
SPANFORGE | WHITE PAPER

The Enterprise AI Delivery Crisis — and the System That Fixes It

“Enterprises don’t have an AI problem. They have a delivery discipline problem.”

42% of enterprise AI initiatives were abandoned in 2025 — up from 17% in 2024.

This paper examines why the pilot trap is accelerating, what it truly costs, and how the SpanForge Exit Gate System™ provides a structured, evidential path to production.

April 2026 • www.getspanforge.com • Sources: S&P Global Market Intelligence (2025); Gartner (2024, 2025); McKinsey & Company (2025)

EXECUTIVE SUMMARY

The AI Pilot Crisis Is No Longer a Prediction

“AI is not failing in the lab — it is failing at the handoff to reality.”

Enterprise AI adoption faces a stark paradox: investment is accelerating while delivery is collapsing. According to S&P Global research spanning more than 1,000 enterprises, approximately 42% of companies reported abandoning the majority of their AI initiatives in 2025 — more than double the 17% recorded the year prior. The scale and velocity of this shift represent one of the sharpest reversals seen in enterprise technology adoption in recent years.

The culprit is not AI itself. Abandonment at this rate is a delivery discipline problem: the absence of a structured lifecycle governing how pilot projects mature into production systems. Without defined gates, clear ownership, and measurable transition criteria, pilots drift indefinitely — consuming budget and producing nothing of lasting value.

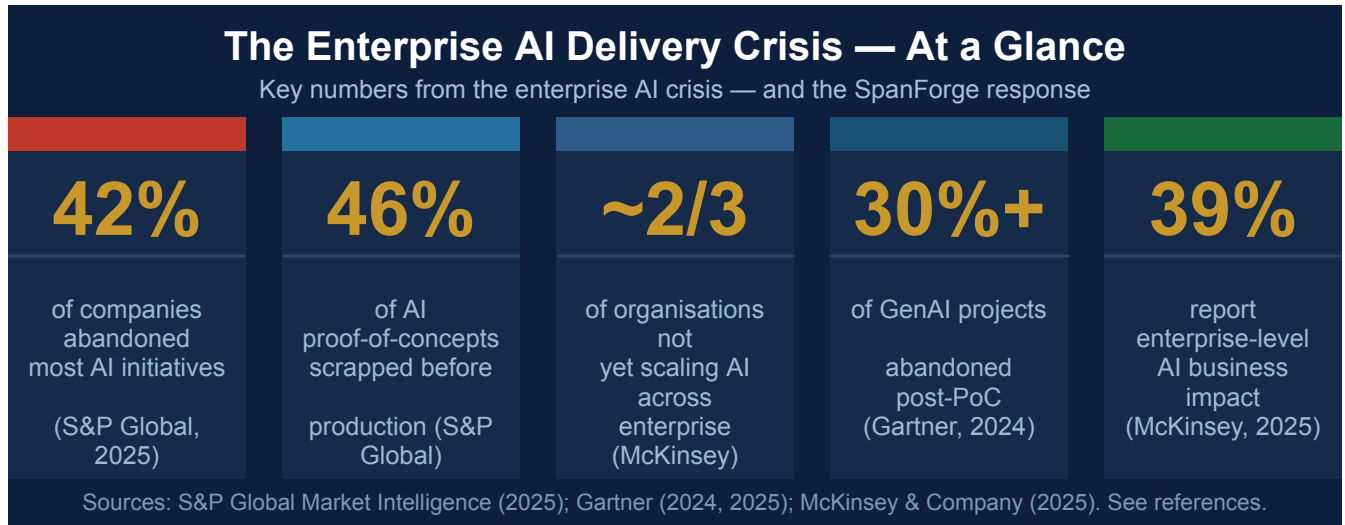
This paper makes three arguments: the ‘pilot purgatory’ problem is structural and worsening; unstructured pilots carry hidden costs most organisations dramatically underestimate; and the SpanForge Exit Gate System™ provides a repeatable, evidence-based mechanism for breaking the cycle.

KEY FINDING

Organisations that implement structured AI lifecycle management — including defined exit gates, clear ownership, and measurable transition criteria — address the root causes of pilot abandonment identified in the S&P Global and Gartner research cited in this paper. The framework in this paper is designed to close the specific governance gaps those sources identify as primary drivers of failure.

AT A GLANCE

The Enterprise AI Delivery Crisis — Five Numbers That Define It

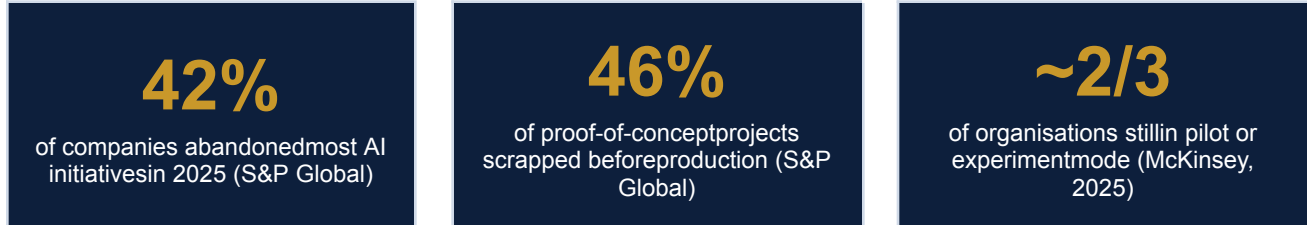


These numbers define the scale of the crisis and the gaps the SpanForge Exit Gate System™ is designed to close. The remainder of this paper examines the evidence behind each and explains the mechanism for addressing them.

SECTION 1

The Evidence: AI Abandonment Has Reached Crisis Proportions

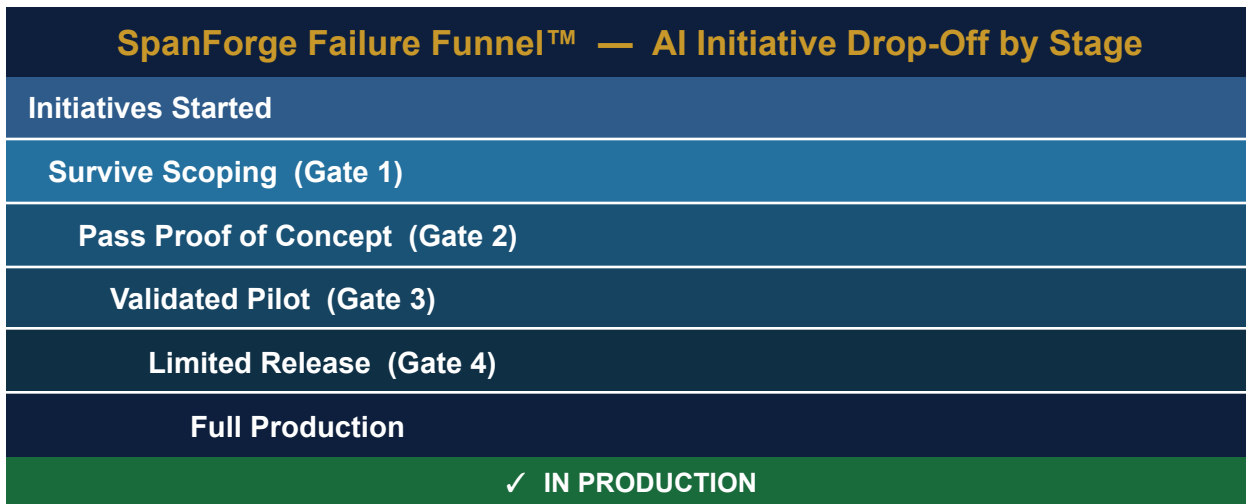
“The failure rate doubled in a single year. That is not a trend — it is a structural collapse.”



Sources: S&P Global Market Intelligence (2025), survey of 1,006 enterprises; McKinsey State of AI (2025), survey of 1,993 organisations. See references.

Figure 1: SpanForge Failure Funnel™

The S&P Global data tells us initiatives are being abandoned at high rates. The SpanForge Failure Funnel™ illustrates where in the lifecycle those losses typically occur — most long before technical limitations become the issue.



“The further from the start, the fewer survive. Most are lost before technical limitations are ever tested.”

Figure 1: SpanForge Failure Funnel™. Illustrative representation of where enterprise AI initiatives are typically lost across the lifecycle. Based on publicly reported adoption data from S&P Global Market Intelligence (2025). Specific drop-off rates at each stage are illustrative.

What the Data Actually Shows

The S&P Global Market Intelligence survey of 1,006 enterprises covers sophisticated organisations with dedicated AI budgets, data science teams, and executive mandates. Approximately 42% reported abandoning the majority of their AI initiatives in 2025 — up from 17% in 2024. On average, organisations scrapped 46% of their AI proof-of-concepts before reaching production. The cost, data privacy, and security risks were cited as top obstacles. Each failed pilot erodes appetite for the next, reduces tolerance for ambiguity, and increases political resistance to AI investment.

The McKinsey State of AI 2025 survey of 1,993 organisations across 105 countries reinforces this picture from a different angle: nearly two-thirds of organisations have not yet begun scaling AI across the enterprise, and only 39% report any EBIT impact at the enterprise level. Widespread adoption has not translated into widespread delivery.

Initiatives in **unstructured exploration phases** — proof-of-concept work without defined success criteria — account for a disproportionate share of casualties. This pattern is consistent with the governance gaps identified across the sources cited in this paper.

Why Now? The Four Accelerants

- **Capability inflation.** Model releases have been so rapid that organisations chronically restart pilots to incorporate the ‘latest’ version, resetting the maturity clock each time.
- **Governance vacuum.** Most enterprises bolted AI onto governance frameworks designed for deterministic software. Gartner (2025) found that 63% of organisations do not have or are unsure they have the right data management and governance practices to support AI — creating unresolvable approval loops and decision gaps.
- **Talent dilution.** As demand for AI practitioners surged, methodology rigour has not kept pace with tool proliferation.
- **Expectation misalignment.** Executive sponsors hold timelines disconnected from enterprise integration realities. When reality diverges, projects are cancelled rather than recalibrated.

INDUSTRY SIGNAL

Gartner predicted (July 2024) that at least 30% of generative AI projects would be abandoned after proof of concept by end of 2025, due to poor data quality, inadequate risk controls, escalating costs, or unclear business value. A separate Gartner study (February 2025) found that 63% of organisations either do not have or are unsure they have the right data management practices to support AI. Without the right data and governance foundations, projects cannot advance beyond the pilot stage.

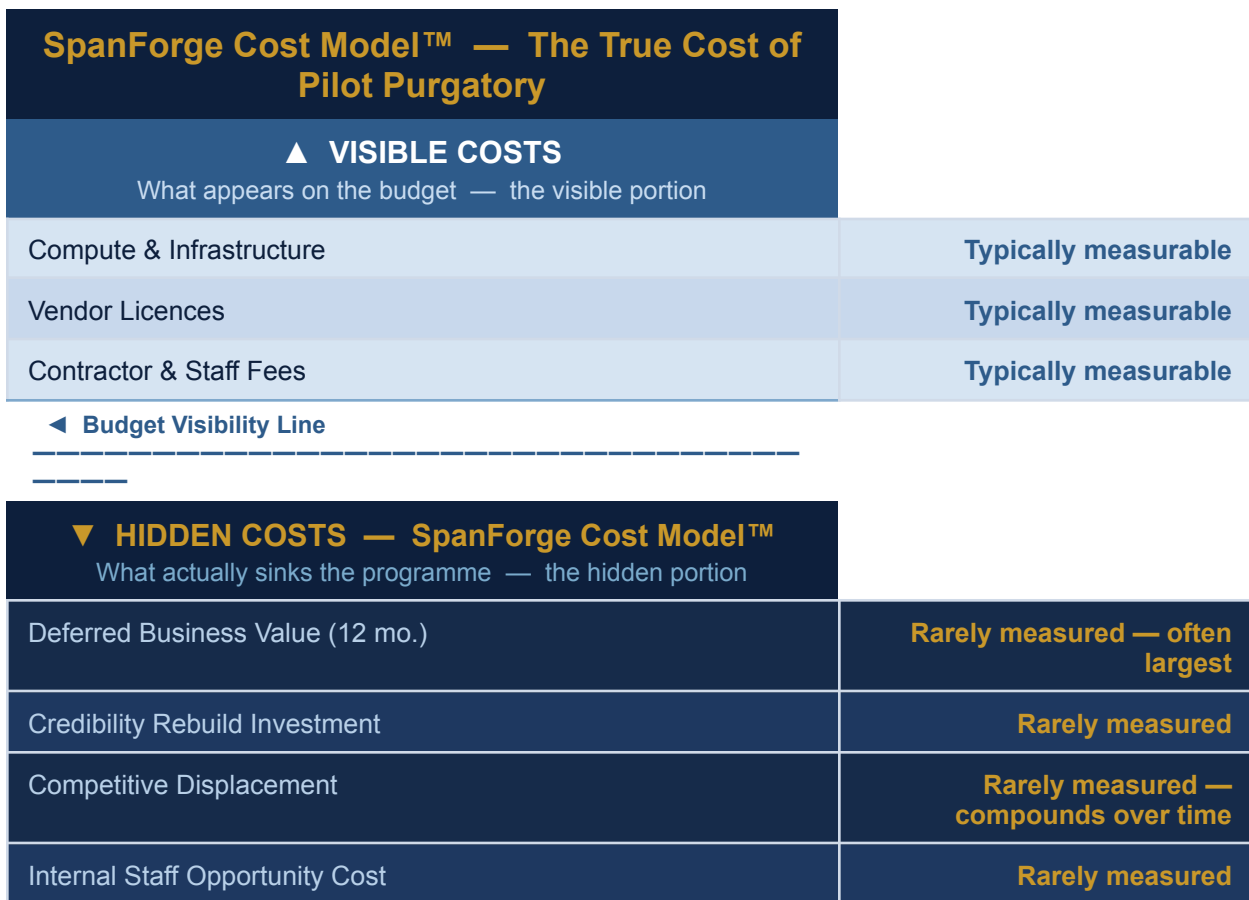
SECTION 2

The Hidden Cost of AI That Never Ships

“Most AI programme cost is invisible on financial statements. That invisibility is why it keeps compounding.”

Organisations typically account for only the direct costs of abandoned pilots: compute, vendor fees, and contractor hours. These visible costs represent only a fraction of the true impact. The larger burden sits below the surface, rarely appearing on a budget report.

Figure 2: SpanForge Cost Model™



“The majority of AI programme cost is invisible — and therefore unmanaged.”

Figure 2: SpanForge Cost Model™. Cost categories and ranges are illustrative, drawn from publicly available industry benchmarks. Estimates; individual results will vary.

1. Direct Financial Waste

The direct costs of an abandoned AI pilot — compute, vendor licences, contractor fees, and internal staff time — are rarely trivial. For mid-market organisations these costs are measured in hundreds of thousands of dollars; for enterprise-scale programmes, in millions. Across a portfolio of several concurrent pilots, the total is significant. But direct costs are only part of the picture.

2. Opportunity Cost of Deferred Value

Every month a pilot spends in purgatory is a month the business value it was designed to capture is not being realised. A 12-month delay on a customer-facing initiative targeting a 15% reduction in service handling time compounds losses far beyond the direct programme cost.

3. Organisational Credibility Erosion

When AI pilots fail visibly and repeatedly, investment appetite contracts. AI becomes politically radioactive — associated with broken promises rather than competitive advantage. Rebuilding sponsor confidence after a series of high-profile failures takes time and organisational effort that is rarely budgeted for. The cost is real, even if it does not appear on a balance sheet.

4. Competitive Displacement

Competitors who have solved the delivery problem are compounding production-system value while others cycle through abandoned pilots. Closing the resulting gap becomes more expensive with each passing cycle.

CONTRARIAN TAKE

Sunk cost is not the primary risk of a failed pilot. The primary risk is the credibility damage that makes the next initiative harder to fund, staff, and deliver.

SECTION 3

Why Most Enterprise AI Approaches Fail (Even With Good Teams)

“The evidence points to a consistent conclusion: governance gaps — not capability gaps — are the primary driver of enterprise AI pilot failure.”

The enterprise response to pilot failure has typically followed one of three patterns, each addressing the symptom rather than the cause.

The “More Rigour” Response

The most common reaction is to add process overhead: more governance reviews, additional sign-off layers, expanded risk assessments. This conflates rigour with bureaucracy. It adds more checkpoints on an undefined journey without clarifying what ‘done’ looks like. The result: a longer, more expensive version of the same unresolved trajectory.

CONTRARIAN TAKE

More governance does not reduce risk — it often obscures it. Adding approval stages without exit criteria is not governance; it is delay with paperwork.

The “Fast Fail” Response

Some organisations time-box aggressively: kill anything without results in 90 days. In enterprise contexts — with complex integration requirements, regulatory constraints, and multi-stakeholder approval chains — 90 days is often insufficient to reach meaningful production-viability assessment.

CONTRARIAN TAKE

Fast fail is a startup principle systematically misapplied to enterprise systems. Premature termination is not learning discipline — it is pattern-matching to the wrong model.

The “Build More Infrastructure” Response

Teams with mature MLOps stacks still abandon pilots at comparable rates when the core governance and lifecycle problem remains unsolved. Tooling is rarely the constraint.

Approach Comparison: Why Alternatives Fall Short

Approach	Pattern	Core Failure Mode	Typical Outcome
Ad hoc piloting	Unstructured iteration	No exit criteria; endless drift	High abandonment; budget exhaustion
Fast fail	Aggressive time-boxing	Insufficient time for enterprise validation	Premature termination; learning lost
Tool-first	MLOps / platform investment	Tooling without lifecycle governance	Infrastructure without delivery
SpanForge Exit Gate System™	Structured evidential lifecycle	Designed to prevent all three failure modes	Structured, documented delivery to production

SpanForge analysis. Failure mode taxonomy derived from patterns identified across S&P Global Market Intelligence (2025), Gartner (2024, 2025), and McKinsey & Company (2025). See references.

ROOT CAUSE

The fundamental failure is the absence of a shared, explicit contract between the AI team and the business about what conditions must be satisfied — and verified — before a pilot advances to the next stage. Without this contract, advancement decisions are political rather than evidential.

Having diagnosed the failure modes, the question becomes: what does a system that actually solves them look like?

SECTION 4

The SpanForge Exit Gate System™: A Structured Lifecycle

“Every advancement decision should be made on evidence. Every closure decision should be documented. Neither should ever be made by momentum alone.”

The SpanForge Exit Gate System™ addresses the root cause of pilot purgatory with a formal, evidence-based lifecycle: defined stages, explicit transition criteria, and non-negotiable gate conditions. It replaces ambiguity with clarity at every step.

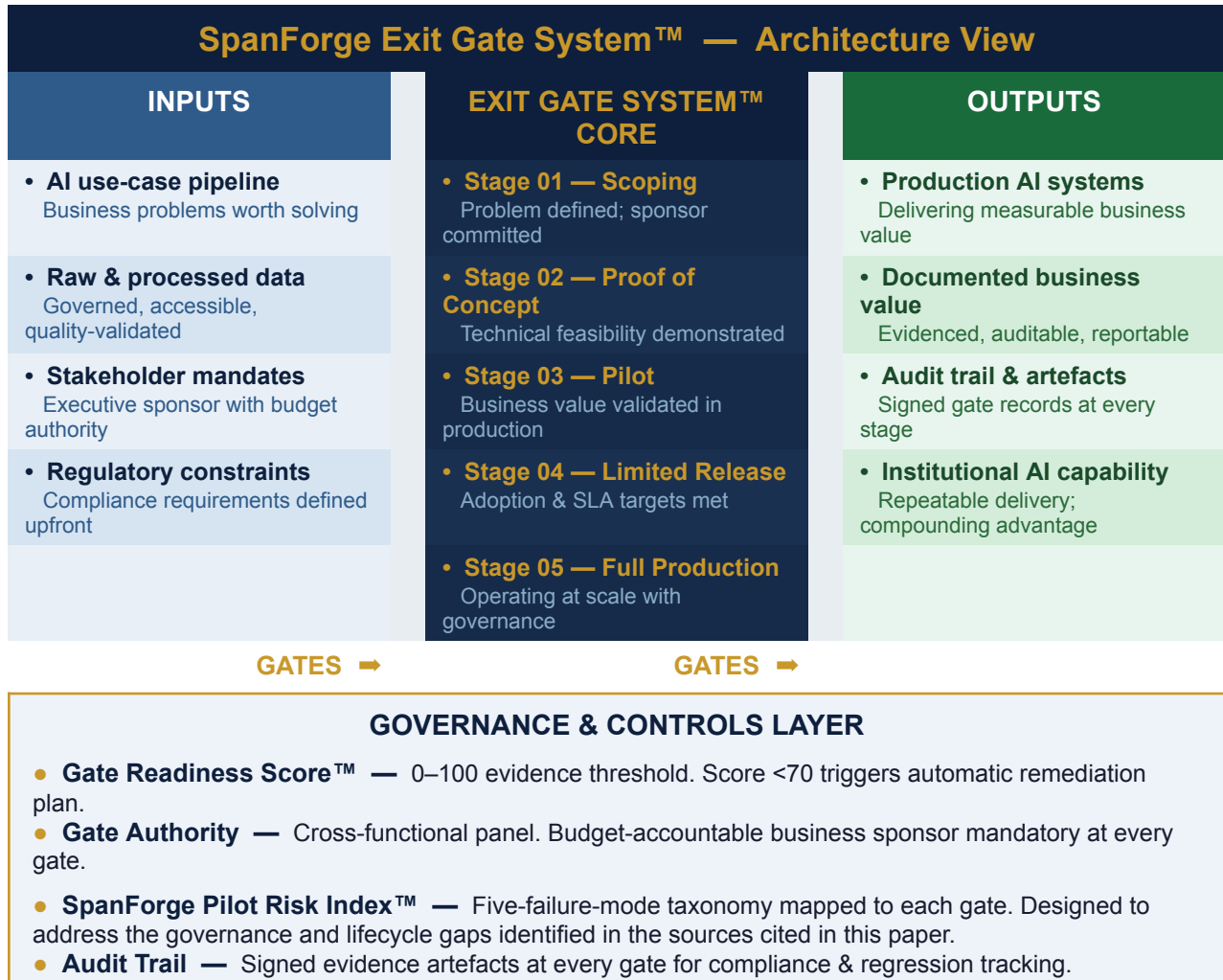
Figure 3: SpanForge Exit Gate System™ — Five-Stage Lifecycle



Each gate requires documented evidence before advancement. Returning an initiative is a designed outcome, not a failure.

Figure 4: SpanForge Exit Gate System™ — Architecture View

The architecture view shows how inputs, the core lifecycle engine, governance controls, and production outputs relate. This is the CTO-level view of how the system operates across an enterprise AI portfolio.



The Five Lifecycle Stages

Stage	Name	Objective	Exit Gate Condition
01	Scoping	Define problem, value hypothesis, and data prerequisites	<i>Signed problem statement; confirmed data access; sponsor commitment letter</i>
02	Proof of Concept	Validate technical feasibility with representative data	<i>Baseline accuracy met on holdout set; risk register reviewed; PoC report approved</i>
03	Pilot	Validate business value in a controlled production environment	<i>KPI targets achieved; operational integration confirmed; compliance sign-off obtained</i>
04	Limited Release	Confirm scalability and adoption with real users	<i>Adoption rate thresholds met; SLA compliance demonstrated; support model defined</i>
05	Full Production	Operate at scale with monitoring and governance	<i>Runbook approved; monitoring live; owner accountabilities documented</i>

SpanForge Pilot Risk Index™ — What Each Gate Prevents

Each gate is designed to prevent one of five specific failure modes. These failure modes are drawn from the governance, data, and lifecycle gaps consistently identified across S&P Global Market Intelligence (2025), Gartner (2024, 2025), and McKinsey & Company (2025).

Gate	Failure Mode Prevented	How the Gate Prevents It
Gate 1 — Scoping	Wrong problem defined	Forces a written, signed problem statement before any technical work begins
Gate 2 — PoC	Technical infeasibility undetected	Requires demonstrated accuracy on representative data before pilot investment
Gate 3 — Pilot	No measurable business value	Mandates KPI validation in a real operating environment, not a sandbox
Gate 4 — Limited Release	No user adoption	Requires measured adoption rates and SLA compliance from actual users before full deployment
Gate 5 — Production	No scalability or governance	Confirms runbooks, monitoring, and owner accountability before removing programme controls

The Gate Readiness Score™

Before each gate review, the programme team produces a Gate Readiness Score™ — a structured 0–100 assessment across five dimensions: technical readiness, business value evidence, operational integration, compliance status, and stakeholder alignment. A score below 70 triggers an automatic remediation plan; no gate review is scheduled until the threshold is met.

The Gate Readiness Score™ serves two functions. It gives teams a clear, measurable target before each review, removing ambiguity about what ‘ready’ means. And it gives the Gate Authority a structured brief rather than a narrative presentation, making advancement decisions faster and more defensible.

The Decision Framework at Each Gate

- **Advance:** All conditions satisfied. The initiative moves to the next stage on a confirmed schedule.
- **Conditional Advance:** Conditions substantially met with time-bounded remediation commitments. Advances conditionally with a defined review trigger.
- **Return:** Conditions not met. Returns to the current stage for a defined remediation period. This is not failure — it is the gate working as designed.

CONTRARIAN TAKE

A gate that every initiative passes is not a governance mechanism. It is a rubber stamp with extra steps. If your gate reviews never result in a Return decision, your evidence thresholds are too low.

Understanding the system is one thing. Seeing its design logic made concrete is another. The following section illustrates the problem it is built to solve.

SECTION 5

Where the SpanForge Exit Gate System™ Should Not Be Applied

“Intellectual honesty about scope is a precondition for credibility. Not every AI initiative needs a five-stage lifecycle.”

The SpanForge Exit Gate System™ is designed for AI initiatives with material business impact, production integration requirements, and meaningful stakeholder accountability. Three categories fall outside its intended scope:

- **Low-risk internal automation.** Scripted workflows and rules-based tools with limited blast radius do not require five-stage governance. A lightweight two-gate model is appropriate.
- **Experimental R&D and horizon-scanning.** Exploratory research where the objective is learning rather than deployment should not be constrained by production-readiness gates. Apply time and budget boundaries instead.
- **Internal prototypes and technical spikes.** Short-duration technical investigations (typically under six weeks) to answer a specific feasibility question are inputs to the Gate 1 scoping decision, not pilot programmes.

If an initiative could cause material disruption if it fails in production, affects customers or regulated data, or requires significant cross-functional coordination to operate, it belongs inside the SpanForge Exit Gate System™.

CONTRARIAN TAKE

Applying enterprise governance to every AI experiment is as damaging as applying no governance at all. Calibrate the mechanism to the risk.

SECTION 6 | ILLUSTRATIVE SCENARIO

What the Exit Gate System™ Is Designed to Change

“SpanForge is at the beginning of its journey. This scenario illustrates the problem the system is built to solve — and what a different outcome looks like when structure replaces ambiguity.”

The pattern described below is not a specific client engagement. SpanForge is a startup and does not yet have a client track record to report. What follows is an illustrative scenario, grounded in the conditions that the S&P Global and Gartner research cited in this paper identify as the primary drivers of enterprise AI pilot failure. It is intended to make the problem concrete and the system’s design logic visible.

The Situation the System Is Built For

Consider a regulated organisation with several AI initiatives in active pilot status simultaneously. Based on the patterns described in the research cited in this paper, the following conditions are typical of organisations experiencing high abandonment rates:

- No written exit criteria exist for any of the pilots.
- Several pilots have no defined business sponsor with budget accountability.
- Progress is tracked by tasks completed, not by business outcomes.
- Some pilots have been restarted due to model churn — each restart triggered by a new model release rather than evidence of failure.
- No pilot has undergone a formal production-readiness review.

These are not exceptional conditions. They are the norm. The Gartner research cited in this paper found that fewer than one in five pilot-phase projects had written, agreed success criteria before development commenced. The S&P Global data shows that approximately 42% of enterprises reported abandoning the majority of their AI initiatives in 2025. The pattern is consistent and well-documented.

THE CORE PROBLEM

In these conditions, the AI team is not failing because of a capability gap. They are failing because no shared definition of ‘done’ exists. Without one, there is no mechanism to declare completion — only exhaustion.

Figure 5: What Changes When Structure Is Applied

The table below shows the contrast between the conditions the research describes and the conditions the SpanForge Exit Gate System™ is designed to create. This is the design intent of the system, not a reported outcome.

Without Structure vs. With SpanForge Exit Gate System™ — Illustrative Contrast	
WITHOUT STRUCTURE — Typical conditions at abandonment	WITH EXIT GATE SYSTEM™ — Intended design outcomes
✗ 8 pilots active, 0 exit criteria written	✓ Every initiative has written, signed gate plans
✗ 5 pilots with no budget-accountable sponsor	✓ Gate Authority includes budget-accountable sponsor
✗ 3 pilots restarted 2+ times due to model churn	✓ Advancement decided by evidence, not momentum
✗ All tracked by tasks, not outcomes	✓ Only initiatives meeting gate criteria advance
✗ 2 pilots > 14 months with no gate review	✓ Closures are documented, not quietly abandoned
✗ CDO confidence: 3.1 / 10	✓ Sponsor confidence in AI delivery increases
✗ AI team: 'constantly busy, never done'	✓ Fewer pilots, but higher proportion reach production

Figure 5: Illustrative contrast. 'Before' conditions are drawn from patterns identified in S&P Global Market Intelligence (2025) and Gartner (2025). 'After' conditions represent the intended outcomes of the SpanForge Exit Gate System™ design.

How the System Changes the Trajectory

When the SpanForge Exit Gate System™ is applied to a portfolio in this condition, the design logic works as follows:

Step 1	Each initiative is mapped against the five stage definitions. A Gate Readiness Score™ is produced for each. In most cases, initiatives believed to be 'nearly ready for production' are revealed to be at Stage 01 or Stage 02. This is the clarifying function of the assessment.
Step 2	A Gate Authority is constituted for each initiative — a cross-functional panel that must include a budget-accountable business sponsor. Initiatives without an identifiable sponsor are flagged for closure or reassignment before proceeding.
Step 3	Gate plans are written: a document specifying the evidence required for the next gate, the accountable owner for each requirement, and the target review date. Advancement becomes a question of evidence, not negotiation.
Step 4	At each gate review, the Gate Authority applies the three-outcome framework: Advance, Conditional Advance, or Return. Initiatives that cannot meet gate

conditions within a defined remediation period are formally closed — not abandoned quietly, but closed with documented rationale and captured learnings.

The expected directional outcome of this process — consistent with the governance gap analysis in the cited research — is that fewer initiatives proceed further, but those that do proceed are materially more likely to reach production. Closing initiatives that cannot pass Gate 1 or Gate 2 is not failure. It is the system working correctly: concentrating resources on initiatives with a genuine path to production value.

SpanForge is actively seeking early adopter organisations to implement the Exit Gate System™ and build the evidence base. If your organisation is navigating the conditions described above, we would welcome the conversation.

SECTION 7

Implementing the SpanForge Exit Gate System™: Practical Guidance

“Organisations that implement structured AI lifecycles do not merely reduce abandonment. They build an institutional capacity to deliver AI at scale.”

Adopting the SpanForge Exit Gate System™ requires three foundational commitments, applied consistently. The commitments below are designed to directly address the governance gaps that the S&P Global and Gartner research cited in this paper identify as the primary causes of enterprise AI pilot failure.

Commitment 1: Retroactive Baseline Assessment

Map every active initiative against the system’s stage definitions. In most cases this produces a clarifying shock: initiatives described as ‘nearly ready for production’ are often revealed to be at Stage 01 or early Stage 02 in practice. The assessment must include business representatives, not only technical staff.

Commitment 2: Gate Authority Designation

Each initiative must have a named Gate Authority before entering the lifecycle model. This must include a business representative with budget accountability — not a delegate. Without this, gate decisions become advisory rather than binding.

Commitment 3: Evidence-First Culture

The most important change is cultural: replacing progress narratives with evidence artefacts. The shift requires explicit endorsement from executive sponsors and consistent reinforcement from programme leadership.

Common Implementation Mistakes

- **No business owner in the Gate Authority.** Without a budget-accountable sponsor, advancement decisions revert to consensus. The gate becomes a formality.
- **Skipping the retroactive assessment.** Applying the system only to new initiatives while existing pilots continue under old governance creates a two-tier structure that undermines credibility.
- **Treating gates as approvals.** The moment a Gate Authority grants advancement based on relationship capital rather than documented evidence, the system is broken.
- **Setting Gate Readiness Score™ thresholds too low.** Teams presenting at reviews with scores below 70 typically have not produced the underlying evidence. The threshold is a filter, not a formality.
- **Failing to document Return decisions.** Undocumented returns create ambiguity about what the team must produce and eliminate the audit trail that makes progress legible.

CONTRARIAN TAKE

If you believe you have no bandwidth to implement lifecycle governance, you almost certainly have more failed pilots consuming that bandwidth than you have realised. The SpanForge Exit Gate System™ creates capacity; it does not consume it.

Recommended 90-Day Adoption Roadmap

- **Days 1–15:** Portfolio audit. Stage map and Gate Readiness Score™ produced for each active initiative.
- **Days 16–30:** Gate Authority designated for all initiatives. Gate plans drafted with evidence requirements and target review dates.
- **Days 31–60:** First gate reviews for initiatives within 60 days of a natural stage transition.
- **Days 61–90:** Portfolio rationalisation decisions. Formal closure of initiatives that cannot meet current gate conditions within a defined remediation period.

CONCLUSION

The Pilot Crisis Is Solvable — But Not By Accident

The doubling of AI pilot abandonment rates between 2024 and 2025 is a diagnostic signal. The dominant approach to enterprise AI — informal, iterative, milestone-free piloting — does not scale. It fails not because technology is inadequate or teams lack capability, but because ambiguity compounds. Without a structural mechanism to resolve it, the weight of unresolved questions eventually exceeds the organisation’s willingness to continue.

The SpanForge Exit Gate System™ resolves ambiguity at scheduled, mandatory intervals. Every advancement decision is made on evidence rather than momentum. Every closure is documented rather than apologised for.

“AI advantage will not be won by those who experiment the most, but by those who operationalise the fastest.”

Organisations that implement structured AI lifecycles will not merely reduce their abandonment rates. They will build an institutional capacity to deliver AI at scale — the durable competitive advantage that separates AI-native organisations from those perpetually stranded in purgatory.

READY TO BREAK OUT OF PURGATORY?

4-Week AI Portfolio Diagnostic

A fixed-scope engagement with three defined outcomes:

- Identify failing pilots within 2 weeks — full stage map against the SpanForge Exit Gate System™
- Gate Readiness Score™ and SpanForge Pilot Risk Index™ for every active initiative
- Reduce abandonment risk within 90 days — Gate Authority design and implementation roadmap delivered

sriram@getspanforge.com • getspanforge.com

REFERENCES

S&P Global Market Intelligence. (2025). Voice of the Enterprise: AI & Machine Learning, Use Cases 2025. Survey of 1,006 midlevel and senior IT and line-of-business professionals across North America and Europe, conducted October–November 2024. S&P Global.

Gartner, Inc. (2024, July). Gartner Predicts 30% of Generative AI Projects Will Be Abandoned After Proof of Concept By End of 2025. Gartner Data & Analytics Summit, Sydney.
gartner.com/en/newsroom/press-releases/2024-07-29-gartner-predicts-30-percent-of-generative-ai-projects-will-be-abandoned-after-proof-of-concept-by-end-of-2025

Gartner, Inc. (2025, February). Lack of AI-Ready Data Puts AI Projects at Risk.
gartner.com/en/newsroom/press-releases/2025-02-26-lack-of-ai-ready-data-puts-ai-projects-at-risk

McKinsey & Company. (2025). The State of AI in 2025: Agents, Innovation, and Transformation. Survey of 1,993 participants across 105 nations, conducted June–July 2025.
mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

ABOUT SPANFORGE

SpanForge is the AI lifecycle platform for enterprise teams — covering every phase from deciding whether to build, through architecture and engineering, to governance and scale. Built around the T.R.U.S.T. Framework (Transparency, Responsibility, User Rights, Safety Guardrails, and Traceability), SpanForge provides the tools, frameworks, and standards that get AI to production and keep it there. Its flagship product, AgentOBS, delivers production observability for autonomous AI agents — baselining behaviour, detecting drift, and enforcing consent boundaries before issues reach regulators or users.

Explore the platform at getspanforge.com.

For enquiries: sriram@getspanforge.com

© 2026 SpanForge. All rights reserved. SpanForge Exit Gate System™, SpanForge Cost Model™, SpanForge Failure Funnel™, Gate Readiness Score™, SpanForge Pilot Risk Index™, AgentOBS™, and T.R.U.S.T. Framework™ are trademarks of SpanForge. getspanforge.com