Where Does Accountability Live When the AI You Bought Doesn't Behave?

Yasin Jabbar · yasinjabbar.com · September 2026

Primary sources (all public, linked on the last page): Gartner press release (25 June 2025) · IBM Institute for Business Value C-suite study (8 June 2026) · IBM / Ponemon *Cost of a Data Breach Report* 2026 and 2025 · Checkmarx, *The Future of Application Security in the Era of AI: 2027 Industry Outlook* (May 2026) · LexisNexis 2026 Future of Work Report as reported by CIO.com (June 2026).

Nothing in this chapter draws on non-public information from any prior employer. General industry practice only.

The Decision

The board wants AI. Legal wants control. Employees want speed. The CIO sits in the middle of all three with a budget the CFO is about to question.

The decision in front of every mid-market CIO right now is **not** whether to do AI. That ship sailed. The decision is **who owns it when it breaks** — because the numbers say it is going to break, and the numbers say most of it is going to get cancelled.

Gartner put a date on it. In June 2025 they predicted that **more than 40% of agentic AI projects will be cancelled by the end of 2027**, for three reasons: escalating costs, unclear business value, and inadequate risk controls.

Read those three reasons again. They are not three problems. They are one problem seen from three chairs — the CFO's, the board's, and the CISO's. The name of that one problem is *accountability*. Nobody wrote down who owns the outcome, so the cost has no owner, the value has no owner, and the risk has no owner. Projects without an owner get cancelled.

This chapter is about how to be the CIO whose projects survive 2027.

The Pain — In Their Own Numbers

IBM's Institute for Business Value surveyed 2,000 senior technology executives across 33 countries between January and April 2026. The results are not subtle:

- **Two-thirds of CIOs and CTOs** report being held accountable for AI systems they do not fully control.
- **70%** say business teams deploy technology faster than IT can monitor it.
- **77%** say AI adoption has already outrun their governance capabilities.
- Surveyed leaders expect the number of deployed AI agents to grow **38% by 2027** — and only **11%** feel completely ready for that scale.
- The average organization logged **54 AI-agent incidents** in the past year.

- **17%** of those were high severity, meaning more than four hours to contain.
- Of the high-severity incidents: **37%** were data exposures or security breaches, **33%** were cascading system failures, **17%** triggered compliance issues.

Do the arithmetic on the averages: 54 incidents at 17% high-severity is **about nine high-severity AI incidents a year** in a typical shop. Most CIOs I talk to could not name three of theirs.

Checkmarx surveyed 2,350 CISOs, AppSec managers and developers across 14 countries for its 2027 Industry Outlook (May 2026). Their finding is worse, because it is about intent:

- **75% of organizations admit they often or sometimes deploy code they already know is vulnerable.**
- **About 30%** admit they ship compromised code and hope the vulnerability won't be found.
- **93%** had at least one breach traced directly to an in-house developed application.
- Teams whose codebases are 81–100% AI-generated ship vulnerable code **3.4x more often** than teams at 20% or less.

The most uncomfortable line in the whole report: organizations that self-identify as **"highly mature" AI shops ship vulnerable code at 42%** — and their breach rates are barely distinguishable from less mature peers. Maturity didn't save them. Process didn't save them. The bottleneck isn't detection. It's the human decision to ship anyway.

This is the world the CIO is being asked to take accountability for.

The False Choice — Lockdown vs. Innovation

Most accountability frameworks pretend the decision is between **locking it all down** (kill shadow AI, force every workload through a governance board) and **embracing the chaos** (trust your people, let velocity win, hope nothing catches fire).

Both fail at scale.

Lockdown kills the speed advantage and breeds the very shadow AI it tries to prevent. The clearest signal in the current data comes from LexisNexis's 2026 Future of Work research, as reported by CIO.com: **74% of AI-trained employees use unauthorized AI tools, versus 17% of untrained employees.** The more training and policy you wrap around AI, the more your most capable people route around it. Mandatory training without a sanctioned path is shadow-AI accelerant. Governance teams didn't see this coming because they assumed humans optimize for compliance. Humans optimize for getting the work done.

Embracing the chaos has a price tag, and IBM's *Cost of a Data Breach* series has been tracking it:

- The 2025 edition found shadow AI added **$670,000** to the average breach cost, and **97%** of organizations breached through their AI lacked proper AI access controls.
- The 2026 edition (602 breached organizations, March 2025–February 2026) found the share of security incidents involving shadow AI **more than doubled to 43%**, **92%** of organizations attacked through their own AI models had failed to control access to them, and **more than two-thirds** had no governance process in place to limit shadow AI. The global average breach now runs roughly **$5 million**.

Twelve months ago shadow AI was a footnote in the breach data. Now it is in almost half the incidents. The chaos is compounding.

The real decision isn't a slider between these two extremes. It's a three-layer architecture of accountability.

The Real Decision — Three Layers of Accountability

Layer	The Question	Who Owns It
Visibility	Do we know what AI is running inside our walls?	CIO + Platform team
Containment	When something breaks, how big does the blast radius get?	Security + SRE
Decision rights	Who gets to say "ship anyway"?	Defined by workload type, NOT job title

Answer all three, write them down, and your accountability surface shrinks even as your AI sprawl grows. Miss any one and you're the CIO in the IBM survey with 54 incidents and no name on any of them.

Layer 1 — Visibility

You cannot be accountable for what you cannot see. The minimum viable product here is not a vendor's magic quadrant. It is a **weekly inventory of AI dependencies** that lists, for every customer-facing or revenue-affecting workload:

- Which model or provider is in the path
- Who in the organization owns that relationship
- What happens if the provider has a four-hour outage
- Which workloads have a manual fallback and which do not

If you can produce that list this Friday, you are already ahead of the 77% who say adoption has outrun their governance. The rest are the ones with nine unaccounted high-severity incidents a year.

Layer 2 — Containment

Plan for absence, not degradation. When a model API goes dark, an AI-dependent workflow doesn't slow down — it stops. Most model providers cannot offer the four-nines commitment a hyperscaler does. That is a contract reality, not a vendor failing, and it should be underwritten like one.

Containment looks like:

- **Quarterly AI failure drills** — kill the model API in staging and see what stops working.
- **Workforce cross-training** — at least two humans per AI-replaced workflow must be able to do the work manually within 24 hours.
- **Vendor SLA scrutiny** — a 99.5% SLA is roughly 44 hours of permitted downtime a year. Underwrite the workload accordingly, or move it to a provider that commits to 99.99%.

This is operational-continuity work, not magic. Mid-market IT teams already know how to run it — they did it for SaaS a decade ago and for cloud after that. They just haven't applied it to AI yet.

Layer 3 — Decision Rights

The Checkmarx finding — the bottleneck is the human decision to ship anyway — is the only finding here that requires a policy change rather than a tools change.

Pre-define, in writing, **who can override a security finding for which workload type.** A working matrix:

Workload type	Who can ship despite a vulnerability finding
Internal-only / low blast radius	Engineering manager
Customer-facing / non-regulated	VP Engineering + CISO co-sign
Customer-facing / regulated	Board sub-committee escalation only
Anything touching PHI/PII at scale	No override — fix or kill

One page, signed by the CEO. It removes most of the "ship anyway" pressure, because the engineering manager can now point at the policy instead of absorbing the personal hit. It also turns the Checkmarx 30% — ship and hope — from a cultural problem into a defined exception with a name attached.

Common Wrong Turns

- **"We'll require AI training, then nothing slips past governance."** The data says the opposite. Trained employees use unauthorized tools at more than four times the rate of untrained ones. Training without a sanctioned path and clear decision rights is gasoline.
- **"We'll buy a shadow-AI detection tool."** Detection is largely solved. The bottleneck is what you do with the detection — and most organizations do nothing. Buy decision rights, not more dashboards.
- **"We'll wait for regulation to clarify."** Your agent count grows 38% by 2027 and Gartner's cancellation deadline is the same year. Acting first is cheaper than acting after the incident that makes the news.
- **"We'll make the security team the gatekeeper."** Security reviewing every AI deployment is exactly the bottleneck pattern that produced the 30% ship-and-hope number. Distributed decision rights with a written escalation matrix is faster *and* safer than a single gate.

What This Costs to Run

Be honest about the price on both sides.

The play — a weekly inventory, a quarterly failure drill, and a one-page decision-rights matrix — is roughly **half of one senior platform engineer's time** for a mid-market organization. It is not a program. It is a habit with three artifacts.

On the other side of the ledger sit the sourced numbers above: an average of 54 agent incidents a year, roughly nine of them high-severity, a global average breach cost near $5 million, shadow AI present in 43% of incidents, and a Gartner-dated cancellation risk on more than 40% of the agentic projects your board just approved.

You don't need an invented multiplier to make the case. The publicly available numbers make it on their own. Put them in next year's risk register with a name next to each one.

Recommendation by Profile

- **Regulated mid-market (healthcare, financial services, credit unions and community banks, gov-adjacent):** All three layers, in writing, signed at board level, audited annually. For you, accountability is a *regulator* problem before it is a CFO problem. Non-negotiable.
- **Growth-stage SaaS:** Visibility layer in 30 days, decision rights in 60, failure drills in 90. Do not wait for the breach to discover what wasn't inventoried.
- **Cost-sensitive operations:** Layer 1 alone — the visibility inventory — surfaces most of the shadow-AI exposure at almost no cost. Start there. Layers 2 and 3 in year two.

The Operator's Bottom Line

> *The board hired you to make AI work. They will fire you for the breach — or cancel the project before it gets the chance.*

The accountability gap is not closed by buying more tools, hiring more security, or training more employees. It is closed by writing down, in three layers, who owns what — and then drilling the failure modes before they go live.

Visibility. Containment. Decision rights. Run the play.

Where to Start — The AI Value Diagnostic

If you want the three layers mapped against your actual workloads rather than a generic matrix, that is exactly what the **AI Value Diagnostic** does.

- **Fixed scope, fixed price: $3,500.**
- **Two weeks**, start to readout.
- You get an inventory of the AI in your walls, a containment assessment of the workloads that matter, and a decision-rights matrix drafted for your organization — a document your board can sign.
- Most AI engagements in this market start at $25,000. Find out what you actually need first.
- You'll talk to me. There is no one else, and no handoff to a delivery team after you sign.

Book it or read the full scope at yasinjabbar.com/diagnostic.html.

Sources

1. Gartner, "Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027," press release, 25 June 2025. https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027
2. IBM Institute for Business Value, "New IBM Study Finds CIOs and CTOs Face Growing AI Control Gap as Enterprise Deployment Scales," 8 June 2026 (2,000 executives, 33 countries, 19 industries). https://newsroom.ibm.com/2026-06-08-new-ibm-study-finds-cios-and-ctos-face-growing-ai-control-gap-as-enterprise-deployment-scales
3. IBM / Ponemon Institute, *Cost of a Data Breach Report 2026*, 29 July 2026. https://newsroom.ibm.com/2026-07-29-ibm-study-one-in-four-malicious-breaches-are-ai-enabled,-costing-companies-6-million-on-average — shadow-AI share (43%) and access-control (92%) figures as reported by Cybersecurity Dive: https://www.cybersecuritydive.com/news/data-breach-costs-ai-governance-ibm/826463/
4. IBM, *Cost of a Data Breach Report 2025*, 30 July 2025 ($670K shadow-AI premium; 97% lacking AI access controls). https://newsroom.ibm.com/2025-07-30-ibm-report-13-of-organizations-reported-breaches-of-ai-models-or-applications,-97-of-which-reported-lacking-proper-ai-access-controls
5. Checkmarx, *The Future of Application Security in the Era of AI: 2027 Industry Outlook*, May 2026 (2,350 respondents, 14 countries). Press release: https://checkmarx.com/press-releases/75-of-companies-ship-vulnerable-code-despite-a-startling-increase-in-threat-velocity-agentic-appsec-unleashed-26-is-changing-that/ — detailed findings as reported by CIO.com: https://www.cio.com/article/4183209/enterprises-know-ai-generated-code-is-vulnerable-theyre-shipping-it-anyway.html
6. LexisNexis, 2026 Future of Work Report, as reported in CIO.com, "Why your most AI-savvy employees are driving shadow AI," June 2026. https://www.cio.com/article/4178359/why-your-most-ai-savvy-employees-are-driving-shadow-ai.html

Next in the series — Decision 4: "AI Resilience — Planning for Absence, Not Degradation." Why your business-continuity plan probably doesn't survive a four-hour model-provider outage, and the 90-day cure.