

The 2027 Enterprise AI Decision Guide

Executive Summary

Before deciding what to buy, decide what to ask

Enterprise AI projects don't fail on the model. They fail on governance—and governance cannot be added afterward.

17%

of organizations have actually deployed AI agents, against more than 60% planning to within two years

Gartner 2026 CIO Survey

40%

of enterprises are expected to downgrade or decommission autonomous agents by 2027, over governance gaps found only after an incident

Gartner, May 2026

2027/12

EU obligations for standalone high-risk AI systems take effect

Regulation (EU) 2026/1744

The error isn't execution. It's attribution.

Spending is surging, deployment is stalling—and the bottleneck isn't the model.

Global enterprises | Gartner 2026 CIO Survey—respondents are enterprise IT decision-makers

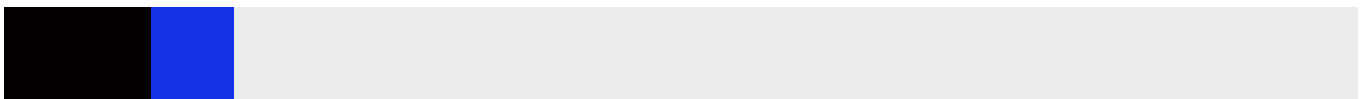
Plan to deploy AI agents within two years

>60%



Have actually deployed—■ segment marks the 11% production-ready

17%



Gartner 2026 CIO Survey; Deloitte, 2026

Local Corroboration

Among 230 of Taiwan's top 2000 enterprises, **74%** are actively planning or running AI projects but only **48%** are past departmental pilot or into production; a separate survey found **56%** still citing security, compliance and privacy as the main constraint on scaling AI.

Nutanix, commissioned to DIGITIMES, July–September 2026; Dell Modern Enterprise Readiness Study 2026. Both vendor-commissioned. The 48% here and the 17% above do not measure the same thing and should not be compared directly.

When your AI project is stuck in pilot and can't move into production, you're not alone—most enterprises stall at the same point.

Look back at cancelled AI projects and the same three reasons keep coming up: cost overruns, business value that can't be articulated, inadequate risk control. Weak model capability is not among them.

Governance gaps rarely surface during a pilot—they tend to appear only after go-live, once something has gone wrong. Gartner expects that by 2027, around 40% of enterprises will downgrade or shut off their autonomous agents because of security risks arising from weak governance. Fixing it at that point costs the most.

So the place to invest next isn't a stronger model. It's putting governance first.

Three forces you don't control

They answer, in turn: why act, why now, and when the cost of not acting comes due.

Inference economy

\$23.3B

2026 global inference spending – the first year above training's \$19B.

Inference is a steady-state workload, and steady state favours owning over renting. The share of enterprises using public cloud as their primary production inference environment fell from 56% to 41% in a year.

Memory supply

+125%

Server cost increase in the first half of 2026.

Around 70% of memory capacity is going to AI data centres, and quote windows have shortened to 1–30 days. The budget approved this year may buy 70% of the specification by the time you execute.

Regulatory clock

Dec 2027

EU obligations for standalone high-risk AI systems take effect.

What regulation asks for is evidence, not a policy statement—and evidence cannot be reconstructed afterward. 2027 is a year for preparation, not observation.

Jan 2026

Taiwan's Basic Act on AI takes effect

Dec 2027

EU high-risk AI rules start to apply

Aug 2028

EU rules extend to AI embedded in regulated products

Which AI systems will be asked to explain and keep records

Under Taiwan's AI Risk Classification Framework, an AI application is to be designated high-risk if it could cause serious harm to **national security, fundamental rights, life safety, property, social order or the environment**.

The test is the objective nature of the application itself.

The possibility of harm is enough; actual damage is not required.

Gartner, August 2026; Gartner and TrendForce, 2026; Broadcom Private Cloud Outlook 2026; Regulation (EU) 2026/1744. Taiwan's sectoral rules are due from each competent authority by January 2028.

The three forces run on different clocks, but they land in the same place: your 2027 AI plan and budget.

Inference runs every day, so AI moves from a one-off project cost to a standing operating cost. The memory shortage is structural, so delay means the same budget buys less. And regulation is the only one of the three with a fixed date, asking for the thing hardest to assemble at the last minute—evidence.

Three paths, one shared gap

None of them treats governance and audit evidence as a default deliverable.

Public cloud, self-building and integrated solutions each have real advantages—but they are missing the same thing: governance and audit evidence.

	Public cloud	Self-built	Integrated solution
Fast to start	●	○	◐
Data sovereignty	○	●	●
Predictable cost	○	◐	◐
Supply stability (2026–27)	●	○	◐
Governance built in	○	Build it yourself	Usually optional
Audit evidence deliverable	○	Build it yourself	Design it yourself

● Present / ◐ Partial / ○ Absent

Integrated solution: hardware and software pre-integrated by a single vendor, delivered into a third-party colocation facility and managed by you.

Each path has real advantages. Public cloud starts fastest and suits anything still being validated; self-building gives the most complete control; integrated solutions are mature, with stable supply chains.

But the last two rows are the point. They either do not provide governance or require the adopter to build it—and that is exactly the reason behind the agent downgrades cited earlier.

As noted earlier, Gartner expects around 40% of enterprises to downgrade or switch off autonomous agents over governance gaps found only after go-live. Those gaps most often sit in these two rows: **governance that isn't built in, and audit evidence you have to design yourself.**

That isn't any one vendor's failing. Most platforms were designed around getting AI running, and AI governance is often left for the adopter to add. But after go-live, you're the one who has to answer to the board, regulators and customers.

So when you compare options, weigh AI governance alongside performance and cost. The eight questions on the next page are where that starts.

So the questions you ask have to change

When AI moves from answering questions to acting on our behalf, specifications are no longer the whole question.

The three paths on the previous page are missing the same layer.

If that's the case, the questions we put to vendors have to change too.

What you used to ask

How large a model can this machine run?



What's the cost per token?



Which models are supported?



Does it have RAG?



Where is the data stored?



Is there access management?



What can this agent do?



When can it be delivered?



What you should ask in 2027

Six months from now, can you still reconstruct the full basis for a specific AI output?

Can you see which department and which agent spent what—and set a ceiling?

**Does switching models require downtime?
Will you be locked into a single supplier?**

Do answers carry sources? After a document is updated, will it still answer from the old version?

**Which data never leaves your network—
enforced by the system, or just an optional setting?**

Can agents at different risk levels run under rules of different strength?

**What can this agent not do—
and who enforces that, and how?**

How long is the quote valid? What is the price adjustment mechanism?

All eight questions on the right ask whether operational performance can be proven, not just whether a feature exists. None of the answers are on a spec sheet; they're in how the AI is managed. And that is exactly where many enterprises fall into the same trap.

Not every AI agent should run under the same rules

Share one foundation, but set the rules by risk level.

The most common mistake is to run every agent under the same rules. Set them too loosely and high-risk agents slip out of control; set them too tightly and even low-risk uses stall.

Taiwan's AI Risk Classification Framework names agents exceeding their granted authority as risk type B6, "AI agents acting beyond authorisation": an agent may drift from its original instructions as goals prove incomplete or conditions change, and can acquire system access beyond what it was granted; in multi-agent systems, agents triggering one another can produce chain reactions neither the developer nor the user anticipated.

To avoid both outcomes, the workable approach is to separate what is shared from what is kept apart.

Upper layer: shared by all agents

Control plane (unified)

One identity system · one audit event format · one usage and cost ledger · one policy enforcement point

Lower layer: managed by risk tier

L1 Read-only

Retrieval, Q&A, summarisation

Scoped data access
Outputs carry sources

L2 Write

Writes to internal systems,
produces formal documents

Named approver
Preview before write,
reconstructable

L3 External

External communication,
triggering transactions, cross-
organisation delegation

Per-item approval
Spending caps, two-person
review

Unify the foundation and you can see everything; tier the rules and you can contain the risk.

None of it happens overnight—but the good news is that none of it has to wait until procurement is complete.

Four things you can start now, without buying anything

Whichever path you eventually take, you can start with these four reviews as the foundation for your 2027 AI plan.

The risk categories are based on Taiwan's AI Risk Classification Framework (Ministry of Digital Affairs, July 2026), issued under Article 16 of the Basic Act on Artificial Intelligence—in particular A2 (lack of transparency or explainability), A5 (privacy and personal data protection) and B6 (AI agents acting beyond authorisation).



Inventory every AI system in use

What AI is actually in use across the company, including tools departments bought themselves?

Phase 1

Output: an AI system register



Classify data

Which data can AI use?
Is the boundary held by the system, or by people's good judgment?

Phase 2

Output: classification standard and enforcement points



Review agent permissions

Are agent limits set in access control, or only written into the prompt?

Phase 3

Output: agent risk tiering table



Assess audit trail gaps

If you were asked to explain tomorrow, could you produce the records?

Phase 4

Output: gap list with remediation priority

THREE TEST QUESTIONS FOR THE AUDIT TRAIL

1. For a specific output three months ago, which model produced it and on what data?
2. Over the past month, what was each department's AI usage and cost?
3. Which data did a given employee access through AI on a given date?

Whatever you can't answer yet is what to fix first.

None of this requires a procurement decision first—and once it is done, the quality of any procurement evaluation improves.

If you remember only three things

- 01** What gets AI into production is rarely the model. It's whether you can govern it.
- 02** Regulators will ask for evidence, and evidence can only be built from today.
- 03** Before deciding what to buy, be clear on what you want.

Start with the four reviews on page 7, then choose the resource below that fits where you are in your evaluation.

STILL EVALUATING

**See the three approaches
and the full context**



aisoportal.com

WANT TO SEE THE OFFERING

**AISO appliance:
options by industry scenario**



aisoportal.com/bundled-solution

AISO stands for **AI for Sovereignty**.

Agent technology is still evolving quickly, and the methods for governing it are still taking shape. What counts as best practice today is not a fixed standard. Your governance approach should evolve with your organization's business needs and competitive priorities.

Technology will keep changing. Whether we can be accountable for it is the question that stays.

service@aisoportal.com