ENTERPRISE TRANSFORMATION

Transforming the enterprise with agentic AI

How a business control plane connects AI capability to business value

Document ID VP-ETRAN-2026-01 | November 2026

For business and technology leaders

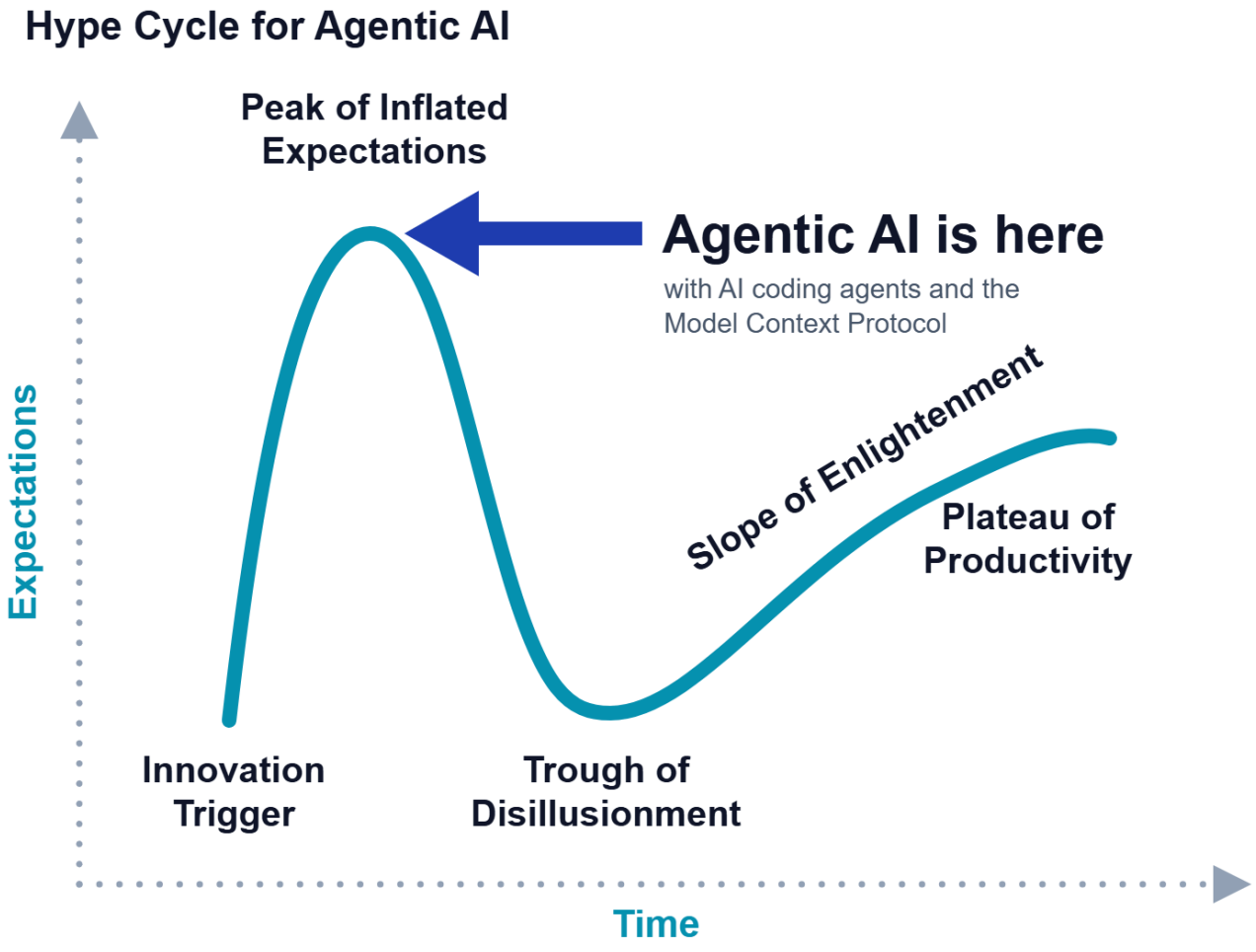
Contents

01 From experiments to operations
02 Why agent programs stall at scale
03 How a business control plane works
04 A refund workflow in practice
05 Staffing human review
06 Verifying decision evidence
07 Scaling from one workflow to a fleet
08 Measuring value and cost
09 A practical path to adoption
10 Conclusion
11 About VeriProof
12 References

Executive summary: A retailer lets an agent issue eligible refunds up to $500. When the agent proposes $900, the payment pauses. A manager can authorize the exception, and the payment system reports whether the money moved. Those are three separate facts a business owner needs before delegating the next refund. A business control plane connects the policy, the person with authority, and the recorded outcome. This paper uses that example to show how teams can measure the cost of a workflow and decide when to extend it to other agents.

The market is moving from experiments to operations

Gartner's Hype Cycle tracks how expectations for a technology rise, fall, and recover as organizations learn what it can deliver. [2] The April 2026 Hype Cycle for Agentic AI places agentic AI at the Peak of Inflated Expectations. AI coding agents and the Model Context Protocol sit beside it. The next phase in the model is the Trough of Disillusionment, when results fall short of early promises and interest wanes.



Source: Gartner, Hype Cycle for Agentic AI, April 2026. Simplified illustration; positions are approximate.

Figure 1. Agentic AI at the Peak of Inflated Expectations. Simplified illustration based on Gartner's Hype Cycle for Agentic AI, April 2026. [2]

Gartner also tracks governance and agent management as distinct practices with their own paths to maturity. Enthusiasm for agents has run ahead of the practices needed to manage them. That gap

shapes what happens in the trough. When funding comes up for review, leaders need evidence of what a workflow delivered and what it cost.

Getting through that phase starts with the work itself. McKinsey argues that capturing value from AI requires reorganizing roles, workflows, and decision rights around outcomes. [3] Gartner's analysis of agentic AI returns points toward specialized agents that expand into workflows across functions. [4] The unit of transformation is a business process, such as a customer journey, a supply chain decision, or a financial close.

Each deployment should also leave something the next one can use: agreed business definitions, tested policies, trained reviewers, and a record of results. That reuse turns a series of projects into an enterprise capability, increasing the pace that new agents can be deployed.

Why agent programs stall at scale

Pilots test whether an agent can perform a defined task under controlled conditions. But daily operation brings four pressures that pilots rarely reveal.

Exceptions are the daily workload

In production, information arrives incomplete, policies change, connected systems fail, and customers make requests nobody anticipated. Every exception needs a response. The time people spend reviewing and correcting agent work is a real operating cost, and it belongs in the business case.

Delegation changes accountability

When an agent drafts text, a poor result costs the user some time. When an agent acts, the stakes are higher. An incorrect payment or an unauthorized change to a customer record has consequences for the business and its customers. The organization needs clear limits on each agent's authority, a way to intervene before an action takes effect, and confirmation of what happened afterward.

Engineering telemetry answers a different question

Traces, logs, and performance metrics explain how an agent ran. They rarely tell a business owner whether an action was permitted, who approved an exception, or whether the payment went through. Without a shared decision record, engineers must piece together the answer. That work is slow, and the answer can depend on who does the reconstruction. Months later, the team may struggle to connect an approval to the payment it authorized.

Coordination costs grow with every agent

When each team builds its own rules, approval steps, and records, every new agent adds another set to learn and reconcile. Risk teams review each design from the beginning. Reviewers work across different queues and formats. Auditors receive evidence in a different shape for every system. Each additional review process adds cost to the business case.

A better AI model does not solve these problems. What's needed is a shared operating model that connects technical execution to business responsibility.

A business control plane connects engineering and business authority

A business control plane provides a unified framework for defining permissible agent actions and maintaining a transparent record of their execution. It joins work that usually sits with separate teams. Business owners define the intended outcome and acceptable limits. Risk and compliance teams specify the controls, and engineers build the agent and its reporting. The model has four elements.

Business telemetry

Business telemetry is the set of structured facts an agent and its application report about a task: the relevant conditions, the proposed action, and the result. For a refund, that includes the amount, eligibility, reason, and payment status. Engineers implement the fields. Business teams define what they mean and which decisions depend on them. Shared definitions let each fact mean the same thing to the policy, the reviewer, and the auditor.

Policy checked before action

Business teams write policies in natural language, such as “Refunds above $500 require approval from an authorized manager.” Shared field definitions give engineers a precise basis for turning that sentence into a condition on the reported amount and the configured review authority. Teams then test that the implemented rule matches the intended policy. At run time, the application asks for a decision before it acts and follows the answer it receives.

Human authority for exceptions

Escalation brings in a person with the authority the agent lacks. The reviewer sees the business facts and the policy that triggered the request, then records a decision. That decision joins the original request, the conditions evaluated, and the application’s reported result in a single history. Business owners, compliance teams, and auditors can use that history without asking engineers to rebuild the case.

Verifiable evidence

A decision record is only as useful as the trust others place in it. Independently verifiable evidence lets an auditor, regulator, or customer examine the decision history and check that it has not changed since it was recorded.

The control plane complements security, identity, and performance monitoring. Technical signals can inform a decision, while business rules define which actions are permitted and who may approve exceptions. Clear ownership and review deadlines keep those responsibilities workable as the number of agents grows.

The model in practice: a refund workflow

VeriProof provides a business control plane for connected AI agents and automated workflows. [5] Its industry vocabulary defines shared fields for agent telemetry, business terms for policy authors, and readable labels for reviewers. [6] Each policy check returns one of five outcomes: Allow, Warn, Transform, Escalate, or Deny. The application then continues, continues with a warning, continues with a corrected version, waits for a person, or stops.

Return to the retailer's refund policy. An eligible refund up to $500 can proceed automatically. The $900 request needs a manager. Here is how the team follows it from proposal to settlement.

1. Prepare the request. The agent proposes a $900 refund. Before submitting the payment, the application reports the amount, eligibility, and reason in the agreed business fields and asks for a policy decision.
2. Apply the rule. The rule compares the reported amount with the $500 limit and returns Escalate. The application saves its state and pauses the payment. An eligible $300 refund would have returned Allow and proceeded without review.
3. Obtain an authorized decision. VeriProof routes the request to an eligible reviewer. The reviewer sees the amount, eligibility, and rule that triggered escalation, then records a decision in the VeriProof inbox or a connected work app. [7]
4. Act on the decision. After approval, the application verifies the signed outcome and rechecks current policy and workflow state before it proceeds. It also guards against a duplicate payment. At this point the approval is recorded, but the refund is not yet confirmed. [7]
5. Record the result. The payment application reports whether the refund completed and supplies a confirmation reference when one is available. VeriProof links that result to the policy and review records. Until the application reports an outcome, completion remains unconfirmed. [5]

Each step records a separate fact. The policy result records whether the action can proceed or needs review, and the approval records a person's decision. The outcome report records what the payment system did. Keeping these facts apart lets the record answer an auditor's questions precisely.

For a team that approves every refund by hand today, this design changes the work. Routine eligible requests proceed under agreed rules, and reviewers spend their time on exceptions that need their authority. Shared records also cut the handoffs needed to explain a decision to a customer or a supervisor. These are expected gains, so the business case should measure them before counting them.

Human review is an operating responsibility

Escalation works only when someone is ready to decide. A review queue that nobody owns becomes a bottleneck, and agents waiting on approvals deliver none of the value they were built for. Treat review as part of the operating model, staffed and managed like any other service.

Before launch, the workflow owner names who is authorized to decide each type of exception, who covers absences, and how long a request may wait. VeriProof supports eligible reviewer groups and due times that follow working hours and calendars. It also supports delegation to another eligible reviewer and escalation of overdue work to a fallback. [7] The organization still decides who staffs the queue and how much time reviewers have.

Separation of duties needs the same planning. Smaller teams, where one person may hold several roles, need a recorded fallback that keeps work moving and still shows who made the decision.

Evidence that others can understand and verify

A useful record tells the decision's story in business terms. For the refund, it preserves the requested amount, the reported eligibility, and the reason, together with the rule version and the conditions

evaluated. It shows why the request needed more authority, who decided, and what the payment application reported afterward. A supervisor can follow it without a developer's help.

Shared vocabulary makes that possible. VeriProof's industry lexicons map the fields that agents emit to readable labels and to the standards they come from. [6] The same terms carry through policy evaluation, human review, and the evidence record. Framework mappings then link recorded decisions to relevant controls and show gaps to their assigned owners. [8]

An auditor checks the exported files against the signed inventory using a trusted public key, then confirms the anchor with the ledger provider. The verifier runs independently of the VeriProof portal. [9]

For a refund, an auditor can verify that the exported approval and reported outcome match the records committed to the ledger. The payment system supplies a separate check: its transaction record shows whether the refund settled and for how much.

Scaling from one workflow to a fleet

A successful workflow leaves a reusable pattern: agreed business fields, tested policies, named authorities, and a record of outcomes. Related processes can adopt that pattern while keeping their own rules and decision owners. Payment and procurement agents might share a supplier identifier, for example, while applying different approval limits set by their own business owners.

Reuse works best when definitions are common. In VeriProof, a rule can apply to a group of applications. The group keeps one field contract, which a developer confirms once for the group. An application that joins the group later is covered by the group's rules from the moment it joins. [10]

Rule changes deserve the same care as code changes. A new or revised rule can run in Shadow first, recording what it would have decided without changing application behavior. Teams compare those results with real activity, enforce the rule on a few applications, and then widen it to the rest. [10] The approvals the organization requires apply at each step.

As adoption grows, each role needs its own view of the same work. VeriProof provides a portal for each role: business owner, risk and compliance officer, governance engineer, developer, auditor, and administrator. Built-in handoffs move work between those portals with an owner and a due date. A governance engineer can assign a missing field to a developer. Overdue reviews follow the configured escalation path. [11]

When a team adds an agent or changes a model, it should confirm the business facts reported and the actions checked before they take effect. [12] It should also confirm how results come back. Retest denied actions, human review, duplicate prevention, and recovery from failures before relying on the revised workflow.

The decision history also guides improvement. Repeated escalations for the same missing fact show where to check the application's reporting. Long waits suggest a staffing shortfall or an authority limit set too low. Assign an owner to each change and test its effect before expanding an agent's authority.

Measuring business value and the full cost of delivery

A credible business case compares the agent-supported workflow with a baseline for the same process and a similar mix of requests. Measure the intended outcome, such as resolution time, processing

capacity, or handoffs per case, alongside the rate of correct completion. Faster responses mean little if correction work outweighs the time saved.

Count the full cost of delivery: the model and platform, human review time, rework, and the ongoing support that integrations require. Gartner cited rising costs in its forecast of project cancellations, and many of those costs sit outside token billing. [1] For the refund workflow, useful measures include:

- The share of refunds completed correctly, reconciled against the payment system
- Reviewer time per escalated request and time spent waiting in the queue
- Incorrect or duplicate payments and the cost of correcting them
- Total cost per correctly completed refund, compared with the manual baseline

State the evaluation period and request volume so the comparison has context. Agree in advance how freed capacity will create value, whether by serving more customers, resolving complex cases sooner, or improving another process. If review queues grow or errors increase, investigate before expanding. Report estimates of losses avoided separately from observed savings.

A practical path to adoption

Progress in deliberate stages, and let each stage end with evidence that justifies the next. The first workflow carries the most setup work, because it establishes the definitions, roles, and review practices that later workflows reuse.

Select one workflow

Choose a process with clear value, measurable outcomes, and an accountable business owner. Early on, bring business, risk, compliance, and engineering together. Agree on the intended result, the telemetry required, the policy conditions, and who holds decision authority.

Prove it under real conditions

Establish the baseline and acceptance criteria before the evaluation begins. Test ordinary requests alongside missing facts, denials, delayed reviews, and failed execution. Run new rules in Shadow before enforcing them, and confirm that reviewers can meet the agreed due times.

Expand by pattern

Expand only after the workflow meets its criteria and its owners accept the operating responsibilities. Apply the proven pattern to a related process, and validate its rules and integration there. Leaders still need to redesign the surrounding work, prepare staff, and direct the capacity that agents release.

Conclusion

Agentic AI will change how enterprises serve customers and make decisions. The organizations that benefit most will be the ones that can delegate work with confidence. That confidence comes from operating discipline: shared business facts, policies checked before action, human authority for exceptions, and evidence that others can verify. A business control plane provides that discipline in a consistent manner, so each new agent builds on the last. Leaders can use those results to decide which workflows merit further investment.

About VeriProof

VeriProof is a business control plane for AI agents and automated workflows. It checks proposed actions against approved business policy, routes exceptions to authorized people, and keeps verifiable records of each decision and its outcome. Business owners, risk and compliance officers, governance engineers, developers, auditors, and administrators each work in their own optimized portal experience from the same decision record. Learn more at veriproof.app.

References

[1] Gartner: Agentic AI project cancellation forecast, June 25, 2025

[2] Gartner: What the 2026 Hype Cycle for Agentic AI Reveals

[3] McKinsey: AI is changing work. Now it has to change the organization

[4] Gartner: Mastering Agentic AI: Multimillion-Dollar ROI Lessons

[5] VeriProof: How VeriProof works

[6] VeriProof: Industry vocabulary

[7] VeriProof: How Human Review works

[8] VeriProof: Compliance Center and framework mappings

[9] VeriProof: Verify exported evidence

[10] VeriProof: Rules, Shadow, and staged rollout

[11] VeriProof: Analyst briefing on role portals and handoffs

[12] VeriProof: Choose an agent integration

Sources accessed November 4, 2026. The refund workflow and its amounts are illustrative.

Try one workflow with your team

Choose a refund policy your team uses today. Test one request within the limit and one that needs a manager. Follow each from the policy decision to the result your payment system reports.

Start your free 60-day trial

veriproof.app/start