



THE STATE OF AI EXECUTION, 2026

WORK, DELIVERED

The shift to AI that delivers work autonomously, with people making the decisions that matter.

Aaron Sipper

Founder & Chief Architect

OCTOBER 2026

PURPOSE AND SCOPE

AI has moved quickly in 2026. Nearly nine in ten organisations now use AI regularly. The big platforms have all announced agents that persist and work in teams, and investors and analysts increasingly expect money to move from software seats to finished work. Running costs are easier to see, too. Some of the largest providers now bill agent use by usage, and managing AI spend has become standard practice.

The results haven't caught up. Fewer than four in ten organisations attribute any profit impact to AI, only about one in six have deployed AI agents, and analysts expect more than 40% of agentic AI projects to be cancelled by the end of 2027. The more useful question now is whether a business can get AI work done end to end, reliably and on its own terms. We call the distance between using AI and getting results from it the delivery gap.

This paper is about that question. We don't compare AI models here, or the agent software built around them. We look at the whole system around the work: the knowledge, procedures and policies that define it, the people who make the calls, the AI teams that do it, and the disciplines that decide whether it actually gets finished.

It draws on 2026 research and our own build learning. The research comes from analysts, investors and platform providers. The learning comes directly from designing, building and running AI teams ourselves.

SUMMARY OF FINDINGS

AI can now do real business work, and in our experience results come when that work runs end to end inside a system that manages it.

- 1

Most organisations use AI. Few get value from it at scale.

Nearly nine in ten use AI regularly, but only 37% attribute any profit impact to it, and several studies put the group getting value at scale at around 5 to 8%. Much of the use is unofficial.

p.4
- 2

The market is moving from selling tools to selling finished work.

Investors estimate six dollars go on services for every dollar on software, buyers are leaving per-seat pricing, and in 2026 investors began buying and building the firms that deliver the work.

p.5–8
- 3

Outcomes depend on the whole system around the work.

Business knowledge, SOPs and policy define the work, and people make the judgement calls. AI teams do the work in between and keep a record that shows it was done.

p.9
- 4

Finishing the work comes down to nine disciplines.

About one in six organisations have deployed AI agents. Models still fall short, but research finds most failures come from how the work is run.

p.10
- 5

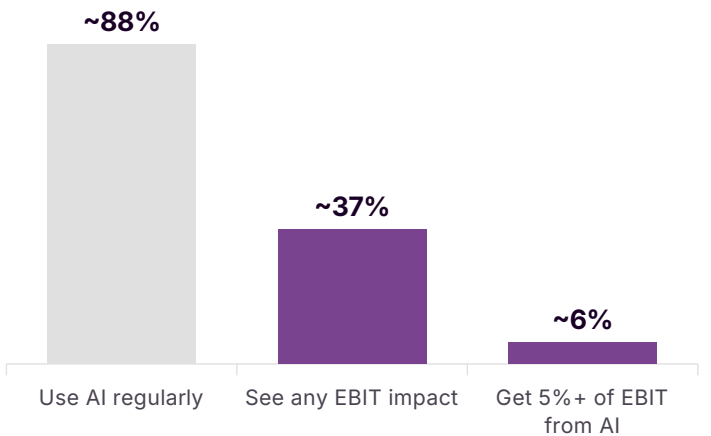
Cost and data location now shape how AI gets deployed.

Agent work uses far more tokens and providers now bill it by usage. Gartner expects sovereign cloud spending to nearly double in mature Asia-Pacific in 2027.

p.11

From using AI to getting a return

% of organisations, 2026 global survey (detail on p.4)



WHAT IT MEANS FOR LEADERS

Most organisations are already using AI, often at real cost and effort. Getting from the first bar to the last is the delivery gap, and a better model alone won't close it: in the 2026 data it follows how work is redesigned, run and owned. Solving it is a leadership decision about how work gets done, not an IT rollout.

AI ADOPTION AND MEASURED IMPACT

AI use has spread faster than organisations have changed how they work, and measurable results have not kept up.

Most organisations use AI. McKinsey's 2026 survey of 1,719 organisations found nearly nine in ten using it regularly in at least one function, and 40% of large organisations now say they are scaling AI agents, up from 27% a year earlier.

Few get value at scale. Only 37% attribute any EBIT impact to AI, about the same as a year ago. Studies draw the line in different places but land in the same range: about 6% are McKinsey's high performers, 7.5% are "future-built" in BCG's 2026 index, and in KPMG's mid-2026 Asia-Pacific survey only 5% have established a return with meaningful outcomes.

How work runs makes the difference. Nearly three-quarters of McKinsey's high performers redesigned their workflows, against about a quarter of everyone else. Deloitte finds only 30% redesigning key processes around AI, while 37% use it at surface level with little change underneath.

A lot of it is unofficial. Two in three office professionals at large firms in Australia, Japan, the UK and the US had used AI they believed wasn't permitted. In WalkMe's 14-country study, a vendor survey, 45% of workers had used unsanctioned AI tools in the past month.

How office professionals at large companies use AI

% of respondents at firms with US\$500M+ revenue (Wakefield / PagerDuty, Jun 2026)



FROM PRACTICE

Activity can look like output. A chart, or a report saying the job succeeded, doesn't show that the information is current or that the action happened.

WHAT IT MEANS FOR LEADERS

For each AI tool and seat, ask which business outcome it's accountable for. If nobody can say, you're paying for activity, not business outcomes.

FROM ASSISTANCE TO DELIVERED WORK

The shift is from buying AI tools to paying for finished work that is delivered, checked and recorded.

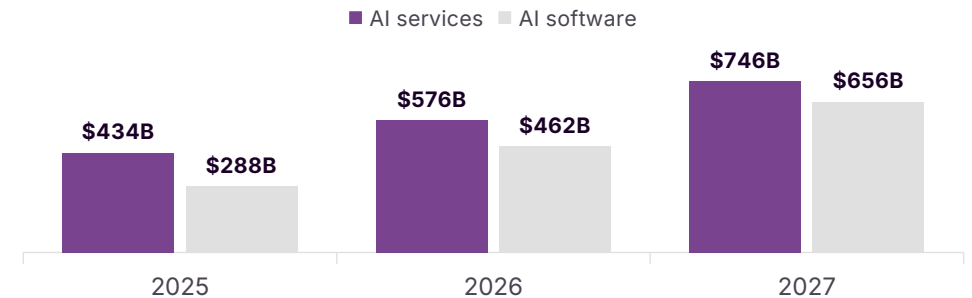
Three stages. At first a business buys a tool or a seat, and an employee does the work prompt by prompt. Next, agents take on specific steps, but people still stitch them together. In the third stage the business buys a result from a defined workflow: AI teams do the work and people approve the decisions that matter.

Why the money follows. Across the economy, Sequoia estimates six dollars go on services for every dollar on software, and that services budget is what AI doing the work can reach. Within AI spending itself, Gartner forecasts more on services than software in 2026, and its fastest-growing software is agents and assistants, the software that does the work.

Others see it from different angles. Y Combinator's Summer 2026 call asked for AI-native service companies, and Bessemer judges them on software-like margins. Gartner expects agents to put up to US\$234 billion of application software spend in play by 2030. In G2's mid-2026 survey of 1,019 B2B buyers, only 9% saw per-seat pricing as fairest for agents, while paying per outcome rose from 11% to 19%.

AI services and AI software spending, worldwide

US\$ billions; 2026 and 2027 are forecasts (Gartner, Sep 2026)



AI agents and assistants more than double in 2027 (+124%). AI services: a US\$1.2 trillion opportunity by 2030.

1 Assists

Buy: Tools or seats

Measured in: Usage, seats, tokens

Connects: Nothing; people move the work

Results in: Personal productivity

2 Does tasks

Buy: Agents for steps

Measured in: Tasks completed

Connects: One system or step at a time

Results in: Faster tasks, still joined by hand

3 Delivers work

Buy: Results

Measured in: Cost per outcome

Connects: Existing systems, end to end, with approvals

Results in: **Measured business outcomes**

WHERE STAGE 3 STARTS

With work the business already pays for. Work done in-house or outsourced already has a budget, an owner and a standard, so the outcome can be defined, bought and checked.

WHAT IT MEANS FOR LEADERS

Change the buying question from which tool to which end-to-end flow, to what standard, at what cost per outcome.

INVESTMENT IN AI-DELIVERED WORK

In 2026, venture and private-equity money began funding firms that deliver work with AI.

- 1 **DEMAND Buyers want results.**
Only 9% of B2B buyers see per-seat pricing as fair for agents, and paying per outcome rose from 11% to 19%. Gartner puts AI services at US\$576B in 2026, ahead of AI software.
- 2 **SUPPLY GAP Few can deliver them.**
About 17% of organisations have deployed agents, and only a quarter have moved even 40% of pilots into production. One AI lab says demand is outpacing any single delivery model; across Asia, talent is a top barrier.
- 3 **CAPITAL Investors fund delivery.**
They buy and rebuild services firms, back companies that build AI into clients' operations, and push AI across businesses they own. The table shows where.

Demand is moving to outcomes faster than businesses can deliver them, and capital is filling the gap.

DRILL-DOWN: SOUTHEAST ASIA

Demand is broad but early. Across seven Asian markets, 65% of firms have started on AI but only 15% are advanced (Thai SMEs: 71% and 17%). AI use per person is still about a tenth of US levels.

Capital hasn't followed demand yet. 92% of startup funding in H1 2026 went to Singapore-based companies, and the median Series A fell to US\$8M.

The foundations are arriving. Thailand approved US\$1.99B of AI-led investment in July, and IDC expects AI spending in Asia-Pacific outside China and Japan to reach US\$162B by 2030.

WHERE THE MONEY IS GOING, BY REGION

Region	AI share of venture funding	Moves into AI-delivered work, 2026
US	82.7% of deal value, Jan-Sep	Two AI labs launched services firms with private equity, one with US\$4B. Long Lake agreed US\$6.3B for Amex GBT. Thrive Holdings raised US\$2B to apply AI across 70+ firms it owns
Europe	75% in Q3, a record	General Catalyst backs founders buying and rebuilding service firms, such as UK lettings agencies and German SME accountants, typically with over US\$100M each
Asia	About 55% in Q3 (estimate)	SoftBank and OpenAI's Japan venture rolled out an AI-run security patching service, from finding flaws to deploying fixes, to 3,000 companies

WHAT IT MEANS FOR LEADERS AND INVESTORS

Capital is following delivered work in every major region, mostly into large firms in a few markets. Demand from small and mid-sized businesses is broad and still early. We think the next delivery firms will be built close to those businesses, running work in-country at a cost they can carry.

OUTCOMES BY BUSINESS FUNCTION

An outcome is work a business would otherwise give to a person or a firm, delivered end to end to an agreed standard.

Three business priorities. Most outcomes fall under priorities leaders already own: growing revenue, saving effort in operations, and protecting approvals, data and continuity.

What makes it finished work. A personal assistant helps one person with one step. Finished work moves a whole case across people, systems and approvals until the result is delivered, and it carries the six marks below the whole way.

Where value shows up. In McKinsey's 2026 survey, revenue gains from AI are reported most often in marketing and sales, and cost savings most often in supply chain management. Gartner finds customer service leaders' AI spending up 38% while their budgets grew 2%. BCG puts 70% of the effort of getting value from AI on people, organisation and process.

Priority	What the end-to-end flow covers	Measured by
Grow	Winning and keeping customers, from first contact through to the signed deal, delivery and renewal	Speed of response, conversion, revenue per customer
Save	Running the business, from a request or trigger through every team and system until the work is done	Cycle time, effort per case, rework
Protect	Keeping the business safe, from a request or risk through checks and approvals to a recorded decision	Issues caught early, policy followed, decisions on record

SIX MARKS OF FINISHED WORK

Owned Every step has an owner, and a named person is accountable	Tracked Status is visible to everyone working on it	Checked Work is verified before it's marked done	Approved Named people make the decisions that matter	Recorded A record shows what was done, by whom and why	Improved Each result feeds back, so the next run is better
--	---	--	--	--	--

WHAT IT MEANS FOR LEADERS

Treat people and AI as one team that shares the outcome. Who does which step will shift as the work proves itself, so setting that balance is part of running the work.

PLATFORM DEVELOPMENTS IN 2026

In 2026 the large platforms each launched AI that works in teams, mainly inside their own ecosystems.

Platform type	What arrived in 2026	What it does well	Where it stops by design
Personal and team AI agents	Persistent bots that keep working, learn routines and join team channels	Fast individual productivity; easy to start	Each person's conversations and memory stay private; vendor's cloud and models
Productivity-suite agents	Agents with their own accounts, process memory and triggers; long-running agents with sub-agent teams	Strong identity and governance inside the suite	Coordinate inside one vendor's ecosystem; much still in preview or early access
Shared AI workspaces	Pages that teams and AI edit together	Shared documents and context	Each member works in their own instance
CRM and line-of-business agents	Multi-agent orchestration with shared context	Deep in their own domain	Orchestration centred on its own platform and data
Automation platforms	AI agents added to workflow builders, with human checkpoints	Dependable for defined, repeatable steps	Judgement and context need design work
AI lab deployment services	New lab-backed services firms whose engineers build AI into a company's operations	Engineering depth and early access to frontier models	Built around their own lab's models; capacity still scaling to meet demand

WHAT END-TO-END DELIVERY STILL NEEDS

Across vendors
Real work crosses CRM, email, finance and messaging; each platform coordinates mainly inside its own.

Where it runs
Vendor cloud and models by default, while demand for in-country and sovereign hosting grows.

Its way of working
Built on the business's SOPs, policies, language and local rules, and kept current as they change.

Run after go-live
Toolkits leave the running to the buyer; Gartner expects 70% of firms to abandon vendor-built agents by 2028.

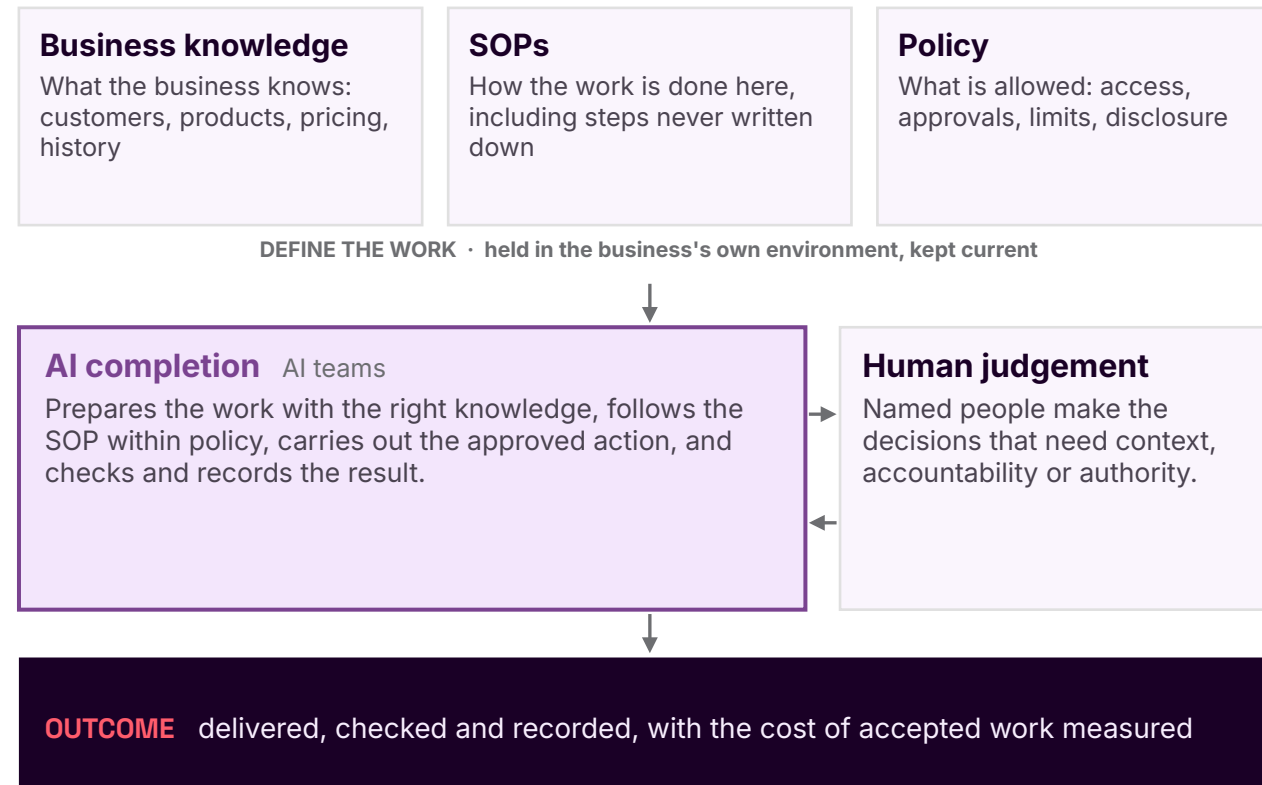
Paid by outcome
Priced by seat or usage, while buyers move towards paying per outcome.

OUR APPROACH
We run the work between platforms.
Purpose-built AI team units take on each job, workflow and the collaboration it needs, across the business's existing systems. They're built on its SOPs and policies, run where it chooses, managed after go-live and measured per accepted outcome.

THE COMPONENTS OF AN OUTCOME

Knowledge, SOPs and policy define the work, people make the calls, and AI teams do it and record it.

FIVE LAYERS OF AN OUTCOME



NINE SUPPORTING DISCIPLINES

DOING THE WORK

- 01 Work design:** steps agreed before anything runs
- 02 Knowledge and memory:** the right knowledge at every step
- 03 Connections and handoffs:** clean links, clear status

KEEPING IT SAFE

- 04 Data boundaries:** each role sees only what it needs
- 05 Control:** limits enforced by the system
- 06 Traceability:** who did what, with what, and why

KEEPING IT WORKING

- 07 Ownership and review:** named owners see the decisions
- 08 Change and learning:** tested, approved, recoverable
- 09 Proof and cost:** accepted outcomes, and their cost

Outside evidence: in 3,100+ test runs, agents failed sharply once a few sub-tasks were chained, and nearly three-quarters of failures came from how the work was run rather than the agent's design.

NINE SYSTEM DISCIPLINES

What 2026 research found, and how an autonomous AI system should be designed to address it.

Discipline and principle		What the research shows	How it should be designed to address it
DOING THE WORK	Work design Redesign the work before automating it.	Three-quarters of high performers redesigned their workflows, against a quarter of others (McKinsey); BCG puts 70% of effort on people and process.	Each workflow starts as agreed SOPs, exceptions and approvals, held in a business knowledge base before any AI runs it.
	Knowledge and memory Give every step the knowledge it needs.	Agents forget earlier instructions and lose context as jobs get longer; memory failures come to dominate (HORIZON study, 3,100+ test runs).	A business knowledge base holds what the business knows outside any one conversation, and the system checks what actually reaches each step, not only what was stored.
	Connections and handoffs Treat links as contracts.	Data and system readiness is the top AI barrier in seven Asian markets, at 47% (UOB); early errors carry through later steps (HORIZON).	A lead agent splits the job and tracks status; every link has a clear contract, and each handoff carries its evidence so work doesn't stall.
KEEPING IT SAFE	Data boundaries Give each role only what its work needs.	34% of office professionals at large firms have put customer data into public AI tools (Wakefield).	Each person and team has its own memory space in the business's environment; only reviewed summaries cross, and refused access is tested both ways.
	Control Let the system enforce limits.	Only 4.4% of written agent rules are backed by an enforcing control (Yan, arXiv); 5% have the full set of agent controls (BCG).	Limits and approvals are enforced by the platform, not left to instructions; work that can't pass its checks stops rather than lowering the bar.
	Traceability Record what was done, and why.	21% of firms planning agents have mature agent governance (Deloitte); by 2030 Gartner expects 80% of the Global 500 to make the CIO accountable for AI evidence.	Every action, input, decision and cost is logged and visible while work runs, so a failure can be traced to its real cause.
KEEPING IT WORKING	Ownership and review Name owners; show them the work.	Only 12% of Southeast Asian firms have an AI owner with budget authority, against 89% of AI leaders (Adobe, vendor survey).	Named people own each workflow; AI teams hand decisions to them with the work and its consequences in front of them.
	Change and learning Test repeats; approve changes.	A quarter have moved 40% of pilots into production (Deloitte); 55% in Asia-Pacific scaled back agent rollouts on cost (KPMG).	Procedures are protected: changes are proposed, tested for repeats and recovery, then approved, with earlier versions kept, because what works once can fail on repeat.
	Proof and cost Measure outcomes and their cost.	Only 5% have established a return (KPMG). Inference cost per agentic workflow is set to rise more than fivefold through 2028 (Gartner).	Results are measured against the business's own baseline, people judge what is accepted (they catch what automated checks miss), and cost is tracked per accepted outcome.

CONTROL, COST AND DATA RESIDENCY

Usage-based billing and in-country demand mean cost and control now matter as much as capability.

More agent use is billed by usage. In 2026, two of the largest providers put agent use on token- or usage-based billing, and a third announced a similar change before pausing it. Token prices keep falling, but agents use far more of them: Gartner expects inference costs per agentic workflow to rise more than fivefold through 2028.

Spend is under closer watch. In EY's mid-2026 survey of US senior leaders, 82% were concerned about token costs, though more were expanding their rollout (37%) than cutting it back (15%). By 2029, Gartner expects 60% of organisations deploying AI to set up a function that maps its total cost to value.

Where data sits now matters. Gartner expects sovereign cloud spending to grow 38% worldwide in 2027, and by 76% to 98% across Asia-Pacific, from a small base. In a September 2026 survey, 92% of Asia-Pacific IT leaders tied where their data sits to their AI plans.

Delivered locally. Telecom operators already run in-country data centres, and operator-led sovereign AI platforms are starting to appear, such as Indosat's AI factory in Indonesia.

FROM PRACTICE

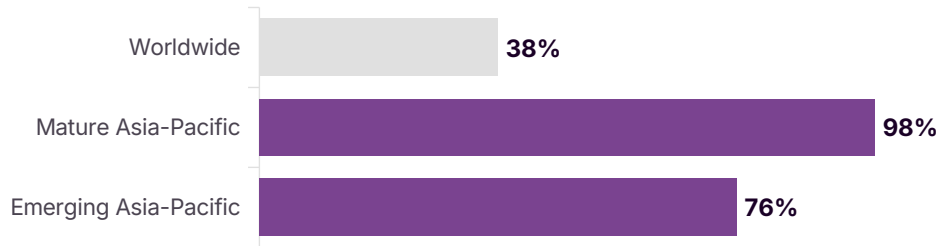
Keep data boundaries explicit. Holding, sharing and acting on information are separate permissions, and refused access needs checking as much as granted access.

WHAT IT MEANS FOR LEADERS

Decide early where the work and its data will run, and budget on cost per accepted outcome.

Sovereign cloud spending growth, 2027

% increase on 2026 (Gartner forecast)

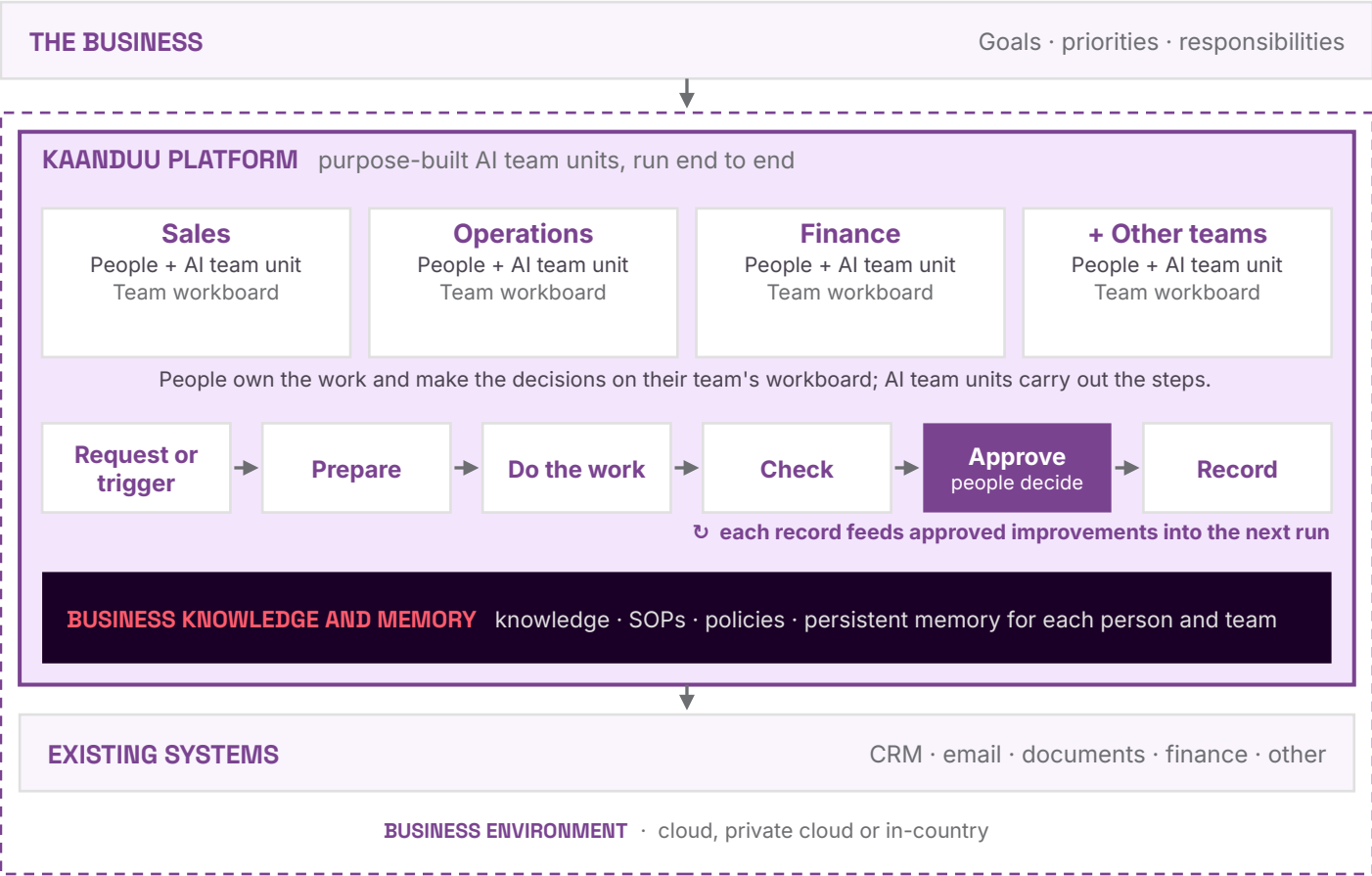


WHERE AI CAN RUN

Option	How it runs	Best for
All cloud	Interactive and background work on cloud AI services	Fastest start; least sensitive data
Hybrid	Routine and sensitive work on the business's own model; cloud AI for conversation	Data control with predictable cost
Fully local	All AI in the business's own or an in-country environment	Strict residency, regulated sectors

THE KAANDUU MANAGED MODEL

KAANDUU Managed builds purpose-built AI team units on a business's own knowledge, SOPs and policies, and runs them in its environment.



Teams own the work. Each team has its own workboard, where people and AI team units share tasks, owners and status, and decisions go to the right person.

Built around its processes. Once the business has agreed the steps, exceptions and approvals, the AI team units are designed to follow them instead of making their own calls.

Knowledge that persists. Business knowledge and each team's memory are kept outside any one conversation, so every step starts from what the business already knows.

Improvement with approval. Each result shows what to keep and change. People in the business approve improvements before they replace an agreed procedure.

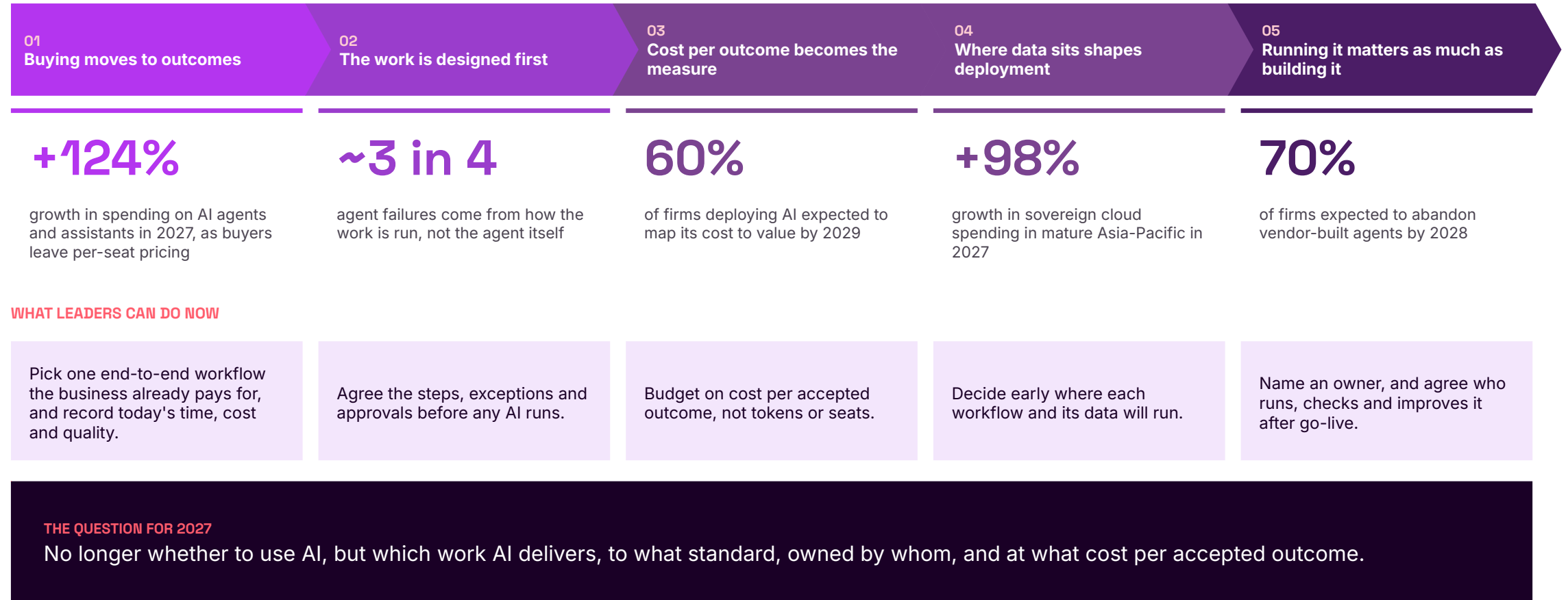
Implemented and managed with the business. We handle deployment, integrations and day-to-day operation. The business sets the priorities and decides who signs off what.

Control stays with the business. It all runs in the business's environment, and the platform is designed to enforce data boundaries and approval rules.

We measure results in outcomes, against the business's own baseline, so it can see what its spend returns.

LOOKING AHEAD TO 2027

The 2026 evidence points to five shifts in 2027, and each suggests a step leaders can take now.



TECHNICAL CHECKLIST

Any provider should be checked against the same list before anything goes live.

FOUNDATIONS

- ☐ **Environment** Runs in a dedicated environment the business chooses (its cloud account, private cloud or in-country) and meets local data-protection rules.
- ☐ **Identity and access** People sign in with single sign-on; each AI team unit has its own identity and only the access its work needs.
- ☐ **Models and path to local** Models set per workflow and swappable; moves from cloud to hybrid to fully local without rebuilding workflows.
- ☐ **Ownership and exit** The business owns its knowledge, SOPs, memory and records, and can export them if it changes provider.

DOING THE WORK

- ☐ **Work design** Steps, exceptions and approvals agreed before anything runs; changes proposed and approved, never silently overwritten.
- ☐ **Knowledge and memory** Knowledge kept outside any one conversation; shows what context actually reached each task, not only what was stored.
- ☐ **Connections and handoffs** Each system link has a clear input contract and useful errors; drafts, approvals and deliveries are distinct states.

KEEPING IT SAFE

- ☐ **Data boundaries** Memory separated by person and team; refused access checked in both directions; sharing decided separately from acting.
- ☐ **Control** Approvals and limits enforced by the platform, not by instructions to the AI; content in emails, pages or files cannot trigger actions on its own.
- ☐ **Security** Encryption in transit and at rest; credentials kept out of prompts and logs; local models secured like any server; regular testing for agent misuse.
- ☐ **Traceability** Inputs, decisions, actions and cost logged and visible while work runs; failures raise an alert to a named person.

KEEPING IT WORKING

- ☐ **Ownership and review** Every workflow and decision has a named owner; reviewers see the work and its evidence first; work the AI cannot finish goes to a person.
- ☐ **Change and learning** Repeats, retries and recovery after an interruption checked; approved improvements applied as the smallest compatible fix.
- ☐ **Proof and cost** Outcomes judged by the people who accept them; cost per accepted outcome measured, with spend limits per workflow.

Ask any provider for cost per accepted outcome, production error rates, recovery after an interruption, and proof that access is refused.

WORK, DELIVERED · THE STATE OF AI EXECUTION, 2026



kaanduu.com

KAANDUU LLC (USA) and KAANDUU Co., Ltd. (Thailand)