACT Microsoft counts registrations. On the FY26 Q4 call of 29 July 2026, Nadella reported "nearly 40 million agents registered across tens of thousands of companies within two months of launch," referring to Agent 365. The transcript contains no definition of an agent and no definition of registration. [7]

ASSESSMENT A registry entry is the cheapest possible unit. It requires no execution, no user, and no business process. It is the correct thing for Microsoft to track internally, because governing agents is the product, and it is the wrong thing to place next to a competitor's production or revenue figure. Nothing in the disclosure prevents an agent registered, tested once and abandoned from remaining in the count indefinitely.

FACT Workday and ServiceNow both report contract value tagged to AI. Bhusri's quote is "AI driving more than 25% of our new ACV." ServiceNow reported AI ACV crossing $1 billion in Q2 2026 with net new AI ACV up more than 40 percent quarter over quarter, and raised its AI ACV target from $1 billion to $1.5 billion. Neither company publishes a definition of what makes a dollar of contract value an AI dollar. [1] [20]

ASSESSMENT Tagged contract value is the metric most exposed to vendor discretion, because the tagging happens inside the vendor and is never seen by the customer. When a customer buys a suite that includes agents, the allocation between the AI line and the base line is an internal decision with no external check. This is the metric an investor should discount hardest and the one that most often anchors a valuation multiple.

Vendor and metric	Definition the vendor gives	Denominator	Disclosure quality
Salesforce, Agentic Work Units	Published verbatim: "a measure of discrete tasks executed by AI agents in production across the Salesforce platform, including Agentforce and Slack"	None. No customer count, no task baseline, no per-customer split	Best in set. Definition published, figures reconcile across quarters, but only in a furnished 8-K exhibit
Salesforce, Agentforce ARR	Blended definition covering Data 360 plus "certain generative AI subscription agreements," with a separate Agentforce-only carve-out given on the call	None. Composition changed in Q2 FY27 to add Slackbot and Headless 360	Poor. A growth rate reported across a composition change, with the added products unquantified
Workday, customers using agents	None. "One or more of Workday's organic agents" is not defined, and "organic" is not explained	Not published. Roughly 11,000 customers, so about half the base	Poor. Threshold of one, cannot decline when agents are retired
Workday, AI share of new ACV	None. Neither "AI" nor "new ACV" is defined in the release	New ACV in the period, itself undisclosed	Poor. Percentage of an undisclosed base, tagged at the vendor's discretion
ServiceNow, AI ACV	None located in the earnings materials	Total ACV, not published alongside	Poor. Now Assist bundled into every tier from 9 April 2026, so the metric spans a packaging change
Microsoft, agents registered	None. Neither "agent" nor "registered" is defined on the call	"Tens of thousands of companies," a range spanning a factor of five	Weak. A registry count with no execution, retirement or activity threshold
Synapxe, AgentSea agents created	Implicit but unambiguous: agents created by professionals, and the subset published for reuse	Roughly 80,000 eligible public healthcare professionals, published	Strongest in set. Numerator, reuse figure and population all disclosed together
IFS, agentic transactions automated	"Digital workers" defined as agents that "autonomously monitor, decide, and execute operational tasks,	Agentic transactions on the IFS platform, count not published	Weak. Ratio published, base withheld, survey

Vendor and metric	Definition the vendor gives	Denominator	Disclosure quality
	escalating to humans only when judgment is required"		sample size undisclosed
Okta	No agentic metric disclosed. Commentary references "several AI-related deals" above average contract value	Not applicable	No metric to grade, which is more defensible than a bad one
nCino	No agentic metric disclosed. Four renewals with "expanded AI commitments," unquantified	Not applicable	No metric to grade, which is more defensible than a bad one

The gap between the numerator and a working system

The second question was how far the headline number sits from something a business depends on. Almost no vendor publishes the pair of figures needed to answer that. One organisation does.

FACT Synapxe launched AgentSea to Singapore's public healthcare workforce at the end of May 2026. By late August the agency reported more than 12,000 agents created and more than 300 made available to other users for reuse and adaptation. The platform is open to more than 80,000 public healthcare professionals. [8]

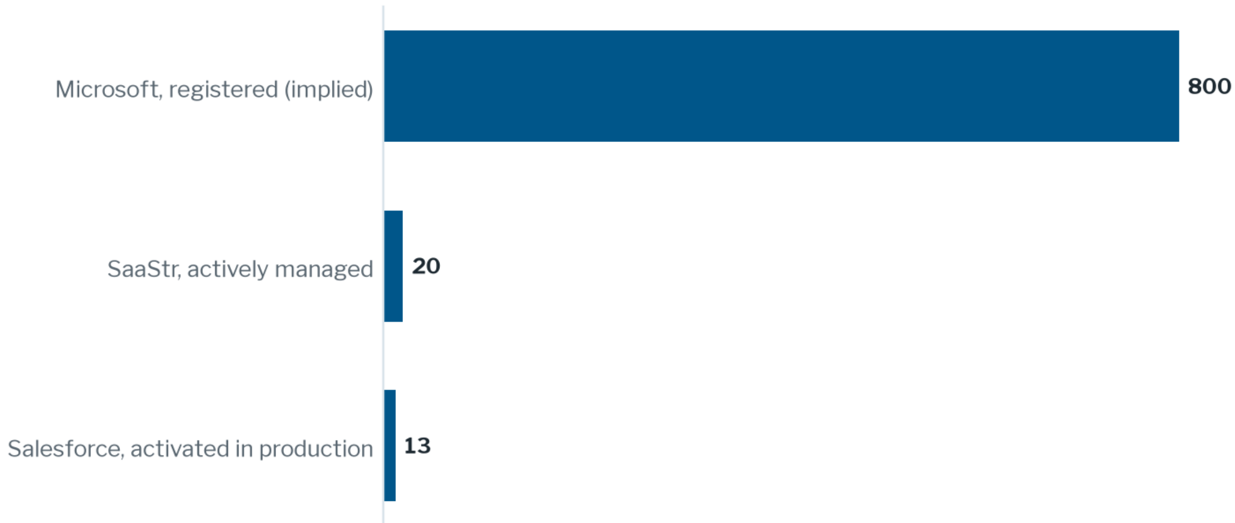
FIGURE 1 **Agents created versus agents shared for reuse, AgentSea, May to August 2026**

FACT Figures as published by Synapxe, Singapore's national healthtech agency. Sharing is a deliberate act by the creator, so the 300 understates agents that are used privately and never published. It is nonetheless the only reuse figure any operator in this study discloses. [8]

ASSESSMENT Two and a half percent of what was built was offered to anyone else. Sharing is not the same as running in production, and an agent used privately by its creator every day would not appear in the 300. But sharing is the cheapest possible proxy for an agent being good enough that a second person should have it, and it is a threshold that ninety-seven and a half percent of these agents did not clear. Read against the eligible population, the platform produced roughly one shared agent for every 267 professionals it was offered to. [8]

ASSESSMENT The general shape is corroborated by weaker evidence. A practitioner survey run by LangChain over 1,340 self-selected respondents in late 2025 reported 57.3 percent with agents in production, but 63 percent of that panel came from the technology sector and 49 percent from organisations under 100 people, which is not the enterprise population these vendor metrics describe. It is directionally useful and quantitatively unusable, and no conclusion here rests on it. [18]

FIGURE 2 Agents per company, by what the source counts



ASSESSMENT The Microsoft figure is the author's arithmetic: nearly 40 million agents divided by a midpoint reading of "tens of thousands of companies" at 50,000. Microsoft published neither the company count nor a definition of registration, so the bar could reasonably sit anywhere between 400 and 2,000. The two smaller figures are measured, not implied. [7] [16] [17]

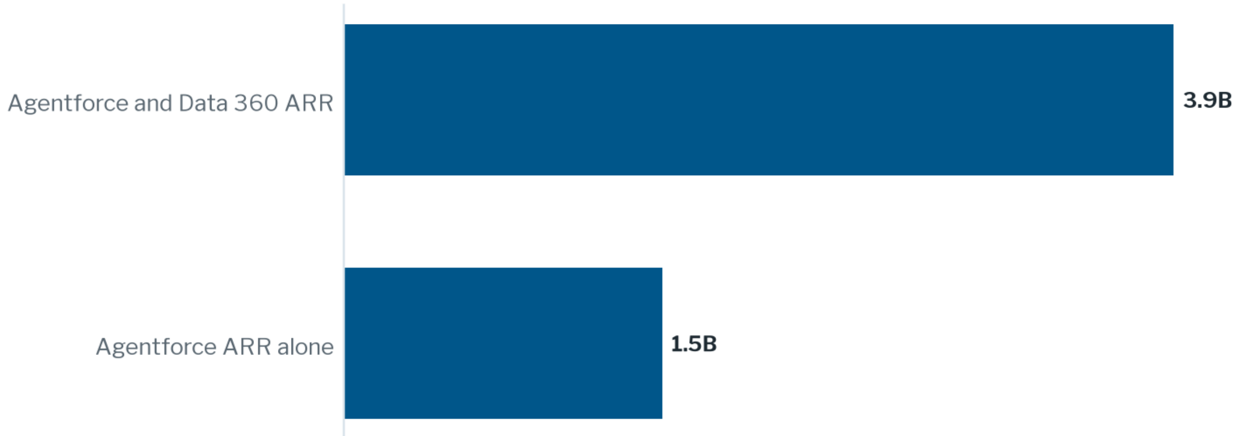
ASSESSMENT The distance between the bars is the finding. Registration, active management and production activation are three different states, and the population of agents shrinks by roughly two orders of magnitude between the first and the last. A CIO comparing a Microsoft number to a Salesforce number is comparing a registry to a runtime. [7] [16]

Definitions that moved while the number was growing

A metric can be well defined and still not be comparable to its own prior period. Salesforce provides the clearest example in the category, and to its credit it disclosed the change in the same document as the growth rate.

FACT The Q2 FY27 release defines the ARR metric as "the annualized recurring value of active Data 360 and certain generative artificial intelligence ('AI') subscription agreements, including those for Agentforce, generative AI products and features, and Headless 360, that were executed at the end of the reporting period," and states separately that "Beginning in Q2 FY27, Agentforce ARR includes our AI offerings Slackbot and Headless 360." The combined Agentforce and Data 360 figure reached nearly $3.9 billion, up over 210 percent year over year. Agentforce ARR alone exceeded $1.5 billion, up over 240 percent. [3]

FIGURE 3 Salesforce Q2 FY27: the metric with a written definition versus the one quoted in headlines

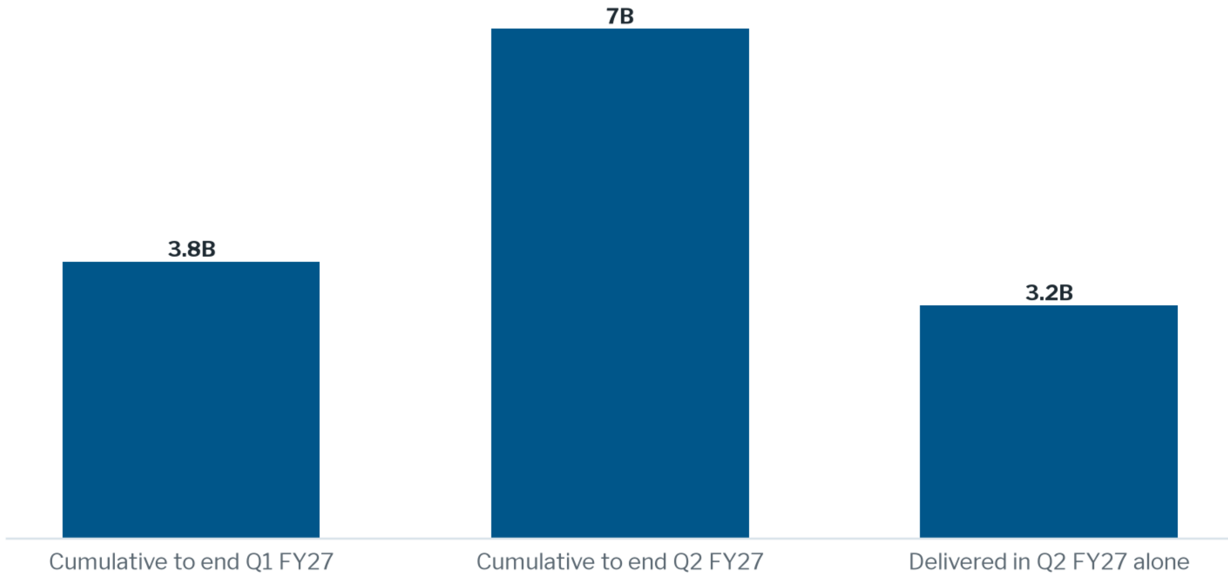


VENDOR CLAIM Salesforce self-reported, unaudited, furnished under Item 2.02 rather than filed. The larger figure carries the published definition; the smaller one is the number most press coverage quotes. Agent-specific ARR is 38 percent of the defined metric. [3]

ASSESSMENT Two problems compound. The metric that carries a written definition is the blended one, and the metric that gets quoted is the carve-out. Then the carve-out absorbed two products mid-year while a 240 percent growth rate was reported over it. Salesforce disclosed the change plainly, which is more than any other vendor here did, and still did not quantify what Slackbot and Headless 360 contributed. Without that figure the growth rate cannot be decomposed into demand and definition. [3]

FACT The task metric holds up better. Salesforce reported 3.8 billion AWUs delivered to date at Q1 FY27 and 7.0 billion at Q2 FY27, with 3.2 billion delivered in Q2 alone. Those figures reconcile: the difference between the two cumulative totals is exactly the in-quarter number. [3] [5]

FIGURE 4 **Salesforce Agentic Work Units, cumulative and in-quarter**



VENDOR CLAIM Self-reported and unaudited. The three figures are internally consistent, which is not true of every operating metric in this study. What they still do not carry is a denominator: no customer count, no tasks-per-customer split, and no baseline for what those tasks cost before an agent did them. [3] [5]

ASSESSMENT This is what a defensible agent metric looks like, and it is worth saying so in a document otherwise critical of the category. A published definition, a stated scope, cumulative and in-quarter figures that reconcile. The remaining gap is the denominator, and until a task count is expressed per customer or against a pre-agent baseline, a billion units is a number without a scale. [3]

FACT ServiceNow's change was to the product boundary rather than the metric. On 9 April 2026 the company replaced Standard, Pro, Pro Plus, Enterprise and Enterprise Plus with three tiers, Foundation, Advanced and Prime. Per ServiceNow's own community documentation, "Now Assist, the Moveworks layer, the Workflow Data Fabric, and the AI Control Tower are bundled into every tier instead of being sold as separate add-ons." Legacy SKUs reached end of sale on 1 July 2026. [19]

ASSESSMENT This is the most consequential disclosure problem in the set, because it is invisible in the metric. Before April, buying Now Assist was a decision a customer made and AI ACV was a demand signal. After April, every tier includes it, and AI ACV becomes an allocation of bundle revenue that ServiceNow performs internally. The $1 billion milestone spans both regimes. Nothing in the reporting lets an outsider separate the two, and a procurement lead being shown that number in a renewal conversation should ask which side of April the comparison sits on. [19] [20]

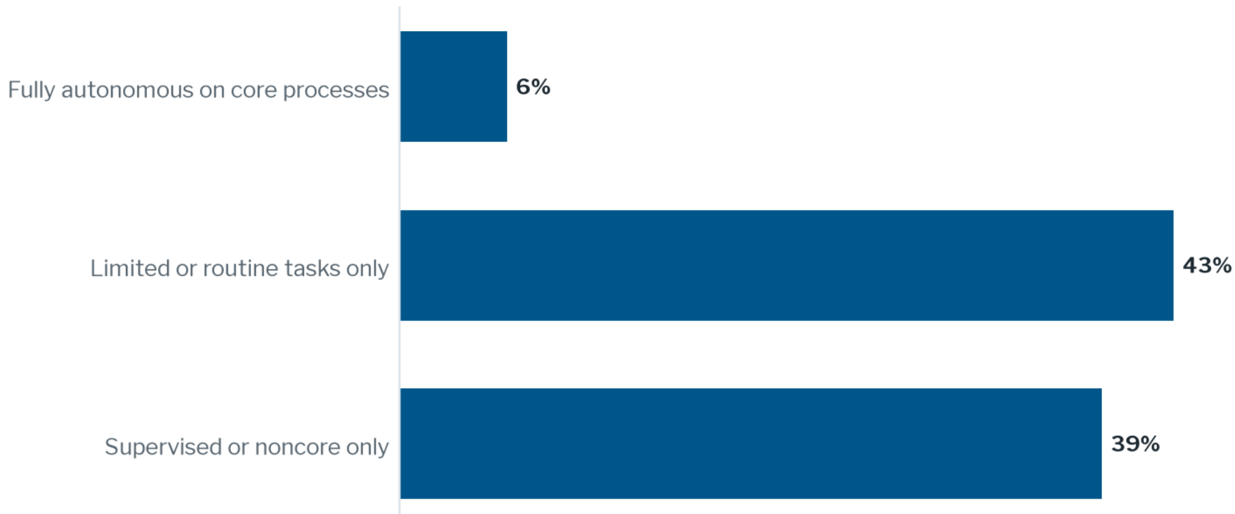
What the trust data does to the operational metrics

The third question was whether the adoption numbers survive contact with evidence about how much autonomy enterprises actually grant. They survive, but only by being reinterpreted as something narrower than they appear.

THIRD-PARTY ESTIMATE Harvard Business Review Analytic Services surveyed 603 business and technology leaders in July 2025, sponsored by Workato and Amazon Web Services, and published in December 2025. Six percent said they fully trust AI agents to autonomously run core business processes. Forty-three percent trust agents with only limited or routine operational tasks, and 39 percent restrict them to supervised use cases or noncore processes. [9]

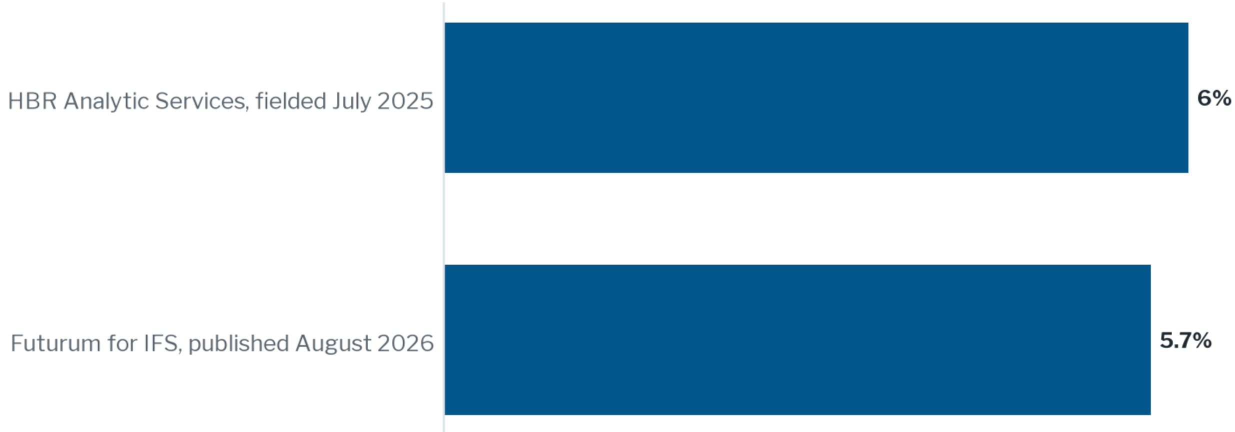
THIRD-PARTY ESTIMATE Separately, research by The Futurum Group commissioned by IFS and published in August 2026 found that 5.7 percent of decision makers trust AI to act fully autonomously, while 66 percent said they were likely to invest in agentic digital workers within the year. The sample size is not disclosed in the materials released publicly, which is a real limitation on a figure quoted to one decimal place. [10] [11]

FIGURE 5 How much autonomy leaders say they permit



THIRD-PARTY ESTIMATE Harvard Business Review Analytic Services, 603 business and technology leaders, fielded July 2025, sponsored by Workato and Amazon Web Services. Self-reported and sponsored by companies that sell agent software, which cuts against the low headline figure rather than for it. The three categories do not sum to 100 percent as published. [9]

FIGURE 6 **Two vendor-sponsored surveys, one year apart, on full autonomy**



THIRD-PARTY ESTIMATE Different panels, different sponsors, different wording, and one of the two does not publish its sample size. The convergence is suggestive rather than conclusive, and both sponsors had a commercial interest in a higher number than they reported. [9] [10]

ASSESSMENT The reconciliation is straightforward once the metrics are read correctly. Workday's 5,500 customers with at least one agent and ServiceNow's $1 billion of AI ACV are entirely compatible with 6 percent full autonomy, because neither metric claims autonomy. They count enablement and contract value. The problem is not that the numbers contradict the trust data. It is that they are routinely presented as evidence of a transition to autonomous operation that the trust data says is not happening. [1] [9] [20]

VENDOR CLAIM IFS is the one vendor in the set that publishes the split directly. It states that its platform automates 60 percent of agentic transactions end to end while the remaining 40 percent require human review. That is a genuinely useful disclosure shape, and it is undermined by the absence of the transaction base, which means the ratio can be reported stably while the volume behind it moves in either direction. [10]

ASSESSMENT Gartner's forecast that more than 40 percent of agentic AI projects will be canceled by the end of 2027 belongs here rather than in the scenarios, because it describes the same constraint from the project side. Gartner also estimates that only around 130 of the thousands of vendors marketing agentic AI are genuine, the rest engaged in what it calls agent washing. Both are analyst judgments with proprietary methodology behind them and should be weighted accordingly. Set against Gartner's own forecast that supply chain software with agentic AI grows from under $2 billion in 2025 to $53 billion by 2030, the firm is simultaneously projecting rapid category spend and a high project failure rate, which is coherent only if a large share of that spend buys something that does not survive. [14] [15]

Culls, consolidation, and the counts that do not reflect them

The fourth question was which vendors have retired agents, and whether the headline count reflects it. One vendor has said so on the record. Its metric cannot show it.

FACT Aneel Bhusri, on returning as Workday CEO, told diginomica: "We had a lot more agents when I came back. We killed a lot or we rolled them into bigger agents." Neither the starting count nor the number removed was disclosed. In the same period Workday reported more than 5,500 customers using one or more organic agents, up more than 35 percent from the prior quarter. [2] [1]

ASSESSMENT A count of customers with at least one agent is structurally immune to a cull. Retire nine of a customer's ten agents and the customer still counts. Merge four agents into one and the count is unchanged. Workday can be simultaneously right that adoption is rising and right that it deleted a large fraction of its catalogue, and a reader has no way to see the second event in the first number. That is not deception. It is a metric that was chosen before the cull and never revisited afterwards. [1] [2]

ASSESSMENT The honest reading of the Workday disclosure is more favourable to Workday than the metric is. A vendor that publicly says it killed agents and folded them into larger ones is describing a real product decision that most of the category has not admitted to. The failure is in the reporting, which gives that decision nowhere to appear. [2]

ASSESSMENT Customer-side evidence points the same way, from a much weaker source. Jason Lemkin of SaaStr, writing publicly in July 2026 about his own eight-figure business, described going from roughly 30 agents to 20 and reporting output up around fourfold after the cut, citing a management ceiling rather than agent quality. This is a single self-reported practitioner account with an obvious selection effect and no audit, and it is included because it is one of the few first-person accounts of a deliberate reduction. It supports the pattern; it establishes nothing on its own. [17]

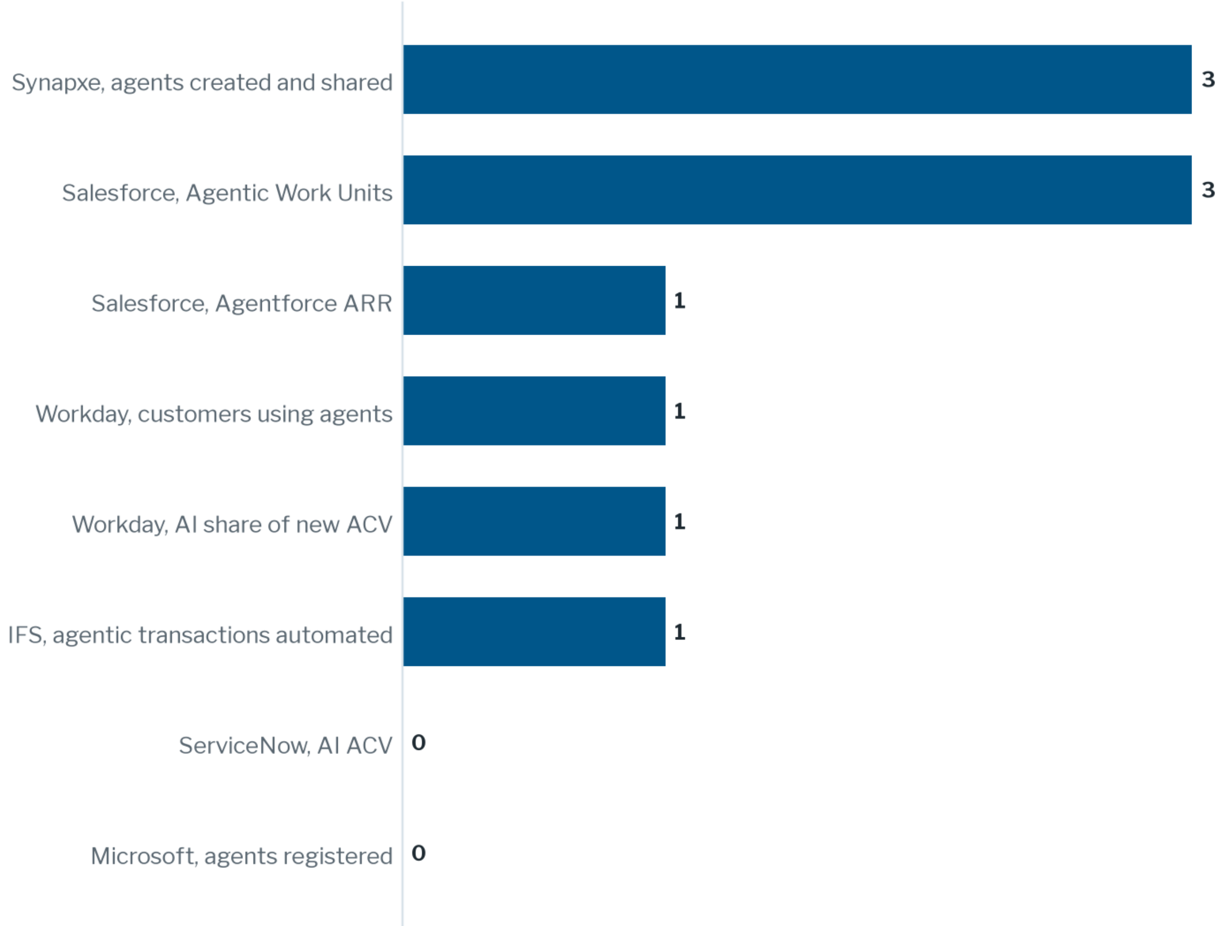
ASSESSMENT There is a quieter version of the same problem in Salesforce's own telemetry study. The Agentic Enterprise Index reports average activated agents per business rising from 5 in February 2025 to 13 in April 2026, and states that to qualify for inclusion a business had to have activated agents in production in every month of the period. Every organisation that tried agents and stopped is excluded from the denominator by construction. The growth is real for the cohort that survived and says nothing about the cohort that did not. [16]

HYPOTHESIS Proposition: agent counts across the category are net of nothing, because no vendor in this study publishes a retirement, deactivation or dormancy figure. This would be refuted by any vendor disclosing agents removed in a period, or by any operator publishing an activity threshold in its definition. It would be confirmed by a vendor reporting a rising agent count in a quarter it also announces a consolidation. Workday's Q2 is close to that second case but not decisive, since its metric counts customers rather than agents. [1] [2]

Disclosure quality, scored

The grading rubric has four criteria, each worth one point. Is a definition published? Is a denominator or population published alongside the numerator? Did the composition stay stable across periods? Can an outsider check the figure against another disclosure? Okta and nCino are excluded because there is nothing to grade, and their absence from the chart should not be read as a failing grade.

FIGURE 7 **Disclosure quality of agentic metrics, scored out of four**



ASSESSMENT The author's scoring against the four-point rubric stated above: definition published, denominator published, composition stable, independently checkable. Subjective at the boundaries, not in the direction. Six of eight metrics score one or zero. [1] [3] [7] [8] [10] [19] [20]

ASSESSMENT The distribution matters more than any single score. Two metrics clear half marks and six do not. A category where three quarters of the published measures fail on definition, denominator or stability is not one where cross-vendor benchmarking is possible, however confidently the comparisons are drawn in board decks and sell-side notes.

ASSESSMENT The pattern in the scoring is that vendors disclose well where the number is small and precisely where they have the least commercial pressure. Synapxe is a government agency with no revenue riding on the figure and it published the reuse rate that makes its own platform look sparse. Salesforce built a rigorous task metric and a loose revenue metric, and the loose one is the one quoted in headlines. That correlation is worth holding on to when reading the next quarter's releases. [3] [8]

Scenarios to the end of 2027

Point forecasts are not appropriate here, because the outcome depends on a small number of discretionary decisions by vendors and regulators rather than on a trend that can be extrapolated.

Three scenarios, with subjective probabilities. The probabilities are the least reliable content on this page. The earliest visible signs are the most useful.

SCENARIO A Quiet retirement 45 percent

As year-over-year growth rates decelerate, vendors stop reporting agent counts and shift the narrative to outcome metrics such as cases resolved or cost per resolution, without ever restating or reconciling the counts they published in 2026.

Consequence. Comparability is lost permanently. Anyone who built a peer benchmark on 2026 agent counts finds the series ends mid-air, and the 2026 figures become uncheckable against anything.

Earliest visible sign. A vendor that reported an agent metric for three consecutive quarters omits it from a release while still discussing agents in the prepared remarks.

SCENARIO B Definitional convergence 30 percent

Pressure from auditors, SEC comment letters on Item 2.02 metrics, or a major analyst firm publishing a standard pushes vendors to footnote agent metrics the way they footnote ARR, with stated denominators and disclosure of composition changes.

Consequence. Cross-vendor comparison becomes possible for the first time, and several 2026 growth rates look materially worse when restated on a consistent basis.

Earliest visible sign. Any agentic operating metric appearing inside a 10-Q rather than only in a furnished 8-K exhibit, or an SEC comment letter naming an agent metric.

SCENARIO C Escalation and a public correction 25 percent

Registry-style counts become the category standard because they produce the largest numbers, until a short-seller report, a restatement, or a customer-side audit publicly contradicts a headline figure and forces a vendor to explain what it was counting.

Consequence. A sharp repricing of AI attach-rate assumptions across the sector, and a rapid, involuntary move to the disclosure discipline of scenario B.

Earliest visible sign. A published bottom-up reconstruction of a vendor's agent count from customer interviews landing materially below the disclosed figure.

Leading indicators

These are observable without private information. Each is a specific event, not a trend, so it either happens or it does not.

If this happens	It means	Moves toward	Where to watch
An agent metric appears in a 10-Q	The vendor has accepted Section 18 liability for the number	Scenario B	EDGAR full-text search for the vendor's own metric

If this happens	It means	Moves toward	Where to watch
rather than only a furnished 8-K exhibit			name, filtered to periodic reports
A vendor publishes agents retired or deactivated in a period	Culls become visible in the count for the first time	Scenario B	Quarterly release operating metrics section
A vendor drops a previously reported agent metric without comment	Growth has decelerated below the level worth reporting	Scenario A	Compare the operating metrics list quarter to quarter, not the narrative
Salesforce reports Agentforce ARR growth without restating the pre-Slackbot base	The composition change is being carried through a full year of comparisons	Scenario A	Q3 and Q4 FY27 releases, footnote to the ARR definition
ServiceNow separates pre-bundle from post-bundle AI ACV	The company is treating the April packaging change as material to the metric	Scenario B	Q3 2026 release and prepared remarks
An operator other than Synapxe publishes agents created alongside agents reused	Reuse becomes a category-standard denominator	Scenario B	Platform operator announcements, especially public sector
An SEC comment letter names an agentic operating metric	The regulator has taken an interest in Item 2.02 metric definitions	Scenario B	EDGAR correspondence filings for large software issuers
A bottom-up count from customer interviews contradicts a headline figure	The disclosure gap has become a public dispute	Scenario C	Sell-side channel checks and short-seller publications

Implications

For CIOs benchmarking peer agent adoption

ASSESSMENT Stop benchmarking against vendor agent counts. There is no comparable series in this study, and a board slide showing your agent count against a published industry figure is comparing your production deployments to somebody's registry. Build the internal metric the vendors will not give you: agents in production, agents used by more than one person, and agents deactivated in the period. The third of those is the one nobody has and the one that tells you whether your programme is learning. [7] [8]

ASSESSMENT If a peer comparison is unavoidable, the AgentSea ratio is the most defensible external anchor available, because it is the only published figure that pairs creation with reuse. Expect the great majority of what your organisation builds never to be used by a second person, and treat that as normal rather than as a failure of the programme. Budget for the cull rather than being surprised by it. [8]

For equity analysts pricing AI attach rates

ASSESSMENT Three specific adjustments. First, discount ServiceNow's AI ACV growth for the April 2026 bundling change, because post-April AI ACV includes revenue from customers who did not choose to buy AI. Second, treat Salesforce's Agentforce ARR growth rate as containing an unquantified contribution from Slackbot and Headless 360 from Q2 FY27, and refuse to model the 240 percent as organic. Third, note that none of these metrics sits in a filed periodic report, so the usual comfort that a number has been through the disclosure controls of a 10-Q does not apply. [3] [6] [19] [20]

ASSESSMENT The metric worth the most modelling effort is Salesforce's AWU series, because it reconciles and it measures execution rather than contract value. The work to do is finding the denominator the company withholds, which means building a tasks-per-customer estimate from disclosed customer counts and testing whether unit growth is coming from more customers, more tasks each, or cheaper tasks being routed through agents. Those three have very different margin implications. [3] [5]

For procurement leads negotiating agentic renewals

ASSESSMENT The single highest-value question in a ServiceNow renewal is what the AI component of the bundle would cost if it were priced separately, given that from 9 April 2026 it is not. A vendor quoting AI ACV growth as evidence of value is quoting a number that rose partly because the vendor moved the product into the base tier. Ask for consumption data instead: assists used against assists allocated, per fulfiller, over the last four quarters. That is a demand signal and the tagged contract value is not. [19]

ASSESSMENT Across every vendor here, ask for a definition in writing before accepting any agent number in a business case, and ask specifically whether the metric can fall. If a vendor cannot say what event would make its published figure decrease, the figure is a cumulative marketing total and belongs nowhere near a renewal justification. Two vendors in this study, Okta and nCino, decline to publish agent metrics at all, which is a reasonable position and worth naming as such when a competitor's count is put in front of you. [12] [13]

For product leaders setting adoption targets

ASSESSMENT Do not set a target on agents created. It is the easiest number in this category to move and the least connected to anything. AgentSea produced 12,000 agents in three months from a workforce of 80,000 and 300 of them were worth sharing; a target on the first number would have been hit spectacularly and meant nothing. Set the target on the reuse rate, or on agents surviving 90 days with more than one user, and accept that the number will look small. [8]

ASSESSMENT Design the retirement path before the launch. Workday's experience and the SaaStr account both describe a management ceiling reached well before a technical one, with consolidation into fewer, larger agents improving results rather than degrading them. A catalogue that only grows will hit that ceiling, and the vendor metrics offer no warning because none of them can decline. [2] [17]

For enterprise architects scoping agent pilots

ASSESSMENT Scope the pilot to the autonomy level the organisation will actually grant, which the survey evidence puts at supervised or routine work for roughly four organisations in five. A pilot designed around full autonomy is testing a configuration that around one enterprise in twenty permits, and its results will not transfer to the deployment that follows. The 60/40 split between end-to-end automation and human review that IFS discloses is a more realistic design target than anything implied by the headline adoption numbers. [9] [10]

ASSESSMENT Instrument for the metrics the vendors do not publish, from day one: per-agent invocation counts, distinct users per agent, and time since last use. Retrofitting that telemetry after a catalogue has grown to a few hundred agents is the point at which most organisations discover they cannot tell which ones matter. The registry the platform gives you counts existence, and existence is the one thing you do not need to measure. [7]

What this document cannot answer

Four questions need primary research that public sources cannot supply.

What proportion of created agents are used at all. The AgentSea sharing figure is a proxy for quality, not for use, and an agent used daily by one person is invisible in it. Answering this needs platform telemetry that no operator publishes: invocations per agent, distinct users per agent, and days since last invocation. A single operator releasing that distribution would settle the central question in this report.

What a unit of agent work is worth. Salesforce's AWU definition explicitly spans resolving a customer case and updating a record, and the commercial gap between those is enormous. Without a value distribution across task types, a billion units cannot be converted into a cost saving, and every ROI claim built on volume is unfalsifiable.

How much of ServiceNow's AI ACV growth after April 2026 is bundle allocation rather than incremental purchase. This is knowable inside the company and inferable from a large enough sample of customer contracts, but not from published materials. It is the difference between a genuine attach rate and an accounting choice.

Whether agent counts are net of retirements anywhere. Every vendor in this study was checked for a deactivation, dormancy or retirement figure and none publishes one. Establishing whether the underlying systems even track it would require vendor interviews rather than document review.

Half-life of this assessment

Decays within six months: every specific figure quoted from a quarterly release, including Salesforce's ARR and AWU totals, Workday's 5,500 customers, and ServiceNow's $1 billion AI ACV. These are superseded every ninety days by construction. The AgentSea creation and sharing counts will also have moved, though the ratio may prove more durable than the levels.

Decays within twelve months: the disclosure-quality scores, which are graded against what vendors published in 2026 and will change the moment any of them adds a footnote. The scenario probabilities, which are calibrated to a market that has not yet had its first public dispute over an agent count. Gartner's cancellation forecast, which resolves at the end of 2027 and can be scored then.

Decays within twenty-four months: the trust figures. Both the HBR and Futurum numbers describe a moment when full autonomy was rare, and the direction of travel in the underlying technology is toward more of it. If those figures have not moved by late 2028, that is itself the more interesting finding.

Architectural, and unlikely to decay: the distinction between counting registrations, activations, executions and tagged contract value. Those four are different objects and will remain different objects regardless of which vendor reports which. The observation that a metric with a threshold of one cannot decline is arithmetic. And the structural point that operating metrics furnished under Item 2.02 carry an explicit disclaimer of Section 18 liability holds until the securities law changes.

Limitations

This analysis was produced independently. No vendor named in it commissioned, reviewed, funded or saw it before publication, and the author holds no position in any company discussed and no commercial relationship with any of them.

ASSESSMENT The evidence base has specific weaknesses that should discount specific conclusions. Two of the sources supplied at the outset could not be retrieved directly. The Straits Times report on AgentSea was substituted with the primary Synapxe announcement carrying the same figures, which is a strict improvement. The Procurement Magazine article returned a server error on repeated attempts, so its statistics reached the author only as search-index fragments; it is listed as second-hand and unverified and no claim in this report rests on it. [22]

ServiceNow's primary earnings release could not be retrieved directly, so the AI ACV figures come from press coverage rather than from the company's own document. The packaging change underlying the ServiceNow finding is sourced to ServiceNow's own community documentation, which is published by the vendor but is not corporate financial disclosure. Both weaken the ServiceNow row of the comparison table relative to the others, and the finding should be read as well supported on the packaging change and less well supported on the exact ACV figures.

The two trust surveys are both vendor-sponsored and one does not disclose its sample size. They are used to establish an order of magnitude, not a precise share, and no conclusion here would change if the true figure were four percent or nine. The LangChain practitioner survey is self-selected, technology-heavy and skewed to small organisations, and is cited only as directional corroboration.

The disclosure-quality scores are the author's judgment applied to a rubric the author wrote. The rubric is stated so the scoring can be disputed on its own terms, and the boundary between a score of one and a score of zero is genuinely arguable in at least two cases. The direction of the distribution is not.

Absence of evidence is doing real work in this document. Several findings rest on a definition not being found rather than on a definition being shown to be absent. Full-text search of the EDGAR index is a strong check for filed documents and no check at all for investor day materials, webcast remarks, or supplemental decks posted to an investor relations site. Where that distinction matters, it is stated in the text.

Sources

[1] Workday, Inc. Form 8-K, Exhibit 99.1, "Workday Announces Fiscal 2027 Second Quarter Financial Results." Filed 27 August 2026. Accession 0001327811-26-000042. *Vendor primary* sec.gov

[2] Lauchlan, Stuart. "Following agentic cull, rising adoption is delivering ROI for Workday customers, says CEO Aneel Bhusri." diginomica, 28 August 2026. *Independent press* diginomica.com

[3] Salesforce, Inc. "Salesforce Delivers Record Second Quarter Fiscal 2027 Results." Form 8-K Exhibit 99.1, 26 August 2026. Accession 0001108524-26-000187. *Vendor primary* investor.salesforce.com

[4] Salesforce, Inc. Form 8-K cover, Item 2.02 Results of Operations and Financial Condition, 26 August 2026. Contains the Section 18 furnished-not-filed language quoted in the text. *Vendor primary* sec.gov

[5] Salesforce, Inc. "Salesforce Delivers Record First Quarter Fiscal 2027 Results." 27 May 2026. Contains the fuller Agentic Work Unit definition. *Vendor primary* investor.salesforce.com

[6] US Securities and Exchange Commission, EDGAR full-text search index. Queries run 30 August 2026 for "Agentic Work Unit" (3 results, all Form 8-K exhibits), "Agentforce and Data 360 annual recurring revenue" (6 results, 8-K exhibits and proxy materials), and "organic agents" (the Workday result is a Form 8-K exhibit). *Regulator primary* sec.gov

[7] Microsoft Corporation. Fiscal year 2026 fourth quarter earnings conference call transcript, 29 July 2026. Remarks by Satya Nadella. *Independent press* fool.com

[8] Synapxe. "New platform to create AI agents available to all public healthcare professionals." 24 August 2026. *Vendor primary* synapxe.sg

[9] Harvard Business Review Analytic Services, sponsored by Workato and Amazon Web Services. Survey of 603 business and technology leaders, fielded July 2025, published 9 December 2025. *Analyst firm (vendor-sponsored)* finance.yahoo.com

[10] IFS. "Industrial workforce capacity gap being filled by agentic digital workers." PR Newswire, 26 August 2026. Cites research by The Futurum Group commissioned by IFS, "Bridging the Capacity Gap." *Vendor primary* prnewswire.com

[11] Automation Magazine. "Industrial workforce capacity gap being filled by agentic digital workers in manufacturing." 28 August 2026. *Independent press* automationmagazine.co.uk

[12] The Futurum Group. "nCino's Q2 FY2027: Agentic AI Banking Thesis Meets Margin Reality." 26 August 2026. *Analyst firm* futurumgroup.com

[13] MT Newswires. "Okta's early traction in agentic AI." Yahoo Finance, 27 August 2026. *Independent press* finance.yahoo.com

[14] Gartner. "Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027." Press release, 25 June 2025. Also the source of the agent washing characterisation and the estimate of roughly 130 genuine agentic vendors. *Analyst firm* gartner.com

[15] Gartner. "Gartner Forecasts Supply Chain Management Software with Agentic AI Will Grow to $53 Billion in Spend by 2030." Press release, 7 April 2026. Cited as context on category spend, not relied on for any finding. *Analyst firm* gartner.com

[16] Salesforce. "Agentic Enterprise Index, Second Edition." Telemetry from a cohort of businesses with activated agents in production in every month from February 2025 to April 2026, plus 4,689 survey responses collected May 2026. *Vendor primary* salesforce.com

[17] Lemkin, Jason. "We Peaked at 30 AI Agents. Now We're Coming Back Down to 20." SaaStr, July 2026. *Practitioner aggregate (self-selected)* saastr.com

[18] LangChain. "State of Agent Engineering." 1,340 self-selected responses collected 18 November to 2 December 2025; 63 percent technology sector, 49 percent organisations under 100 employees. *Practitioner aggregate (self-selected)* langchain.com

[19] ServiceNow Community. "ServiceNow AI-Native Licensing in 2026: A Practical Guide to Assists, Consumption, and Governance." 29 June 2026. Vendor-hosted community documentation, not corporate financial disclosure. *Vendor primary (community)* servicenow.com

[20] TechTimes. "ServiceNow AI Revenue Crosses $1B as Non-Seat Deals Hit 50% of New Business." 27 August 2026. Used because ServiceNow's primary release could not be retrieved directly. *Independent press* techtimes.com

[21] Workday, Inc. Form 10-Q for the quarterly period ended 31 July 2026, filed 27 August 2026. Cited for the installed base of more than 11,000 organizations and for the absence of any agent operating metric in the periodic report. *Vendor primary* sec.gov

[22] Procurement Magazine. "Agentic AI in Procurement: Gartner, Bain, Unilever & Revolut." The page returned HTTP 403 on repeated direct retrieval; its figures reached the author only through search-index snippets and are relied on nowhere in this analysis. *Second-hand (unverified)* procurementmag.com

Rights and permitted use

Commissioned Market Assessment reports, commercial licensing, and briefings on this material are available.

This document is analysis, not investment, legal, or procurement advice. It was not commissioned, reviewed, or funded by any company named in it, and the author holds no position in any of them.

Market Assessment reports, commissioned research, and briefings: **davidsoden.com/market-assessment**