

Institutional Research Group



Rudy Torrijos
Director, Industry Research
rudy.torrijos@pitchbook.com



Derek Hernandez
Senior Research Analyst,
Enterprise SaaS and
Infrastructure SaaS
derek.hernandez@pitchbook.com

pbinstitutionalresearch@pitchbook.com

Published on March 9, 2026

Contents

Executive summary	1
Key takeaways	2
Market context and drivers	4
Convergence	4
The need for AgentOps	8
Enterprise adoption constraints	10
Regional adoption variability	11
Cost of intelligence	12
SaaS-ability	14
Transition risk	15
Self-cannibalization	15
Product margins	15
Customer revenue potential	16
Framework taxonomy	17
Methodology limitations & data confidence	17
Framework 1: Business model transition score	18
Framework 2: Data moat & sovereignty score	19
Framework 3: Incumbent resilience score	20
Framework 4: Technical autonomy score	21
Framework 5: Enterprise control plane score	22
Framework 6: Cognitive architecture score	23
Framework 7: AgentOps & workforce maturity score	24
Framework 8: Agent fidelity matrix (product agency score)	25
The new vocabulary for enterprise AI	29
Glossary	29

EMERGING TECH RESEARCH

Through the Looking Glass: The Race to Build Enterprise AI

Unlocking the labor market with SaaS

PitchBook is a Morningstar company providing the most comprehensive, most accurate, and hard-to-find data for professionals doing business in the private markets.

This report is the second installment of the series *SaaS is Dead, Long Live SaaS*.
Read Part I [here](#).

Executive summary

The enterprise software industry is undergoing a profound structural shift, more transformational than the move to the cloud 20 years ago. The seat-based subscription model on which software as a service (SaaS) has relied is dead. Service as software (SaS), defined by digital labor outcomes, is its replacement. This transition is both technological and economic: It represents the merger of the over \$900 billion global enterprise software market with the over \$50 trillion global knowledge worker labor market.^{1,2}

The race to build enterprise AI is defined by two parallel tracks: the SaaS vendor racing to capture the payroll budget via outcome-based pricing and the enterprise agent operations (AgentOps) team racing to capture the intelligence via proprietary curation. Success will be measured by expanding revenue and margins on one side, and by compounding value creation from the digital workforce on the other.

This report provides a rigorous framework for identifying incumbents and startups capable of migrating payroll budgets into software spend while maintaining strong long-term viability. We define eight specific scoring frameworks stakeholders can use to identify winning business models and durable moats across the SaaS and SaS landscapes today.

We also include market drivers, context, and a comprehensive glossary of essential enterprise AI terms. Our core thesis is that generative AI (GenAI) has commoditized code production, eroding many of the defensive moats of the SaaS era. Existing

1: "Market Share Alert: Enterprise Software Market Size Reached \$900B in 2024," Gartner, Mudit Sharma, et al., June 4, 2025.

2: According to PitchBook estimates.



software vendors must transition from selling access to tools that require human operators to selling outcomes that replace entire employee workflows.

The primary barrier to this transition is what we call the “autonomy gap,” the operational cost incurred when an autonomous agent fails and requires human intervention. Until this gap is closed, the economics of labor replacement remain unattainable. Therefore, success in this new era depends on three simultaneous requirements: (1) Vendors must close the autonomy gap and (2) adopt outcome-based pricing to capture the value of labor, while (3) enterprises must adopt AgentOps to govern the risk of autonomy.

The pivot from SaaS to SaS introduces a period of financial risk, especially for incumbents. As enterprises deploy digital workers, seat counts may reduce seat-based revenue before outcome-based revenue can scale to offset the loss. This is the SaS J-Curve. To navigate this transition carefully, vendors must manage three distinct financial levers that determine their long-term viability: self-cannibalization, the cost of providing intelligence, and the true revenue potential of their customers.

We believe we are at least 18 months ahead of when agentic systems will be ready to replace basic employee workloads given software development cycles, enterprise operational readiness, minimum large language model (LLM) expert intelligence level requirements, economically viable cost per token pricing, and the time-to-power requirements for next-generation AI datacenter computing capacity.

Key takeaways

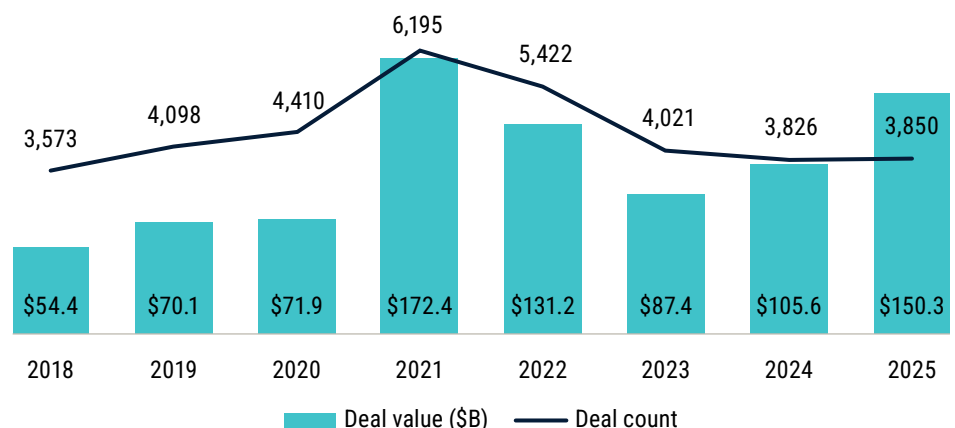
- **The autonomy gap is the gating factor:** The defining challenge for enterprise IT is managing the cost of human intervention required to keep agents running. Success is measured by SaS’s ability to move quickly from requiring human validation for many AI decisions to only requiring human oversight in rare exceptions.
- **AgentOps is the new enterprise IT mandate:** Deploying autonomous software requires a fundamental retooling of the workforce. Enterprises must establish platform-centric infrastructure and new roles such as “reliability architects” and “risk governors” to manage the life cycle, behavior, and skill atrophy of the digital workforce.
- **Intelligence is elastic:** Unlike legacy software, agentic performance scales linearly with compute cost. The physics of autonomy dictates that higher-fidelity reasoning consumes exponentially more compute. Intelligence is no longer a fixed salary-based asset; it will be a variable cost. Vendors must balance model expert intelligence with inference cost-efficiency to maintain margins and prevent enterprise pricing from exploding.
- **The market has bifurcated:** This taxonomy distinguishes between incumbents leveraging data moats and distribution to bundle AI and architecturally adaptive AI startups building compounding, workflow-embedded advantages. Only those who successfully capture the future SoR and proprietary data feedback loops will survive the eventual commoditization of intelligence.



- **The J-Curve is the primary financial risk:** Investors must prepare for near-term volatility as seat-based revenue necessarily declines. Winners will be those who successfully implement hybrid metering to increase average revenue per account as they transition fully to outcome-based pricing.
- **Net labor capture is the new key performance indicator (KPI):** Traditional benchmarks like net dollar retention are insufficient for the SaaS era, as bookings may continue to grow despite an indecisive transition to outcome-based revenue. Success will be defined by a vendor's ability to migrate a customer's payroll budget into their software spend, effectively turning software from a line-item expense into a labor-replacement asset.
- **Human-in-the-loop is a permanent requirement:** Closing the autonomy gap is not a purely technical challenge. Enterprises must formalize new high-value roles, domain experts, real-time interventionists, and governance auditors, to provide the proprietary ground truth and ethical oversight required for systemic reliability.
- **Value pivots toward human judgment:** As generative models drive the marginal cost of technical execution toward zero, economic premium shifts to human oversight. Organizations must integrate technical architecture with high-judgment governance, using human intervention as the essential risk-management layer for scaling enterprise AI.
- **Maintaining margins enables survival:** Unlike legacy SaaS, digital workers incur a continuous "cost of intelligence." Vendors must pass these inference costs to customers via outcome pricing or risk degrading into low-margin services businesses, until such time that inference costs decline substantially.
- **Incumbent moats are conditional on speed:** Distribution and data ownership provide the foundation for incumbent account control, but scale without rapid AI bundling is a value trap. Investors must prioritize vendors capable of closing the autonomy gap before their competitors, both other incumbents and startups, erode their customer base.

Our PitchBook data cited herein is reflected in our combined SaaS coverage, including the Enterprise SaaS and Infrastructure SaaS segments. These can be found on the PitchBook Platform under each respective analyst-curated vertical.

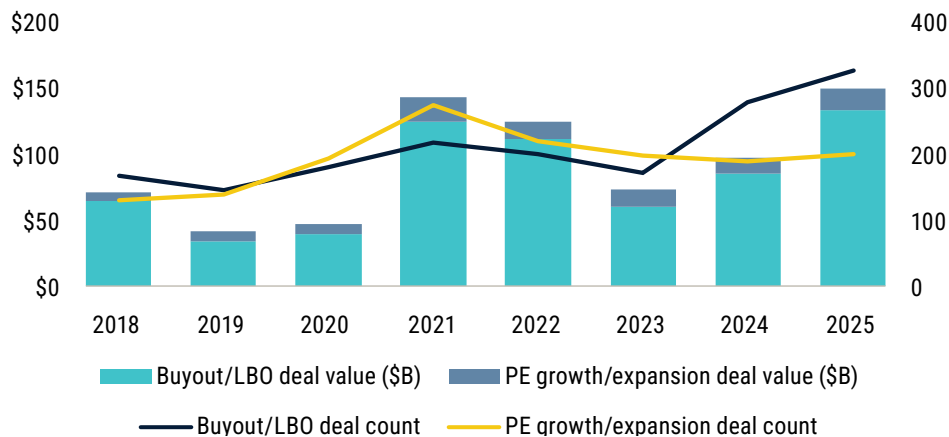
SaaS VC deal activity



Source: PitchBook • Geography: Global • As of December 31, 2025



SaaS PE deal activity by type



Source: PitchBook • Geography: Global • As of December 31, 2025

Market context and drivers

Convergence

The race to build enterprise AI is defined by two parallel tracks: the SaaS vendor racing to capture the payroll budget via outcome-based pricing, and the enterprise AgentOps team racing to capture the intelligence through proprietary curation. Success will be measured by expanding revenue and margins on one side, and compounding value creation of the digital workforce on the other side.

As AI expert intelligence continues to ascend, the shift to autonomous software agents accelerates. To both defend and increase revenues, software vendors must transition from selling tools as subscriptions to selling digital labor as outcomes, as many AI-native challengers have already adopted. Simultaneously, enterprises must retool their operations stacks to create, deploy, and manage these autonomous software entities. This is the rise of enterprise AI. With this report, we have developed a taxonomy that equips stakeholders to assess which vendors are structurally capable of surviving and leading through this convergence.



SaaS VC deal summary

	Quarterly activity					TTM activity	
	Q4 2024	Q1 2025	Q2 2025	Q3 2025	Q4 2025	Q1 2024-Q4 2024	Q1 2025-Q4 2025
Deal count	910	991	988	971	900	3,826	3,850
QoQ change	-0.1%	8.9%	-0.3%	-1.7%	-7.3%	N/A	0.6%
Share of total VC	8.8%	8.7%	10.2%	9.9%	9.3%	8.8%	9.5%
Deal value (\$B)	\$44.1	\$64.9	\$26.4	\$32.2	\$26.8	\$105.6	\$150.3
QoQ change	100.3%	47.3%	-59.4%	22.1%	-16.7%	N/A	42.4%
Share of total VC	34.4%	47.2%	23.0%	25.7%	20.9%	26.9%	29.7%
Exit count	136	132	138	135	137	460	542
Acquisitions	100	96	94	85	88	329	363
Buyouts	29	28	39	38	38	102	143
Public listings	7	8	5	12	11	29	36

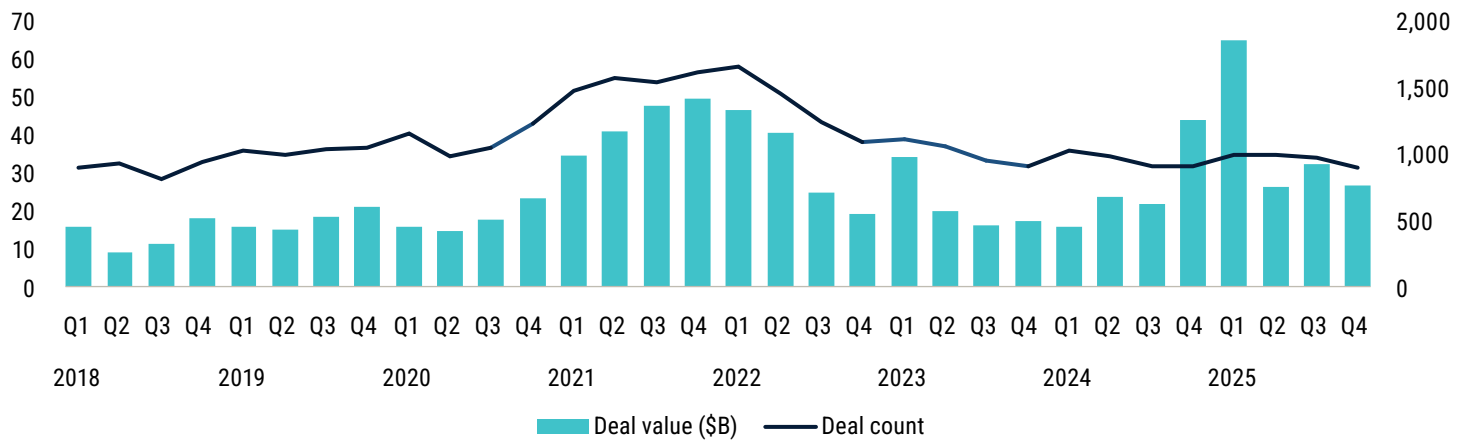
Source: PitchBook • Geography: Global • As of December 31, 2025

Vendor incentives

The shift to SaaS represents a fundamental business model revolution. Success ultimately requires a vendor's revenue model to align with the customer's new operational needs. The SaaS industry is transitioning from the seat-based license model, which caps revenue based on headcount, toward an outcome-based pricing model that monetizes completed work, enabling access to expense lines traditionally served by employee salaries. By merging software budgets with payroll budgets, successful vendors can realize dramatic revenue upside. Our taxonomy evaluates a vendor's SaaS-ability, specifically, their structural capacity to pivot from selling access to selling results. This thesis began with our note [SaaS is Dead, Long Live SaaS](#), which provides further detail.



SaaS VC deal activity by quarter



Enterprise requirements

The greatest barrier to agent adoption is the autonomy gap. This is defined as the operational cost incurred when an agent requires human intervention. This metric encompasses the costs of manual fixes, retraining, and error recovery. This gap is currently far too wide in many sectors for enterprises to confidently deploy agents across their workflows, creating a significant barrier to broader adoption. We expect this gap to narrow significantly in the near term and believe enterprises must prepare for this new reality immediately.

As this gap closes through further AI and agentic advancements, IT organizations must evolve their workforce from operators to supervisors and deploy platform-centric infrastructure to manage agent life cycles. IT will be able to transition its role from a cost and reliability center to a strategic revenue partner. By automating fixes and implementing technology, IT can focus on supporting strategic planning, finance, and business analytics. There will also be an opportunity for a bifurcation of traditional enterprise IT to these broader revenue partner support services and traditional digital infrastructure support services, although in the near term they will remain under one roof as this agentic transformation develops further.

The immediate priority of IT operations should be to implement a platform-centric infrastructure that enables central control across on-premises, edge, and public cloud environments. A unified platform will provide the necessary structure to manage and enable agents to function most effectively. To start, organizations must deploy software-enforced human-in-the-loop protocols for all high-risk operational tasks. In this model, agents propose actions, but human engineers must validate them prior to execution to prevent catastrophic errors or data leaks. As system guardrails are developed and trust matures, the model evolves to a human-aligned model, where operators supervise fully autonomous agents and only intervene during exceptions.

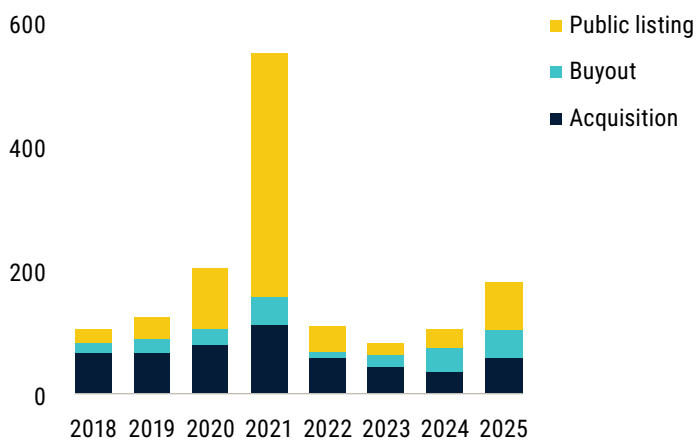


Establishing these robust guardrails is essential. This is the primary responsibility of the reliability architecture team. This team will be responsible for high-level design, governance frameworks, and systemic resiliency of the entire enterprise stack. For any agent, input guardrails must be deployed to filter out unsafe prompts, while output guardrails verify the safety of generated code or commands before they are executed within IT infrastructure. Without these initial safety nets, the risk of operational failure significantly outweighs the potential efficiency gains.

Gartner predicts that human involvement in operational workflows will drop from 95% today to under 40% by 2028,³ driven by this shift. We make the case in our recent report [SaaS is Dead, Long Live SaaS](#) that the “SaaS-pocalypse” market event has accelerated this transition, especially within IT operations, as vendors are reevaluated for their ability to recreate and replace entire employee workflows with AI.

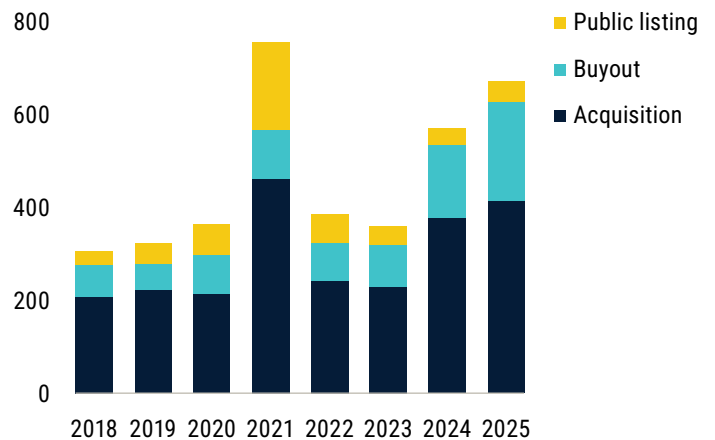
We believe we are at least 18 months ahead of when agentic systems will be ready to replace basic employee workloads given development time, operational readiness, and the time-to-power requirements of AI datacenters for cloud computing.

SaaS PE and VC exit value (\$B) by type



Source: PitchBook • Geography: Global • As of December 31, 2025

SaaS PE and VC exit count by type



Source: PitchBook • Geography: Global • As of December 31, 2025

3: "Predicts 2026: AI Agents Will Transform IT Infrastructure and Operations," Gartner, Cameron Haight, et al., December 4, 2025.



The need for AgentOps

The deployment of AgentOps is essential for building the last mile of AI autonomy. We see AgentOps as the most important investment area today within the enterprise. This last mile is designed to manage the life cycle and behavior of autonomous AI agents while ensuring enterprise resilience and compliance. AgentOps will also ensure that platform-centric infrastructure is deployed throughout the extended enterprise and maintain interoperability across hybrid clouds.

AgentOps also requires new skills, shifting enterprise IT operators toward roles such as autonomy curators, policy engineers, and risk governors. They provide the control layer, ensuring that autonomous actions remain governed, auditable, and aligned with measurable business outcomes. The primary function of the IT manager will evolve from execution to intervention, requiring the enterprise to formalize three distinct human-in-the-loop workflows:

- **The enterprise must employ autonomy curators:** subject matter experts who finetune agents on complex workflows. Vendors support this effort by providing learning interfaces that capture and encode domain expertise for their solutions. This collaboration converts the customer's proprietary business logic into a defensive moat against competitive substitution by a challenger vendor.
- **The enterprise will need policy engineers to intervene when autonomous agents encounter novel scenarios or fall below confidence thresholds.** Vendors must support this via seamless handover protocols, enabling operators with advanced issue-resolution skills to assume immediate manual control and guide the system through edge-case issues.
- **The enterprise will require risk governors to act as compliance officers for the digital workforce.** Vendors enable this oversight by generating forensic reasoning logs and flight recorder-style data. This transparency allows auditors to enforce strict data sovereignty and feed failure data back to curators (policy engineers), transforming unpredictable outputs into auditable, de-risked assets that improve over time.



As generative models drive the marginal cost of coding, data entry, and error resolution toward zero, the economic premium shifts toward the uniquely human traits of nuanced judgment, customer empathy, and ethical accountability. Enterprises will require new standards where human judgment serves as the essential control layer for scaling AI agents.

AgentOps allows the enterprise to optimize cost per task while freeing high-value human capital for future strategy and revenue growth. Our taxonomy assesses whether a vendor provides the necessary agent fidelity to allow an enterprise to move from validation protocols to supervision models without risking operational stability. The enterprise IT workforce must adapt to manage the gap by rigorously evaluating agent reasoning and adaptability. Success will be defined by building a platform-centric foundation that allows humans to govern these agents effectively at scale.

SaaS heatmap by segment

Segment	TTM total VC deal value (\$M)	TTM total VC deal count	YoY VC deal value growth	YoY VC deal count growth
Analytic platforms	\$11,405.8	528	68.3%	14.3%
Application infrastructure	\$4,135.0	163	-0.2%	9.4%
Customer relationship management	\$14,906.5	699	5.7%	-4.1%
Data software & systems	\$4,122.6	102	-69.6%	-21.5%
Devops	\$6,145.2	134	53.9%	-14.1%
Enterprise resource planning	\$31,500.1	1,174	22.0%	-4.0%
It ops	\$2,590.9	65	50.1%	0.0%
Knowledge management systems	\$8,944.7	355	78.8%	0.3%
Other application software	\$59,975.7	450	117.8%	19.7%
Supply chain management	\$6,896.3	195	115.5%	3.2%

Source: PitchBook • Geography: Global • As of December 31, 2025



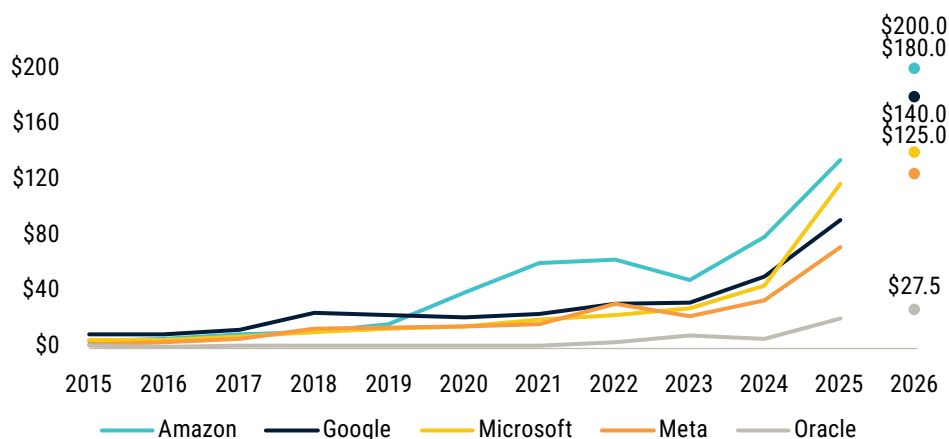
Enterprise adoption constraints

Enterprise adoption will not be gated by model capability alone. The binding constraints are organizational throughput, procurement, governance, and labor-market readiness. This is especially true during the transition from human-managed protocols to a human-aligned operating model.

- **Procurement and deployment timelines.** Even when return on investment (ROI) is clear, agent deployments still route through security reviews, vendor risk management, data residency requirements, and change-control boards. This creates a practical “deployment lag” between infrastructure investment and realized application value, which can materially slow the conversion of pilots into production rollouts. The net effect is that adoption will be lumpy: concentrated in a small number of high-confidence workflows first, with broader expansion occurring only after repeatable governance patterns are proven.
- **CFO cost predictability under outcome pricing.** Outcome-based pricing is directionally aligned with SaS, but it introduces a new budgeting problem: variable unit costs tied to compute and exception handling rather than fixed per-seat subscriptions. The cost of intelligence is elastic and can rise rapidly as tasks require higher-fidelity reasoning, which creates margin and budget uncertainty if metering is not explicit. Vendors that win in enterprise will de-risk this with transparent metering structures (credits, caps, floors) and strong inference efficiency so customers can forecast spend per work unit with CFO-grade confidence.
- **Regulatory and compliance gates.** In regulated environments, autonomy must be bounded. Enterprises will demand a control plane with identity, policy, audit logs, and kill switches, plus deterministic logic and flight-recorder-grade observability to satisfy auditability and data sovereignty. Without these guardrails, the autonomy gap becomes an unacceptable operational subsidy rather than a productivity lever.
- **The AgentOps talent gap.** The market is early in staffing the new operating roles required to supervise digital workers at scale—reliability architects, autonomy curators, policy engineers, and risk governors. This talent gap will slow adoption even where tooling exists, pushing vendors to productize “last mile” AgentOps capabilities and provide implementation patterns that reduce the burden on scarce internal experts.



The “AI Industrial Revolution”: Hyperscaler and others’ capital expenditures (\$B)



Source: Company Filings & 2025/2026 Guidance/Estimates • Geography: Global • As of February 2026
Note: 2026 figures are estimated.

Regional adoption variability

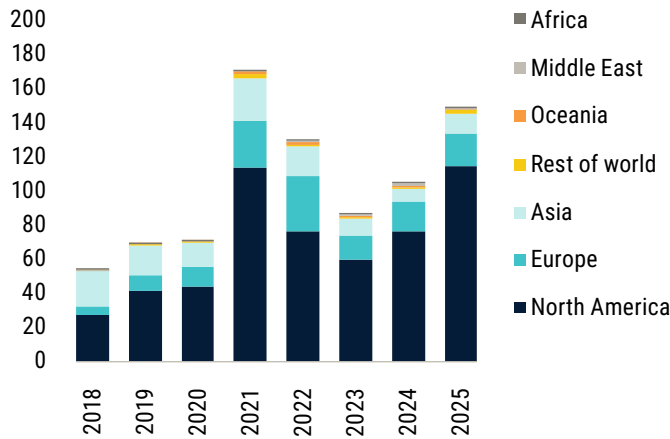
The SaaS transition will not proceed uniformly across geographies. The US is expected to lead adoption, driven by concentrated hyperscaler infrastructure, an established venture ecosystem funding AI-native challengers, and a regulatory environment that currently permits rapid deployment of autonomous agents. Europe faces a structurally different adoption curve. The EU AI Act imposes classification requirements on high-risk AI systems, extending compliance timelines for autonomous agent deployment in regulated verticals. General Data Protection Regulation constraints on training data limit the learning loop mechanisms that underpin the data moat & sovereignty framework. Vendors targeting European enterprises must build sovereignty-first architectures to achieve competitive SaaS-ability scores in that market.

China operates within a separate AI ecosystem with distinct foundation models, data governance regimes, and platform economics that are outside the scope of this taxonomy. Emerging markets across Asia-Pacific, Latin America, and the Middle East will follow with adoption lag, constrained by enterprise IT maturity and local infrastructure availability.

Investors evaluating regional SaaS portfolios should apply the SaaS J-Curve with geography-specific assumptions. The deployment lag between US and EU markets may extend 12 to 24 months for equivalent SaaS-ability penetration. This variance is material for PE investors managing multiregion software holdings, as portfolio companies operating primarily in regulated European verticals will navigate a longer and more capital-intensive transition than US-based equivalents.

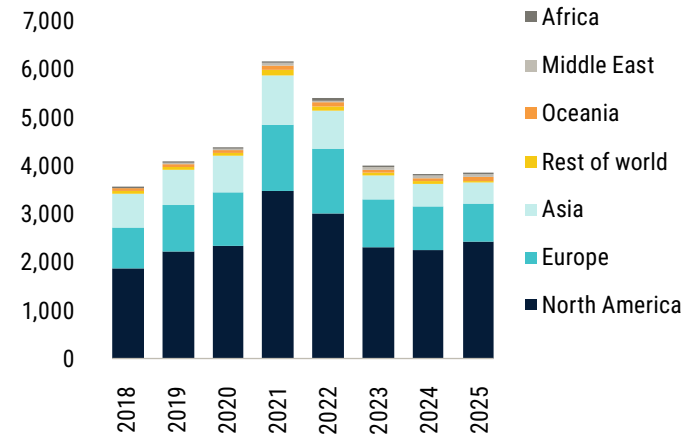


SaaS VC deal value (\$B) by region



Source: PitchBook • Geography: Global • As of December 31, 2025

SaaS VC deal count by region



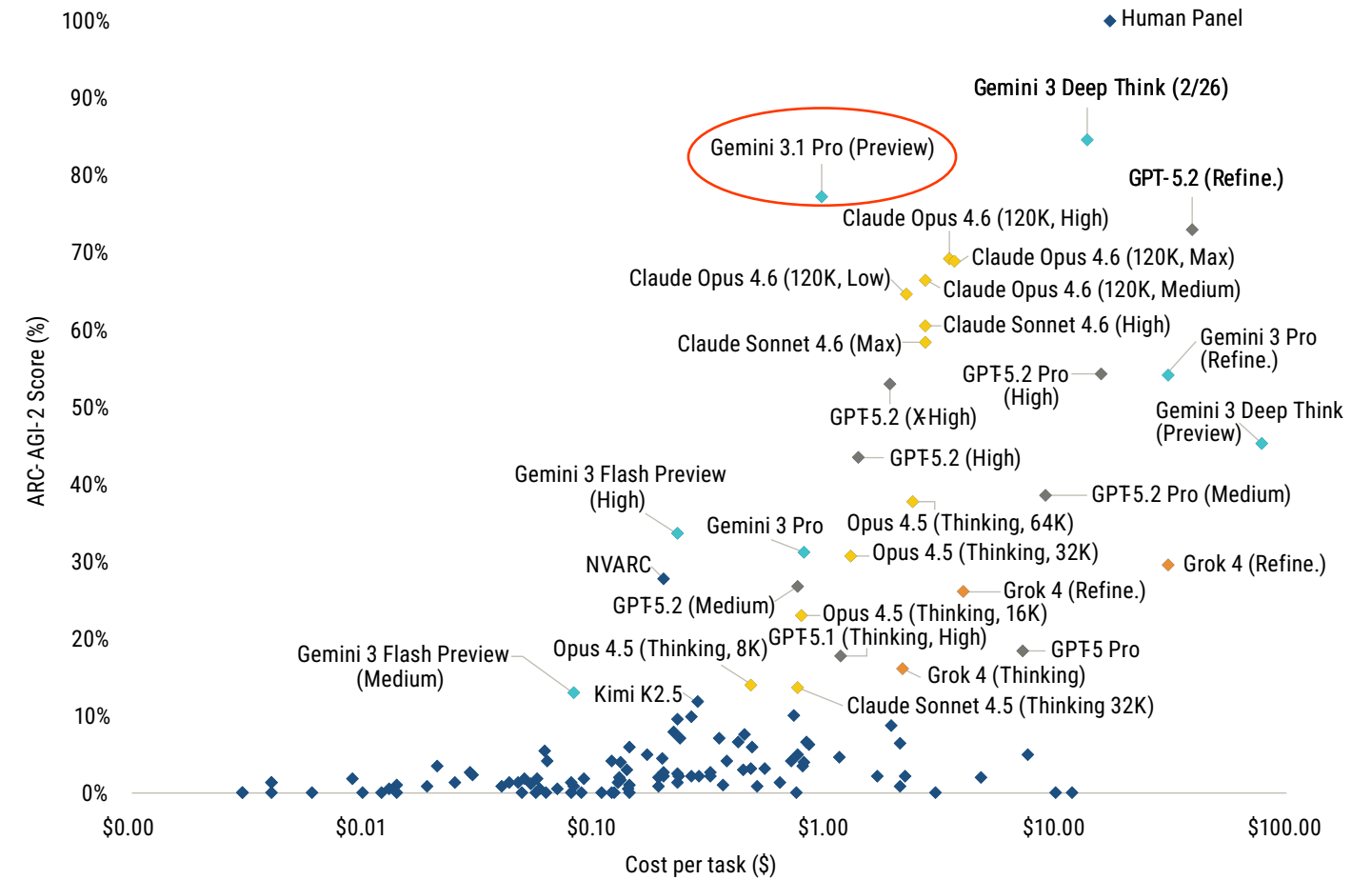
Source: PitchBook • Geography: Global • As of December 31, 2025

Cost of intelligence

Underpinning both vendor incentives and enterprise requirements is a powerful economic constraint: intelligence elasticity. Agentic intelligence scales directly with the amount of computing power allocated to thinking before acting. This means that solving novel, high-value problems consumes exponentially more inference resources. If left unmanaged by the vendor, this variable cost of intelligence will erode gross margins until such time that inference costs decline substantially. Vendors and enterprises must collaborate to optimize this new cost. Our taxonomy identifies the platforms capable of balancing high-fidelity reasoning with the financial pragmatism required for mass deployment. For further details on the dynamics of this cost evolution, please see our recent [Launch Report: Advanced Computing](#).



ARC-AGI-2 leaderboard



Source: [ARC Prize](#) • Geography: Global • As of February 26, 2026



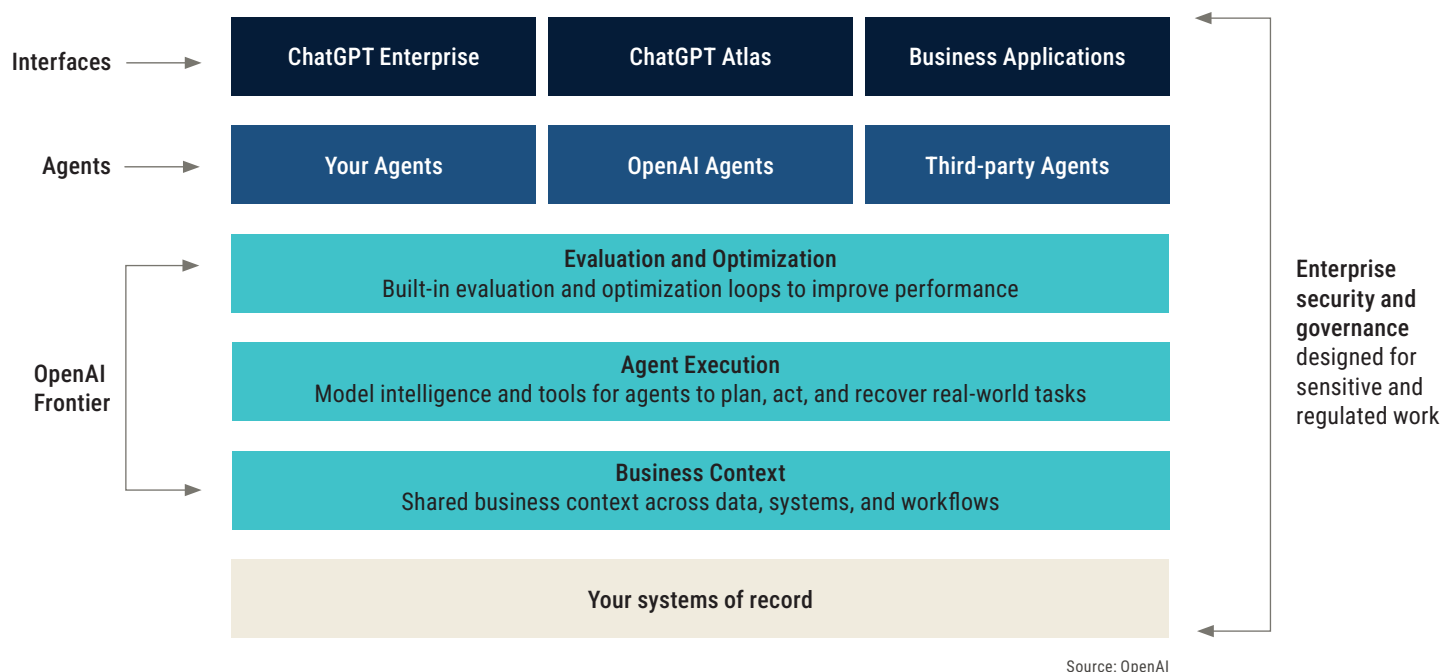
SaS-ability

Ultimately, our taxonomy is designed to identify a vendor's potential to deliver SaS. By successfully addressing the operational autonomy gap for its customers, a vendor will expand its total addressable market (TAM) and capture incremental labor budgets. This taxonomy identifies potential leaders and laggards in the enterprise AI era by driving a structured, repeatable methodology. With AI reasoning tools available today, we evaluate a vendor's potential to meet enterprise requirements. We have defined eight frameworks to accomplish this goal:

1. **Business model transition:** Evaluates the vendor's economic elasticity to pivot from seat-based subscriptions to outcome-based pricing models that capture the value of labor replacement.
2. **Data moat & sovereignty:** Determines if a vendor owns the proprietary system of record (SoR) and data rights necessary to ground agents in enterprise truth and prevent commoditization.
3. **Incumbent resilience:** Measures the structural advantages of distribution, capital depth, and switching costs that allow incumbents to bundle AI features and consume point solutions.
4. **Technical autonomy:** Distinguishes between passive co-pilots and active orchestrators by assessing the agent's ability to safely and securely execute "write" actions within systems of record.
5. **Enterprise control plane:** Audits the maturity of the governance layer, specifically verifying the existence of "Flight Recorder" logs and reversibility tools required for CIO approval.
6. **Cognitive architecture:** Evaluates the operational efficacy of the agent's reasoning loop, measuring its ability to handle ambiguity and maintain contextual continuity across sessions.
7. **AgentOps & workforce maturity:** Assesses the vendor's readiness to support the customer's operational transition to an agent-centric workforce through simulation training and role-based access controls.
8. **Agent fidelity matrix:** Applies the rigorous 10-point product audit to measure the specific cognitive performance and reliability of the software agent.



OpenAI's strategy for competitive control



Transition risk

The pivot from SaaS to SaS introduces a period of financial risk, especially for incumbents. As enterprises deploy digital workers, seat counts may reduce seat-based revenue before outcome-based revenue can scale to offset the loss. This is the SaS J-Curve. To navigate this transition carefully, vendors must manage three distinct financial levers that determine their long-term viability.

Self-cannibalization

The primary risk to an incumbent is efficiency, in which their own AI product reduces demand for the per-user licenses that currently drive revenue. To bridge this gap, successful vendors are implementing hybrid metering. By maintaining a base floor for platform access while layering on success fees for completed work units, vendors will defend and potentially grow their revenue as human headcount declines, so that average revenue per account may rise. This mechanism turns the deflationary pressure of AI into a primary driver of revenue growth for software vendors. While seat-based revenue will almost certainly decline in the short term, the absolute gross margin dollars per account should expand as the vendor captures a portion of the redirected labor cost.

Product margins

Unlike legacy software, where the marginal cost of a new user is near zero, a digital worker incurs a continuous cost of intelligence through inference. Management teams must create systems to route simple tasks to high-efficiency models while reserving

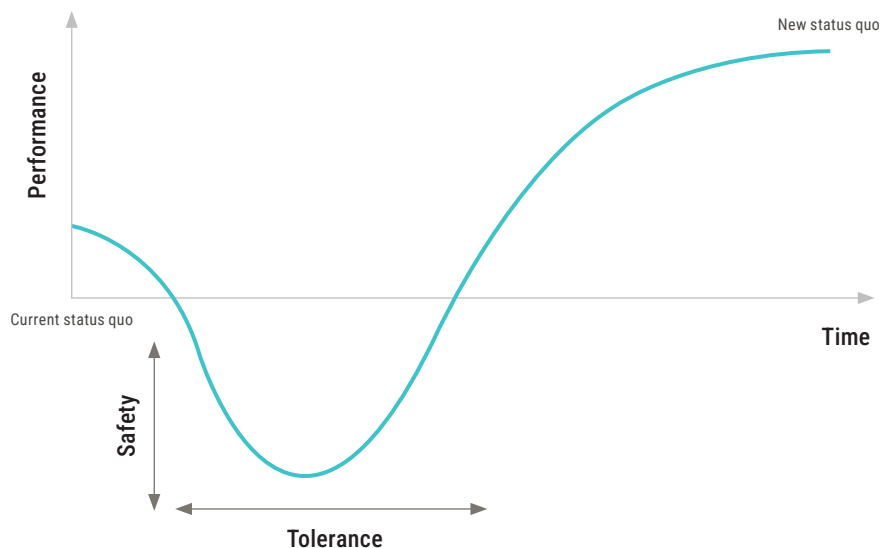


high-cost computing for complex reasoning. If a vendor cannot pass these compute costs through to the customer via consumption pricing, they risk degrading their value proposition to that of a low-margin services business. Maintaining the structural gross margin threshold will be the definitive test of a vendor's technical maturity in this new environment.

Customer revenue potential

Investors must identify the inflection point at which work-unit revenue growth outpaces seat revenue decay. This requires a fundamental realignment of sales incentives and performance metrics. Vendors may deemphasize total contract value based on headcounts and prioritize potential digital labor outcome volume. During this transition, traditional metrics such as net dollar retention will experience significant volatility. We recommend that stakeholders focus instead on net labor capture, which measures the percentage of a customer's total payroll budget the vendor has successfully migrated into their software spend. This type of revenue forecast will require a deeper understanding of the customer's business process needs and will require additional vendor disclosures to investors. This will allow vendors to become more ingrained as strategic partners. Net labor capture metrics will provide the clearest signal of a vendor's success in capturing customer payroll budgets through efficient workflow replacement.

J-Curve



Vendors that help enterprises close the autonomy gap at scale will be the next technology incumbents. Our analysis is that most legacy moats will dissolve into markdown files that AI will execute autonomously. Winners will be those who sell the most reliable digital labor to accomplish this goal.



Framework taxonomy

The transition from seat-based SaaS to outcome-based SaaS is a measurable operational pivot. To assist investors and enterprises in quantifying this shift, we have designed the following eight frameworks to function as a machine-readable scoring engine. With the following analyst prompt and subsequent framework definitions, stakeholders can generate a rigorous “SaaS-ability scorecard.” We note that this prompt is a very basic outline of the considerations required to use our framework. As is the nature of prompt engineering, many iterations will be required to convert qualitative strategic signals into quantitative risk assessment scores to suit individual client workflows. At this stage in LLM development, all output must be rigorously tested and confirmed by human judgment before any subsequent determinations may be trusted.

Methodology limitations & data confidence

We would strenuously note that this SaaS-ability scorecard relies on the depth of a vendor’s digital footprint. We designed these frameworks for companies with established market presence, including publicly traded, PE-owned, and later-stage venture-backed companies. Accurate scoring depends on specific details about pricing models, technical architecture, and feature specifications, which are typically gathered through open-source intelligence.

Stakeholders must recognize that LLM-based scoring is a probabilistic model, not a deterministic audit. The resulting scores should be viewed as a high-signal filtering mechanism rather than investment conclusions. We explicitly recommend that all machine-generated scores be validated against primary research. The nuances of agent fidelity and cognitive architecture often reside in nonpublic engineering documentation or are only visible during live stress tests. Therefore, these scores must be triangulated with customer reference interviews, technical product demonstrations, and hands-on “red teaming” of the agent’s failure modes.

By design, the scoring engine relies on the depth of a vendor’s digital footprint. Private companies operating in stealth or with limited public documentation will structurally underscore relative to public incumbents with extensive investor relations resources. To account for this variance, we recommend that analysts assign a data confidence rating (high, medium, low) alongside each framework score. A low confidence score indicates that the assessment relies on sparse marketing claims rather than verifiable technical documentation, signaling the need for deeper due diligence before underwriting the vendor’s transition to SaaS.

Analyst prompt: “Act as a Senior Enterprise Software Analyst. I will provide you with the ‘8 SaaS-ability Frameworks’ defined in this report. Use these frameworks as your ground-truth scoring methodology. I will then provide you with the name and public details of a software vendor. You will evaluate that vendor against each of the eight frameworks, assigning a score of 1-100 for each category based on their structural readiness for SaaS. Finally, calculate the geometric mean of all of the eight frameworks to create a ‘total SaaS-ability score’ and provide a ‘gap analysis’ summary identifying the specific operational changes required for that vendor to bridge the autonomy gap.”



Framework 1: Business model transition score

Description: This framework evaluates the software vendor's economic structure to determine whether it sells tools or labor. It targets the transition point at which outcome-based pricing enables vendors to capture a substantial revenue uplift by replacing low-cost seats with high-value digital workers. The scoring penalizes legacy seat-based models that disincentivize AI adoption because they are inherently capped by human headcount. Conversely, it rewards outcome-based models that align vendor revenue with the customer's desire for automation and labor replacement. A high score indicates the vendor is successfully merging the software budget with the payroll budget. The assessment comprises the following component dimensions, each evaluated on a scale of 1 to 100. The final composite rating is derived from the geometric mean of these constituent scores.

1. **Unit economics (pricing page analysis):** This differentiates between legacy per-user subscriptions and modern per-outcome models. It awards maximum points to vendors charging by the resolution, report, or filing, while penalizing rigid seat-based licenses. This shift allows the vendor to capture the value of the work performed rather than just access to the tool.
2. **Interface intensity (product demo):** This assesses the necessity of human presence in the application, positing that minimizing time-in-app is now a positive metric. High scores favor asynchronous workflows that function as invisible workers rather than engagement-heavy dashboards. It rewards systems that complete tasks successfully without requiring constant human attention or login activity.
3. **Service capture (services page):** This evaluates the company's ability to capture professional services spend through high-touch automated implementation. A high score reflects a technology-enabled service model that absorbs the budget previously allocated to external vendors, including consulting firms. It indicates the vendor is replacing systems integrators with automated onboarding and configuration.
4. **Growth proxy (ROI calculator):** This examines whether the vendor's value proposition emphasizes efficiency or labor replacement. A score of 100 is reserved for companies explicitly marketing cost savings equivalent to full-time employee salaries or hiring avoidance. It penalizes calculators that focus only on soft metrics like time saved per task.
5. **Metering Structure (Terms of Service):** This checks for the existence of credits or usage meters to ensure margin protection against high inference costs until such time that inference costs decline substantially. A high score indicates a native ability to pass through AI compute costs to the customer via consumption pricing. This mechanism is essential for preserving the critical 60% gross margin threshold in the AI era.



Framework 2: Data moat & sovereignty score

Description: This framework distinguishes defensible systems of record from disposable wrappers by evaluating the ownership of the underlying data asset. It operates on the premise that algorithms are commodities, making proprietary context the only sustainable competitive advantage. This scoring aggressively penalizes companies that rely on public application programming interfaces (APIs) without owning the data layer where business truth resides. It rewards vendors that enforce strict sovereignty, preventing data leakage to generalist models. A high score implies the company owns the “truth” database required to ground AI agents. The assessment comprises the following component dimensions, each evaluated on a scale of 1 to 100. The final composite rating is derived from the geometric mean of these constituent scores.

1. **SoR ownership (integration docs):** This assigns high scores to primary databases like CRMs or ERPs where data originates and resides. It assigns low scores to overlay tools that rely on bi-directional syncs to function. Ownership of the SoR is the primary defensive moat against platform commoditization.
2. **Vertical exclusivity (technology page):** This assesses the ratio of proprietary data used to finetune the company’s models versus public data. A score of 100 implies a moat built on millions of unique, nonpublic records, such as legal or healthcare files. This proprietary tuning creates a barrier to entry for generalist foundation models.
3. **Scraper defense (API docs):** This evaluates the technical measures implemented to prevent unauthorized agents from accessing the platform’s data. A high score indicates aggressive API gating and anti-scraping clauses to prevent competitors from training on the company’s IP. It ensures the vendor retains the value of its aggregated data assets.
4. **Deployment sovereignty (security page):** This rewards walled-garden, private-cloud, or air-gapped deployment options that prevent data leakage. It penalizes multitenant public cloud models that co-mingle customer data with general training sets. Sovereignty is now a critical requirement for enterprise adoption in regulated industries.
5. **Learning loop (engineering blog):** This measures whether user actions and corrections directly retrain the proprietary model, thereby building a unique advantage. A high score means the data advantage compounds with usage, capturing the deep business logic of the enterprise. It creates a virtuous cycle in which the product improves faster than competitors can replicate it.



Framework 3: Incumbent resilience score

Description: Designed for public or late-stage VC and PE-backed companies, this framework assesses the structural moats that protect incumbents from AI-native disruption. It helps investors identify “compounders” that will successfully bundle AI features versus “value traps” that are structurally short of the AI revolution. The scoring evaluates the power of distribution and capital depth. Crucially, this framework treats these moats as conditional on speed. Incumbents that possess scale but fail to bundle AI features rapidly should be penalized as being at risk of displacement by vertical-native startups. A high score indicates the company possesses both the scale to out-distribute challengers and the agility to deploy “good enough” AI defensively before customers churn. The assessment comprises the following component dimensions, each evaluated on a scale of 1 to 100. The final composite rating is derived from the geometric mean of these constituent scores.

1. **Bundle power (product portfolio):** This measures the ability of a multisuite platform to offer AI features for free to preserve market share. A high score implies the vendor can out-compete point solutions by bundling its capabilities into an existing enterprise license. It rewards platform breadth that creates demonstrable barriers for single-feature startups.
2. **Distribution entrenchment (customer page):** This reflects a Fortune 500-standard status, with deep vendor approval integration and multiyear contracts. A high score strongly demonstrates resistance to churn because replacing the vendor requires a massive organizational lift. It disadvantages startups that must fight for every pilot through complex procurement processes.
3. **Switching cost severity (compliance page):** This awards high scores to mission-critical infrastructure where failure results in operational disaster. It assigns low scores to productivity tools where replacement is considered more of an annoyance to end users. These high switching costs provide incumbents with the time needed to catch up on AI features.
4. **Capital depth (financials):** This evaluates the balance sheet strength and free cash flow available to sustain massive AI capital expenditures. A high score indicates the company can outspend venture-backed challengers on compute infrastructure and strategic acquisitions. It ensures the incumbent vendor can survive the capital-intensive transition to agentic architectures.
5. **AI monetization (earnings call):** This checks for distinct AI revenue line items or premium SKUs that are able to sustain price hikes. A high score is given for direct monetization, while a low score applies if AI features are given away solely to reduce churn. It confirms that the vendor can capture value from AI rather than just deploying it defensively.



Framework 4: Technical autonomy score

Description: This framework judges the AI criticality of the platform, differentiating between passive co-pilots and active agents. It uses multiple signals of language model fidelity to determine whether the technology is robust enough for labor substitution. This scoring rewards systems that can execute independent write actions across the enterprise stack safely and successfully. It penalizes “chat-only” interfaces that require a human driver for every step. A high score means the software acts as a worker executing workflows, not just a tool. The assessment comprises the following component dimensions, each evaluated on a scale of 1 to 100. The final composite rating is derived from the geometric mean of these constituent scores.

1. **Agentic level (product claims):** This scores systems based on their ability to replace existing roles rather than just assisting tasks. A score of 100 is for systems marketed to hire an agent or replace Tier 1 support functions. Low scores go to co-pilots that require a human in the loop for every decision.
2. **Orchestration capability (integration page):** This measures the platform’s ability to execute write actions across external enterprise systems safely and successfully. A high score is for agents that can update CRMs, file tickets, or execute code independently. Low scores apply to agents trapped inside a chat window with no ability to affect the enterprise or its systems.
3. **AgentOps readiness (careers page):** This looks for organizational investment in agent operations, annotation, or model evaluation roles. A high score indicates the vendor is building the human infrastructure needed to manage autonomous systems. It suggests a pivot from pure software engineering to reliability engineering for probabilistic systems.
4. **Error handling & fidelity (trust center):** This checks for specific metrics regarding hallucination rates and built-in verification mechanisms. A high score indicates the presence of output guardrails that verify generated code or commands before execution. It separates robust enterprise-grade agents with successful deployment from fragile demos that crash in production.
5. **Task complexity (case studies):** This evaluates whether the agent handles high-compute System 2 reasoning tasks such as planning or coding. A high score is awarded for complex problem-solving that requires fluid intelligence and multistep logic with meaningful memory capacities. Low scores are given for simple summarization or text generation tasks that rely on crystallized intelligence.



Framework 5: Enterprise control plane score

Description: This framework measures the maturity of the governance layer, which is the primary moat for enterprise adoption. Enterprises will not deploy autonomous agents without a comprehensive audit trail and reliable undo capabilities. The scoring evaluates the visibility, reversibility, and security controls available to enterprise IT administrators. It penalizes “black box” systems that conceal their reasoning or act with broad permissions. A high score indicates the platform is safe enough for a CIO to approve for wide-scale deployment. The assessment comprises the following component dimensions, each evaluated on a scale of 1 to 100. The final composite rating is derived from the geometric mean of these constituent scores.

1. **Observability (product docs):** This looks for a “flight recorder” that tracks the agent’s reasoning chain and decision process. A high score requires transparent logs that explain exactly why an agent took a specific action. This visibility is essential for debugging and trusting autonomous systems in real-world production.
2. **Reversibility (admin page):** This indicates whether the system allows for granular rollback of agent actions. A high score is awarded for features that allow enterprises to undo a cascade of autonomous decisions. It mitigates the operational risk of an agent running amok and corrupting data, or worse.
3. **Identity mapping (security docs):** This ensures agents inherit existing role-based access control (RBAC) and user identity. A high score means the agent cannot access data or tools that the human user in its place would be restricted from. It prevents agents from bypassing security protocols via broad service account permissions.
4. **Evaluation harness (engineering blog):** This checks for the presence of vendor-supplied evaluation suites that test agents against regressions. A high score implies the vendor provides tools to continuously audit agent performance as models update. It ensures that improvements in one area do not degrade reliability in another.
5. **Policy guardrails (admin console):** This evaluates the existence of an admin console that allows managers to set strict bounded autonomy rules. A high score enables hard limits, such as “never refund more than \$50 without human approval.” It provides the “constitution” that governs the digital workforce’s behavior.



Framework 6: Cognitive architecture score

Description: This framework evaluates the operational efficacy of the agent, specifically its ability to plan, execute, and recover from errors. It separates fragile intelligence that breaks under stress from robust digital employees. The scoring assesses the agent's memory, tool use, and ability to handle ambiguity without crashing. It rewards architectures that integrate perception, reasoning, and action into a seamless loop. A high score indicates the agent possesses the fluid intelligence required for complex work. The assessment comprises the following component dimensions, each evaluated on a scale of 1 to 100. The final composite rating is derived from the geometric mean of these constituent scores.

1. **Goal decomposition (product demo):** This measures the agent's ability to break down high-level, ambiguous goals into executable microtasks. A high score is awarded for agents that can generate a logical plan from a vague prompt without human help. It indicates the agent understands the steps required to achieve a business outcome, as well as understanding the context of that request and why it would be given.
2. **Tool dexterity (APIs/docs):** This evaluates the agent's ability to drive other applications via an application programming interface (API) or directly via the computer. A high score means the agent can manipulate external software interfaces to complete tasks. It distinguishes agents that can "do" things successfully from chatbots that can only "say" things.
3. **Contextual continuity (features page):** This assesses whether the agent maintains stateful memory across sessions or projects. A high score is for agents that recall past decisions and user preferences over long periods. It prevents the frustration of having to re-explain context to a stateless bot with every instance.
4. **Adaptive resilience (white papers):** This measures the agent's ability to self-correct when a tool fails or an error occurs. A high score implies the agent can retry strategies or choose alternative paths without crashing. It separates brittle scripts from robust agents that can handle real-world ambiguity and complexity.
5. **Active perception (product features):** This evaluates the agent's ability to ingest real-time telemetry, logs, or visual data to build a world model. A high score means the agent proactively seeks out information rather than waiting for a text prompt. It allows the agent to react to changing environmental conditions in real-time as it observes them.



Framework 7: AgentOps & workforce maturity score

Description: This framework measures the organization's ability to manage the autonomy gap and the operational realities of a digital workforce. It assesses whether the vendor provides the infrastructure to govern agents and the training to retool human workers. The scoring focuses on transparency, role management, and cost governance for autonomous operations. It rewards vendors that treat agents as a new asset class requiring specialized management tools. A high score indicates the vendor is ready to support the customer's transition to an agent-centric operating model. The assessment comprises the following component dimensions, each evaluated on a scale of 1 to 100. The final composite rating is derived from the geometric mean of these constituent scores.

- 1. The autonomy gap metric (SLA/trust center):** This checks if the vendor transparently reports intervention rates or resolution success rates. A high score is for vendors that treat autonomy failures as a KPI to be minimized. It provides the enterprise customer with a realistic view of the labor subsidy required.
- 2. Workforce retooling (training page):** This looks for the availability of GenAI Simulators to prevent skill atrophy in human operators. A high score indicates the vendor offers training environments for humans to practice high-stakes supervision. It ensures humans remain capable of intervening when the agent fails.
- 3. Specialized role support (docs):** This evaluates RBAC support for new roles such as policy engineers, autonomy curators, and risk governors. A high score indicates that the platform recognizes the new division of labor in an AI-native enterprise. It allows organizations to assign specific governance responsibilities to distinct human operators.
- 4. Inference efficiency & cost governance (pricing page):** This assesses mechanisms to optimize the cost of thought and to set budget caps per task based on each task's complexity and criticality within the enterprise. A high score is for platforms that prevent runaway inference costs by managing the compute resources allocated to agents. It ensures the economic viability of replacing human labor with digital labor.
- 5. Hybrid fleet control (architecture docs):** This rewards platform-centric infrastructure that manages agents across on-premises, edge, and cloud environments. A high score indicates the vendor supports a unified control plane for a distributed agent fleet. This is essential for enterprises with complex, hybrid IT landscapes.



Framework 8: Agent fidelity matrix (product agency score)

Description: This framework applies a rigorous 10-point product audit to measure the specific cognitive performance of a software product or a collection of products acting as one solution. It serves as a deep-dive reality check on the product's actual ability to function as a fiduciary operator. A failure in any single dimension significantly penalizes the overall score via the geometric mean calculation. The scoring demands excellence across cognitive, operational, and governance layers to ensure reliability. A high score reflects that the agent is not just a demo, but a deployable digital employee. The assessment comprises the following component dimensions, each evaluated on a scale of 1 to 100. The final composite rating is derived from the geometric mean of these constituent scores.

1. **Goal decomposition:** This measures the capacity to accept a high-level, ambiguous objective and autonomously break it down into a logically sequenced plan. It evaluates whether the agent can translate business intent into technical steps without human intervention. A failure here means the agent requires constant micromanagement.
2. **Bounded autonomy:** This evaluates the degree to which an agent operates independently while strictly adhering to pre-defined safety protocols. It checks whether the agent recognizes its limits and pauses for approval before executing high-risk actions. This is the primary safety valve preventing runaway automation.
3. **Active perception:** This assesses the ability to proactively query and ingest real-time data from logs and metrics to build a world model. It captures whether the agent can independently seek out necessary context rather than relying solely on prompts. Without this, the agent is blind to the environment's changing state.
4. **Causal reasoning:** This measures the application of System 2 thinking to infer actual causality rather than just correlation. It differentiates between simple pattern matching and the deep reasoning required to diagnose root causes. This capability is essential for handling novel "unknown unknown" incidents.
5. **Tool proficiency:** This evaluates an agent's dexterity in manipulating external APIs, CLIs, and GUIs to implement changes. It measures the agent's technical effectiveness in interacting with the existing software stack. High proficiency ensures the agent can actually do the work, not just describe it.
6. **Adaptive resilience:** This assesses the capability to detect execution failures, self-correct strategies, and attempt alternative paths without crashing. It measures robustness when faced with unexpected environmental changes or tool errors. A resilient agent can recover from mistakes without requiring human rescue.
7. **Contextual continuity:** This measures the maintenance of persistent semantic memory, allowing the agent to recall past decisions and user preferences. It ensures the agent learns from historical events to inform future actions across different sessions. This builds a "long-term memory" of the specific enterprise environment.



8. **Collaborative orchestration:** This evaluates the ability to recognize task boundaries and seamlessly hand off control to other specialized agents or humans. It checks how well the agent functions as a coordinated team player. This prevents agents from working in silos and conflicting with one another.
9. **Verifiable explainability:** This assesses the generation of human-readable audit trails and reasoning logs that explain exactly why an action was taken. It ensures that every autonomous decision is transparent and can be forensically reviewed for compliance. This builds the trust necessary to remove humans from the loop.
10. **Inference efficiency:** This measures the ability to optimize the cost of thought by dynamically adjusting compute resources based on task complexity. It evaluates whether the agent uses “fast mode” for simple tasks and “reasoning mode” only when necessary. This dimension determines the financial viability of mass-deploying the agent.

Top 20 SaaS VC investors by deal count since 2019

Investor	Total deal count	Pre-seed/seed	Early-stage VC	Late-stage VC	Venture growth	Investor type
Alumni Ventures	582	193	180	176	33	VC
Sequoia Capital	517	130	185	140	62	VC
Gaingels	517	125	144	203	45	VC
Accel	468	84	168	166	50	VC
Andreessen Horowitz	460	92	192	136	40	VC
Tiger Global Management	450	20	143	197	90	VC
General Catalyst	373	91	134	103	45	VC
FJ Labs	341	101	118	100	22	VC
Lightspeed Venture Partners	330	78	128	88	36	VC
Salesforce Ventures	326	25	87	163	51	CVC
Index Ventures	286	57	123	83	23	VC
Global Founders Capital	271	134	94	33	10	VC
Soma Capital	242	132	88	20	2	VC
BoxGroup	210	94	84	27	5	VC
GV	206	32	65	69	40	CVC
Khosla Ventures	202	43	69	61	29	VC
Kleiner Perkins	198	27	73	66	32	VC
8VC	196	54	72	54	16	VC
Founders Fund	192	36	73	57	26	VC
Bain Capital Ventures	192	42	76	58	16	VC

Source: PitchBook • Geography: Global • As of December 31, 2025



Top 20 SaaS PE investors by deal count since 2019

Investor	Deal count	Investor type
TA Associates Management	71	PE/buyout
Hg	67	PE/buyout
Thoma Bravo	61	PE/buyout
Insight Partners	56	PE/buyout
Vista Equity Partners	54	PE/buyout
Blackstone	47	PE/buyout
PSG	46	PE growth/expansion
Accel-KKR	42	PE/buyout
Kohlberg Kravis Roberts	40	PE/buyout
Warburg Pincus	39	PE/buyout
TPG	38	PE/buyout
Sixth Street Partners	37	PE growth/expansion
Francisco Partners	34	PE/buyout
Aquiline	34	PE/buyout
Silver Lake	29	PE/buyout
Clearlake Capital Group	29	PE/buyout
Advent International	27	PE/buyout
General Atlantic	26	PE growth/expansion
The Carlyle Group	25	PE/buyout
K1 Investment Management	24	PE/buyout

Source: PitchBook • Geography: Global • As of December 31, 2025



Top 20 SaaS strategic investors by deal count since 2019

Investor	Deal count	Investor type
Fullsteam Operations	23	PE-backed
Descartes Systems Group	19	Corporation
Salesforce	18	Corporation
The Access Group	13	PE-backed
Avalara	11	PE-backed
Fiserv	11	Corporation
ServiceNow	11	Corporation
International Business Machines	11	Corporation
Motorola Solutions	11	Corporation
NVIDIA	10	Corporation
Snowflake	10	Corporation
Verisk Analytics	10	Corporation
SPS Commerce	9	Corporation
Visma	9	PE-backed
Shift4 Payments	9	Corporation
Tyler Technologies	9	Corporation
Atlassian	9	Corporation
Workday	9	Corporation
Sage Group	9	Corporation
Medallia	9	PE-backed

Source: PitchBook • Geography: Global • As of December 31, 2025



The new vocabulary for enterprise AI

The transition from SaaS to SaS also requires a new, universally understood lexicon. Growth and operational metrics, such as “seats” and “daily active users,” will eventually be replaced with digital labor outcomes metrics and benchmarks.

This glossary attempts to define the specific economic, technical, and operational mechanisms that determine a company’s ability to survive the merger of software and labor. We strongly recommend stakeholders become aware of these concepts, as they provide the foundational logic for the SaS-ability frameworks and the future of enterprise valuation.

Glossary

1. **Active perception:** The ability of an agent to proactively query and ingest real-time data to build a world model. It captures whether the agent can seek out necessary context independently rather than waiting for a prompt. This prevents agents from working with stale information and allows them to react to changing environmental conditions.
2. **Adapters (low-rank adaptation):** Specialized layers added to a general AI model that encode deep domain expertise, such as legal citation rules or medical billing codes, without the need to retrain the entire system. They enable the rapid, inexpensive creation of thousands of vertical agents tailored to highly specific industries and niche tasks. This technical approach enables enterprises to build a proprietary improvement loop by constantly refining the agent’s performance on specialized data.
3. **Adaptive resilience:** An agent’s ability to detect execution failures and self-correct strategies without crashing. It distinguishes robust agents from brittle scripts that fail at the first error. It assesses the agent’s robustness in the face of unexpected environmental changes or tool failures.
4. **Agent fidelity:** The comprehensive assessment of an agent’s performance across the key dimensions of reasoning, accuracy, tool use, reliability, safety, speed, cost-efficiency, and long-term memory. It describes the extent to which a digital worker can operate in a production environment autonomously without human oversight or intervention. High fidelity indicates a trustworthy digital employee capable of executing complex workflows with the same consistency and institutional nuance as a high-performing human professional.
5. **Agent orchestration:** The software layer that manages the handoffs and coordination between multiple specialized agents. It ensures that complex workflows requiring different skills are executed smoothly. This layer is essential for preventing conflicts and ensuring task continuity.



6. **Agent-to-agent (A2A):** The interaction protocol enabling specialized autonomous agents to recognize task boundaries and seamlessly hand off control to other agents. This allows for the creation of “crews” (for example, a researcher, a writer, and an editor) that collaborate to solve cross-functional problems that exceed any single agent’s domain.
7. **AgentOps:** The “last mile” operational practice responsible for managing the life cycle, behavior, and resilience of autonomous agents. Distinct from DevOps, it focuses on agent fidelity, compliance, and outcome reliability. It is the new IT support function required to govern the digital workforce.
8. **Agentic AI:** The evolution of software from passive tools to autonomous entities capable of independently executing tasks such as planning and tool use. It unlocks high-value automation markets requiring reliability and security beyond simple chatbots. This shift moves software from an advisor role to an active worker role.
9. **Artifacts:** Tangible outputs produced by an agent, such as documents, code, or charts. These serve as proof of work for audit trails and verifiable explainability. They distinguish a tool that helps you work from a worker that produces actual results.
10. **Autonomy curator:** A human role focused on managing the knowledge base and context of agents to prevent drift. They maintain the “brain” of the digital workforce by pruning and updating semantic memory. This prevents the agent from making decisions based on stale or incorrect data.
11. **Autonomy gap:** The quantifiable operational cost (manual fixes, training, recovery) incurred when an AI agent requires human intervention. It represents the subsidy required to keep a self-driving tool running and is the primary metric IT must manage. Reducing this gap is the primary goal of the AgentOps function.
12. **Best of suite:** An incumbent software company that successfully bundles AI capabilities into a broad product portfolio to preserve market share. These consolidators leverage their distribution power to out-compete point solutions and capture service revenue previously flowing to external vendors, including consulting firms.
13. **Bounded autonomy:** The degree to which an agent operates independently while strictly adhering to pre-defined safety protocols. It evaluates the agent’s ability to recognize limits and pause for approval. This creates a safe sandbox for autonomous action, preventing runaway automation.
14. **Causal reasoning:** The application of System 2 thinking to infer actual causality rather than just correlation. Essential for diagnosing root causes in IT operations and distinguishing signal from noise. It enables the agent to solve novel incidents that are not in its training data.



- 15. Chain of thought (CoT):** A reasoning technique where the model generates hidden “thinking tokens” to plan before answering. This increases cost and latency but significantly boosts accuracy for complex logic. It enables models to move from reflexive answers to deliberate, multistep reasoning.
- 16. Collaborative orchestration:** The ability of an agent to recognize task boundaries and hand off control to other agents or humans. It evaluates how well the agent functions as a coordinated team player. This prevents agent-on-agent interference in complex, multistep workflows.
- 17. Command center:** The centralized management interface providing real-time visibility and human oversight across a firm’s entire fleet of digital workers. It functions as a single pane of glass for risk governors and executives to monitor performance, review intervention rates, and manage high-level control over agent activities. This layer focuses on operational transparency and decision-support, allowing human leaders to redirect agent goals or approve high-liability actions.
- 18. Compute volume:** The metric of value capture for model providers, representing the total cost of thinking. It differentiates the model provider business model from the SaaS application model. As models think longer (System 2), this volume increases exponentially per task.
- 19. Computer use/pixel agent:** An agent’s ability to control a mouse and keyboard to interact with legacy software user interfaces (UIs). It bridges the gap between modern APIs and “green screen” enterprise apps. It is the universal adapter for automating workflows in systems without API access.
- 20. Context window:** The amount of information (tokens) a model can hold in short-term memory. Larger windows enable analysis of entire books or codebases, a critical feature for enterprise use. It reduces the need for complex retrieval systems by holding more data in active memory.
- 21. Control plane:** The underlying technical infrastructure and governance layer that enforces security, identity, and policy protocols across all agent interactions. It provides back-end connectivity, managing authentication, data sovereignty rules, and the immutable audit logs required for regulatory compliance. This architecture enables policy engineers to program deterministic guardrails that automatically prevent agents from accessing unauthorized systems or violating corporate mandates.
- 22. Cost of intelligence:** The quantifiable price per task to achieve a specific level of performance on reasoning benchmarks. Intelligence is highly elastic to computing power, meaning higher fidelity requires exponentially more resources. Currently, in Q1 2026, achieving human-level reasoning costs between \$1 and \$10 per task according to the ARC-AGI-2 assessment.⁴

⁴: “ARC-AGI-1 Leaderboard,” ARC Prize, 2025, accessed on February 24, 2026.



- 23. Crystallized intelligence:** Pre-learned or memorized knowledge (training data) that an agent draws upon. Current models are expert-level here, making knowledge retrieval a commodity. It contrasts with fluid intelligence, which is required for novel problem solving.
- 24. Data fabric:** A unified architecture connecting data across on-premises, cloud, and edge environments—a prerequisite for AI-native platforms to access information anywhere. This breaks down the silos that prevent agents from seeing the whole picture, both within and externally to an enterprise.
- 25. Data ingress/egress:** Pathways for moving data in and out of model environments. Control over this is a key walled-garden feature and a variable in cloud costs. Incumbents often charge for egress (or deny it altogether) to prevent startups from training on their data.
- 26. Deep business logic:** The complex, implicit rules and decision trees that govern high-value enterprise workflows, such as structuring a complex credit derivative or handling a compliance exception. Historically hardcoded by engineers, this logic can now be captured in natural language files that domain experts can write directly.
- 27. Deployment lag:** The typical 18-month gap between infrastructure capital expenditure and application value realization. This explains the current disconnect between massive hyperscaler spending and steady enterprise adoption rates. Investors must account for this lag when modeling ROI for AI hardware.
- 28. Deterministic logic:** Hardcoded “if-then” rules that an agent cannot override. Used for compliance and safety where creativity is unacceptable. Incumbents use this to wrap deep business logic around unpredictable LLMs.
- 29. Digital worker:** An autonomous software entity designed to replace human roles (SDR, QA engineer, for example) rather than merely assisting them. Unlike traditional software tools that require human operation, digital workers execute end-to-end workflows independently.
- 30. Engine manufacturers:** Providers of raw LLM intelligence (for example, OpenAI, Anthropic) distinct from the domain-specific platforms that govern them. They capture value via compute volume (the cost of thinking) rather than workflow stickiness. They provide the fuel, while SaaS vendors build the car.
- 31. Enterprise AI:** The systemic deployment of autonomous agents and cognitive infrastructure designed to execute core business workflows rather than merely assisting human users. Unlike consumer models, it requires strict adherence to corporate governance, data sovereignty, and determinative safety protocols to operate within regulated environments. This ecosystem integrates stateful architectures with legacy systems of record, allowing digital workers to read, reason, and write back to the enterprise database with full auditability. It represents the transition from software as a productivity tool to software as a labor substitute,



capable of closing the gap between autonomous potential and the human intervention currently required to fix errors. Ultimately, it is the strategic imperative to build a proprietary digital workforce that compounds in value through the continuous curation of institutional knowledge.

- 32. Exception handling:** Pre-programmed workflows for when an agent fails to complete a task. Robust handling prevents brittle crashes and allows for graceful degradation. It is the difference between a demo and a production product.
- 33. Expert intelligence:** The level of capability where an AI model achieves 100% scores on human-expert benchmarks in knowledge-based fields. This marks the first phase of AI commercialization, enabling high-ROI automation in research & development and coding.
- 34. Federated context:** Accessing data across multiple siloed apps without moving it to a central lake. It trades deep integration for broader visibility and faster deployment. This approach accelerates adoption by skipping the data migration phase.
- 35. Finetuning:** Training a base model on specific domain data to create a specialized expert. Essential for vertical agents to build a moat against generalist foundation models. It is a critical step for rebuilding deep business logic and capturing proprietary context.
- 36. Flight recorder:** An observability tool that logs the reasoning trace and thought process of an agent. Essential for auditing why an agent made a decision, a requirement especially for deployment within regulated industries. It provides the transparency needed for verifiable explainability and compliance.
- 37. Fluid intelligence:** Novel reasoning capabilities required for agents to handle unknown scenarios, as opposed to memorized knowledge. This is the new frontier for agentic capabilities and requires massive compute scaling. Achieving expert-level fluid intelligence is required for replacing high-value human labor.
- 38. GenAI simulators:** Virtual training environments for human operators used to practice high-stakes decisions, used to mitigate skill atrophy in those operators. They are essential for preventing skill atrophy in a world where agents do most of the work. They prepare the workforce for the supervisor role.
- 39. Goal decomposition:** The ability of an agent to break down ambiguous high-level objectives into logically sequenced, executable technical plans. It measures the agent's ability to translate business intent into technical steps without human help. This capability is the primary indicator of operational fluid intelligence.
- 40. Grounding:** Anchoring AI responses in retrieved enterprise data to prevent fabrication (hallucinations). The technical mechanism that turns a creative storyteller into a reliable analyst. It forces the model to cite its sources from internal documents and work from them.



- 41. Human-aligned model:** The mature operating model where humans supervise autonomous agents and only intervene during exceptions. This model requires high system trust and low error rates. It represents the target state for the SaS transition.
- 42. Human-managed protocols:** The initial operating model where agents propose actions but require human validation before execution. Used for high-risk tasks to prevent catastrophic errors during the early adoption phase. This is the starting point for most enterprise AI pilots.
- 43. Inference efficiency:** The ability to optimize the cost of thought by dynamically adjusting compute resources based on task complexity. It involves routing simple queries to smaller, faster models and complex problems to larger reasoning models. This metric determines the financial viability of mass-deploying autonomous agents.
- 44. Input/output guardrails:** Safety filters deployed to detect and block unsafe prompts (input) and verify generated outcomes (output). Without these, the operational risk of autonomy outweighs the efficiency gains. They act as the firewall for the agent's cognitive process.
- 45. Intelligence elasticity:** The economic concept that AI intelligence and fidelity scale directly with the amount of computing power allocated to it. Demand for compute is effectively uncapped because more compute is understood to always yield higher fidelity and utility. It refutes the idea that AI demand will plateau.
- 46. Intervention rate:** The percentage of agent actions requiring human correction. The inverse proxy for autonomy, the lower this rate, the higher the SaS value. Investors use this trendline to value this technology's deflationary potential.
- 47. Knowledge graph:** A map of entities and relationships serving as the agent's long-term memory. Essential for contextual continuity and understanding hidden connections. It allows the agent to reason about the business structure, not just process text.
- 48. Labor arbitrage:** The legacy business model of BPOs (billing for human hours), which SaS targets for replacement. AI agents attacking this \$50+ trillion market drives the deflationary economics of the sector by replacing human services with software. It represents the largest TAM expansion in software history.
- 49. Last mile of autonomy:** The specific operational gap between a general-purpose model and a reliable, governed enterprise agent. This gap is addressed by AgentOps through tools that manage life cycle, compliance, and resilience, ensuring agents can function in production environments.
- 50. Language model fidelity (LMF):** A metric measuring how accurately a model follows user intent and adheres to factual ground truth. It is the single most critical



performance metric for agentic systems, as low fidelity leads to costly failures. Perfect language model fidelity is required to remove human-in-the-loop oversight.

- 51. Model context protocol (MCP):** An open standard that enables AI models to securely connect to and access data from content repositories and business tools. It standardizes how agents discover and consume external context, solving the “many-to-many” integration problem between models and data sources.
- 52. Markdown skill files:** Text-based instructions written in natural language that define an agent’s specific capabilities and workflows, such as “How to run a cash flow analysis.” This format replaces complex, hardcoded logic in traditional software, allowing nontechnical domain experts to “program” agents by simply writing a document. It represents an elimination of the engineering barrier to entry, reducing development cycles from months to days.
- 53. Model routing:** Dynamically dispatching tasks to the cheapest or best model for the job. Optimizes cost per task and preserves margins for the vendor. It allows vendors to arbitrage intelligence costs behind the scenes.
- 54. Multi-agent systems (MAS):** Architectures where specialized agents (such as researchers and writers) collaborate to solve complex goals. This allows for goal decomposition by breaking down tasks and delegating these into expert domains. This approach generally outperforms single massive models on complex benchmarks.
- 55. Multimodal:** AI that can process and generate multiple data types such as images, audio, and text simultaneously. Essential for physical AI and agents that need to “see” computer screens to navigate legacy software. It allows the AI to perceive the work environment as a human does.
- 56. Net dollar retention (NDR):** A traditional SaaS metric measuring revenue growth from existing customers, which is becoming insufficient for valuing AI-native companies. While historically the gold standard for software health, it fails to capture the massive expansion of the TAM for labor replacement. In the SaaS era, this metric must be augmented with net labor capture to reflect the true economic value of the platform.
- 57. Net labor capture:** A new KPI measuring the percentage of a customer’s total payroll budget successfully migrated into their software spend. Unlike traditional net dollar retention, which tracks software budget expansion, this metric specifically quantifies the substitution of human labor costs with digital labor fees. It provides the clearest signal of a vendor’s strategic value and ability to monetize the autonomy gap.
- 58. Outcome-based pricing:** A business model where vendors charge for the successful completion of tasks rather than per seat. This is the primary mechanism for capturing the value of AI, as vendors must pivot to this to avoid



revenue cannibalization as AI reduces headcount. It aligns vendor incentives with the customer's desire for automation.

- 59. Platform-centric infrastructure:** A unified foundation enabling central control of agents across on-premises, edge, and public cloud environments. The last mile infrastructure for AgentOps. It creates a seamless control plane for the distributed digital workforce.
- 60. Policy engineer:** A new enterprise role responsible for coding the rules and compliance boundaries that govern agent behavior. They translate corporate policy into machine-readable constraints (the constitution). This ensures agents act within legal and ethical limits.
- 61. Policy store:** A database housing the constitution, or rules for agents. This allows for global governance updates without retraining models. It enables nontechnical managers to update safety rules instantly across the fleet, and thus must be securely protected.
- 62. Retrieval-augmented generation (RAG):** A technique that enhances model accuracy by fetching external data to include in the prompt. RAG solves hallucination problems and is a critical enabler for enterprise adoption. It allows the model to “take an open book test” by anchoring answers in enterprise truth.
- 63. Regression:** When a model's performance degrades on existing tasks after an update. This requires continuous evaluation harnesses to detect regression before it impacts revenue. This phenomenon makes trust a moving target in AI operations.
- 64. Reliability architect:** A senior IT role responsible for designing the governance frameworks and guardrails for the agentic enterprise stack. They ensure systemic resiliency and prevent agent-induced cascades. This role bridges the gap between technical operations and corporate risk policy.
- 65. Retrieval schemata:** Blueprints defining how an agent searches and interprets data during RAG. The “librarian logic” ensuring correct citations and relevance. A well-designed schema ensures the agent finds the specific clause needed, not just a keyword match.
- 66. Risk governor:** A role ensuring that autonomous agent actions remain auditable and compliant with enterprise risk standards. They have the authority to trigger kill switches if agents misbehave. They provide the control layer for the digital workforce.
- 67. Reinforcement learning from human feedback (RLHF):** Improving a model by having humans rank its outputs to align it with human preferences. A key method for building a data moat by capturing proprietary user feedback. It creates a virtuous cycle in which the model gets smarter and better aligned with usage.



- 68. SaaS-pocalypse:** A market phenomenon where SaaS valuations collapse due to the fear of AI displacement. It marks the moment investors began repricing software based on labor-replacement potential rather than just productivity. This fear is the accelerator forcing legacy vendors to pivot to SaaS.
- 69. Service as software (SaS):** The post-SaaS business model where vendors sell labor outcomes rather than tool access. Revenue is derived from work units completed (for example, per resolution) rather than seats, capturing the 8x revenue uplift of labor replacement. This shifts the value metric from software access to business outcome.
- 70. SaS-ability:** A vendor's structural capacity to pivot from selling access to software tools to selling completed digital labor outcomes. It measures the technical and operational maturity required to minimize the need for human intervention, allowing software to function as an independent worker rather than a passive instrument. This capability determines a vendor's potential to capture customer payroll budgets by replacing expensive human effort with automated results. It distinguishes companies that merely wrap public data from those that own the proprietary data and workflows needed for defensible automation. Ultimately, it indicates a sustainable business model that aligns revenue with the value of work performed rather than user headcount.
- 71. Shadow AI:** Unregulated or undocumented AI agents running within an enterprise environment without central governance. It creates significant IP leakage risks and compliance violations. CIOs must counter this by demanding walled gardens and enterprise-grade privacy.
- 72. Sidecar pattern:** An integration strategy where an agent runs alongside a legacy app, observing and acting without altering core code. A strategy to bypass switching costs and modernize legacy systems. It allows agents to automate systems of record without needing deep API permission.
- 73. Skill atrophy:** The degradation of human technical skills when agents handle all routine tasks. A major talent risk that organizations must mitigate through simulation training. It creates a dependency where humans can no longer intervene during catastrophic failure.
- 74. Stateful architecture:** The technical design that allows agents to retain memory and context over long periods, unlike stateless chatbots that reset after every session. This structure is required to replace human project managers and handle multiday workflows, as it enables the agent to pick up exactly where it left off. It turns the agent from a transactional tool into a continuous employee capable of complex, long-running tasks.
- 75. System of action:** The layer of the operational stack where agents execute changes (for example, ServiceNow and Jira). Agents must integrate here to be useful, moving beyond the system of intelligence (Chat). This is where the actual work units are completed and value is realized.

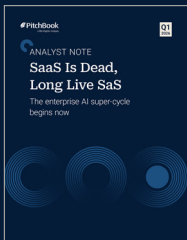


- 76. System of record (SoR):** The “truth” database (CRM, ERP) where critical business data originates and resides. Owning the SoR is the primary defensive moat for incumbents, as agents cannot function effectively without deep access to this proprietary context. It prevents AI solutions from being commoditized by generic foundation models.
- 77. System of workflow automation:** The target state for SoR providers to aggressively pivot to the agentic AI era. It describes platforms that use deep domain expertise to automate entire processes rather than just recording data. This shift allows incumbents to compete directly with AI-native challengers.
- 78. The 60% rule:** A financial covenant metric stating that if a software company’s gross margin drops below 60% due to AI costs, the model is broken. It forces companies to align pricing with their compute consumption to ensure long-term viability. Investors use this to identify vendors failing to pass inference costs to customers today.
- 79. Value trap:** A legacy company that looks cheap but faces structural obsolescence because its revenue is tied to human seats. As AI reduces headcount, their TAM shrinks, making them a dangerous investment despite, in some cases, high current cash flows. Investors are urged to sell these assets unless they aggressively pivot to outcome pricing.
- 80. Verifiable explainability:** The generation of human-readable audit trails and reasoning logs explaining exactly why a specific action was taken. This ensures that every autonomous decision is transparent and can be forensically reviewed for compliance. It builds the trust necessary to remove humans from the loop.
- 81. Vertical agent:** An AI agent trained on proprietary, niche data (dental billing, for example) that outperforms generalist models. These are the primary targets for VC funding as they build defensible data moats against foundation models. They succeed by rebuilding deep business logic specific to that vertical that generic models cannot replicate.
- 82. Workflow wrapper:** A startup developing a thin UI layer over a public foundation model that facilitates processes without owning the underlying data. These companies are considered high-risk investments because they lack defensive moats and are vulnerable to rapid commoditization by incumbents.
- 83. Work unit:** The new atomic unit of economic value (for example, a resolved ticket) in SaaS models. Replaces seats as the growth metric and focus of pricing. It shifts the focus from user engagement to labor completion.



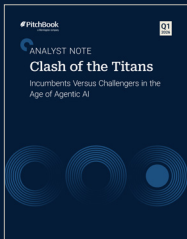
PitchBook provides actionable insights across the global capital markets.

Additional research:



Q1 2026 Analyst Note: SaaS is Dead, Long Live SaaS

Download the report [here](#)



Q1 2026 Analyst Note: Clash of the Titans

Download the report [here](#)

PitchBook Insights is an online compendium of in-depth data, news, analysis, and perspectives that shape the private capital markets.

PitchBook subscribers enjoy exclusive access to a comprehensive suite of private market insights, including proprietary research, news, data, tools, and more on the [PitchBook Platform](#).

Nizar Tarhuni

Executive Vice President of Research and Market Intelligence

Paul Condra

Senior Director, Global Head of Private Markets Research

James Ulan

Director, Industry & Technology Research

Report created by:

Rudy Torrijos

Director, Industry Research

Derek Hernandez

Senior Research Analyst

Oscar Allaway

Senior Data Analyst

Caroline Suttie

Senior Manager, Design

Learn more about [PitchBook's Institutional Research team](#).

Click [here](#) for PitchBook's report methodologies.

COPYRIGHT © 2026 by PitchBook Data, Inc. All rights reserved. No part of this publication may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, and information storage and retrieval systems—without the express written permission of PitchBook Data, Inc. Contents are based on information from sources believed to be reliable, but accuracy and completeness cannot be guaranteed. Nothing herein should be construed as any past, current or future recommendation to buy or sell any security or an offer to sell, or a solicitation of an offer to buy any security. This material does not purport to contain all of the information that a prospective investor may wish to consider and is not to be relied upon as such or used in substitution for the exercise of independent judgment.