# THE CONTROL TOWER FALLACY

## Visibility Is Not Governance

July 2026

By: Tristan Cox, Head of AI, Workato APAC

> Why enterprise AI governance is an infrastructure problem, not a monitoring problem

## Executive Summary

The AI industry has a vocabulary problem. The term 'control tower' has been touted as shorthand for enterprise AI governance, implying that if you can see your AI agents and how they are behaving, you have the governance problem solved.

You do not. Visibility is a precondition for governance. It is not governance itself.

This paper draws a clear distinction between what a Control Tower does and what a Control Plane does, using an analogy that avoids technical language: The airport. An airport control tower performs a specific, valuable function. It does not run the airport. Running the airport requires more. Control must span deep Execution layers if we are to harvest true, scalable value.

The distinction matters because organisations investing in AI governance are frequently being sold control towers when what they need is a Control Plane. Some of the most confident voices in this market are, to extend the analogy, the baggage handling division, the aircraft manufacturer, or the check-in desk. Each has real expertise. None of them should be running the airport.

Three appendices address specific Control and Execution capabilities the industry most urgently needs to improve: a cost paradigm to govern AI token consumption; an auditability paradigm that produces evidence a regulator can actually use; and a connectivity paradigm that extends the safeguards provided by the Control Plane into the Execution Plane, allowing your agents to span all of your organisational systems and assets effectively, to take execute deep action across those assets, and thereby drive tangible, measurable business value.

**The control tower tells you the planes are in the air.**

**The control plane tells you who authorised the flight, whether they had the right to do so, and exactly what happened from gate to gate.**

## In this paper


| | |
|---|---|
| The Control Tower vs The Control Plane | 3 |
| Three Mistakes Organisations Are Making | 7 |
| The Gaps the Industry Needs to Close | 9 |
| What Senior Leaders Should Be Asking | 12 |
| A1. The Cost Paradigm: Governing AI Token Consumption | 15 |
| A2. The Auditability Paradigm: Evidence, Not Logs | 22 |
| A3. The Connectivity Paradigm: Delivering Measurable Value, Organisation-Wide | 28 |


**“The goal is end-to-end AI governance, both organisational and technological, fully operational in step with your agent rollout. Partial implementation may be easier to justify and deliver, but also presents a fallacy this paper is about: the capabilities that are hardest to see are often the ones most critical to the success of strategic agent initiatives.”**

— Massimo Pezzini, Head of Research, Future of Enterprise, Workato

## The Airport That Has No Security

Imagine an airport with a world-class control tower. The air traffic controllers are experts. The radar is state-of-the-art. Every aircraft in the surrounding airspace is tracked in real time, and the controllers can direct any plane to any runway, clear departures with confidence, and reroute inbound traffic when conditions change. The sky above this airport is, by any measure, well governed.

Now imagine the same airport has no passenger security screening. The baggage handling operation is outstanding, sophisticated and automated, but the company running it has also been appointed, for reasons of procurement convenience, as the authority on who can access the terminal security scanners, the aircraft cockpit, the administration offices, flight control, and the wine cellar for the Chairman's lounge. The check-in staff have been told they are now also responsible for certifying planes as airworthy before each departure. The aircraft manufacturer has noted, helpfully, that their planes come with excellent onboard safety systems, and that this should count for something.

No rational person would board a flight from this airport.

This is not a contrived scenario. It describes, with reasonable accuracy, the state of enterprise AI governance in many organisations today. The control tower - the monitoring layer that shows what AI agents are doing - has improved considerably and is now useful in better implementations. **The control plane,** the infrastructure that governs identity, access, data flow, and accountability across everything AI touches, enabling deep action and execution, is largely absent. Let's make that plain: not an aeroplane, a plane. A Control Plane.

**Visibility is a precondition for governance. It is not governance itself.**

## What a Control Tower Actually Does

An air traffic control tower has a carefully bounded mandate. It manages the safe movement of aircraft in the airspace around the airport: which runway to use, when to land, when to take off, how aircraft navigate around each other in the air. When two are converging, the tower intervenes. When a pilot declares an emergency, the tower coordinates the response. Without it, commercial aviation would not function.

But the tower's authority ends where its mandate ends.

The tower does not screen passengers boarding those aircraft. It does not verify that the catering company loading meal trolleys holds the security clearance to operate airside. It does not know whether the fuel supplier's staff have been background-checked. It has no role in managing the retail tenants in the terminal, the conditions applying to the ground handling contractor, or the qualifications of the maintenance crew. Those functions belong to the broader governance infrastructure of the airport, what engineers call the control plane.

In enterprise AI, a 'control tower' is the monitoring layer that watches what AI agents do. It sees which agents are running, what actions they are taking, whether they are staying within defined parameters. When an agent behaves unexpectedly, the control tower flags it. When an agent needs to be halted, the control tower issues that instruction. Organisations that cannot see their AI agents have a serious problem. You need this capability.

An even more serious problem emerges when organisations have concluded that visibility is the same thing as governance. This error is harder to detect because, from the outside, it looks like the problem has been addressed! The reality is that a Control Tower is just one layer of a much broader and necessary capability:

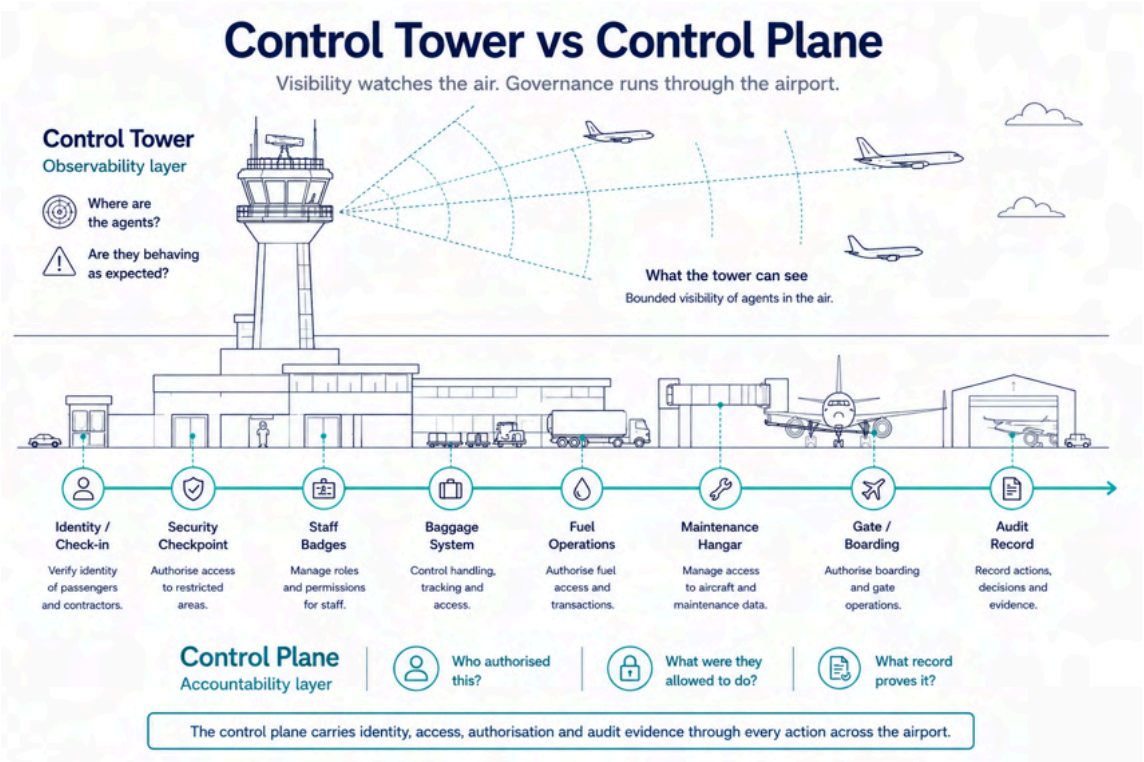
## The Control Plane: The Infrastructure You Don't See

The average passenger arriving at an airport interacts with perhaps a tenth of the infrastructure making their journey safe. They check in, pass through security, find their gate, and board. The process feels simple.

What they do not see is the governance system covering every person who touches their journey. The ground crew who loaded their bags holds an airside pass specifying exactly which areas they may enter, during which hours, and under what conditions. The fuel truck driver operates under equivalent constraints. So does the caterer, the maintenance technician, and the cleaning contractor. Every individual with airside access has been verified, credentialled, and their movements are auditable. When something goes wrong, investigators can reconstruct who was where, with what authorisation, and what they did. The record names people, not job titles or system labels.

This infrastructure is not visible to passengers and does not appear in the airport's marketing materials. Without it, the airport cannot be governed.

Where the tower watches what agents are doing, the Control Plane governs who the AI is acting on behalf of, what systems it can access, and whether a named human has authorised each significant action. These are different functions. One is observability. The other is accountability. The Control Plane is the fundamental enabler of  agents to execute *deep action;* securing and governing the ability for agents to operate within the organisation with scaled value.



The identity layer is where this matters most. When an AI agent executes an action in an enterprise system, that action should be attributable to the human who authorised it, the same way any airside access event is attributable to the specific badged individual who triggered it. Today, most AI agents act under a generic system credential. Think of it as a master key: it opens every door, is issued to no one in particular, and nobody audits it. The AI acts, the record shows a system name, and the human who directed the action is invisible.

> **The governance question is not 'can we see what the AI is doing?' It is 'can we prove who authorised it, and demonstrate that they had the right to do so?'**

That question requires a control *plane*. A control tower cannot answer it.

---

## Three Mistakes Organisations Are Making Right Now

### 1. Putting the baggage handling division in charge of the airport

Baggage handling is indispensable to aviation. Modern baggage systems are sophisticated: real-time tracking, automated sorting, integration with airline scheduling. The professionals running them are skilled. But no airport appoints the head of baggage operations as its chief executive, because the baggage domain is too narrow for the whole-of-airport role. The baggage division has deep expertise in one area. It does not have the perspective to manage security clearances, negotiate terminal lease arrangements, oversee customs integration, or coordinate the independent contractors across the full airport boundary.

The enterprise AI equivalent is asking an IT service management platform to govern an organisation's entire AI estate. ITSM platforms are valuable. They manage IT incidents, service requests, and operational workflows with real competence, and several major

vendors have built capable AI monitoring on top of that foundation. But the mandate of an ITSM platform is service management. Baggage is its core capability: important, specialised, and not responsible for the whole airport.

Asking an ITSM platform to provide access controls, identity governance, and audit infrastructure for AI agents across HR systems, financial platforms, citizen-facing services, and regulatory reporting is a scope mismatch. Baggage Handling will be well-run. The airport will not be well-governed.

## 2. Asking the aircraft manufacturer to run airport security

Boeing builds extraordinary aircraft. Their engineering rigour and technical depth are real. No airport asks Boeing to run its security screening operation, though, because Boeing's expertise is in building aircraft, not in managing the identity and access systems that determine who goes where in an airport and what they are permitted to do when they get there.

This error has a direct equivalent in AI governance. The companies that build large language models have serious expertise in AI research and model development. Their safety work is genuine and technically deep. Their models carry input filters, content guardrails, and output constraints. These are the aircraft's onboard safety systems: they matter and they work.

They operate at the model level. They manage what the AI will say and will not say, how it handles certain requests, and what content it declines to produce. They do not manage what enterprise systems the AI is permitted to access. Just like the control tower, they do not verify that the human directing the AI has the authority to do so. They do not create the audit record that a legal or compliance function requires to demonstrate accountability.

**The aircraft's onboard safety systems are necessary. They are not sufficient for airport security.**

Treating a model's built-in safeguards as enterprise AI governance is the Boeing-running-security error. The aircraft is safe. The airport is not governed.

## 3. Asking the check-in staff to perform aircraft maintenance

The third mistake is subtler and more widespread. It does not come from a vendor making claims beyond its expertise, but from the genuine complexity of the tools organisations reach for when they decide to take governance seriously.

The major cloud providers offer powerful platforms for building and governing AI systems. For organisations with deep cloud architecture expertise (the rare combination of specialists who understand identity systems, access control frameworks, network architecture, compliance tooling, and AI agent management together), these platforms provide an opportunity for real Control Plane capability.

The problem is the abstraction level. Aircraft maintenance requires specialist certification, specific equipment, years of domain experience, and a mindset entirely different from customer service. A competent, motivated check-in agent cannot become an aircraft maintenance engineer through additional training on top of their existing duties. The gap is too large.

Governing AI through a major cloud provider's native tooling requires the same kind of specialist expertise. It means navigating multiple separate administrative surfaces, each with different access control models, different audit log formats, and different policy enforcement mechanisms, while staying current with service configurations that change monthly. For most technology teams, this simply does not match their available expertise or operational capacity.

The outcome is predictable: organisations purchase sophisticated governance tools, deploy a fraction of their capability, and convince themselves they have addressed a problem they have only partially touched. The check-in desk now knows that aircraft maintenance exists. The aircrafts are still not maintained.

**The point of a Control Plane**

The purpose of a Control Plane is not to give organisations more powerful tools. It is to **provide governance capability at the right level of abstraction**, one that operations and compliance teams can manage without requiring specialist infrastructure engineering for every governance decision. It **must be unified and simplified** across all components in order to enable safe, deep action and execution across the enterprise.

## The Gaps the Industry Needs to Close

The AI governance market has made real progress on visibility. Monitoring dashboards are more capable than they were two years ago. Detecting agents operating outside expected parameters, tracking token usage, and alerting on anomalous behaviour are all improving. Control tower capabilities are largely immature from a technical point of view, however if developed and implemented well over the next twelve months, would be useful within their limited span of control.

The much broader Control Plane market is also maturing. The capabilities enterprises require cluster around the following four most impactful areas, none of which are technical novelties. All of them are standard enterprise governance requirements applied to a new type of actor, and extend their reach deep into the Execution layer:

### Identity at the point of Execution

When an AI agent acts on an enterprise system, the identity recorded in that system's audit trail should be the human who authorised the action, not a generic service account or system label. Governed identity pass-through means the AI acts on behalf of the authenticated human, carrying that person's permissions and leaving that person's verifiable record in the downstream system. It also means no action the AI takes can

exceed what the authorising human is themselves permitted to do. The AI cannot access systems or data that the directing human could not access directly.

Accountable enterprise systems have always worked this way. Applying the same standard to AI agents has simply not happened consistently, and its absence is the most significant governance gap in most current deployments.

## Governed memory

AI agents that operate across multiple interactions accumulate context: previous conversations, past decisions, data encountered in prior sessions. Without governance, this memory becomes an unmanaged liability. It persists beyond its useful life, may contain sensitive information never intended for long-term retention, and provides no record of what was stored, when, by whom, or why.

An airport that allowed contractors to accumulate sensitive documents with no retention schedule, no disposal procedure, and no record of what was held would face immediate regulatory breach. Enterprise AI memory requires the same disciplines applied to any other category of business information: defined retention periods, documented purge procedures, records of what was held and when it was destroyed, and protection against unauthorised access.

## Prompt governance

The instructions given to an AI agent, the rules and guidance that determine how it behaves, are a policy document. Changing those instructions without an approval process, without a version history, and without the ability to revert to a prior state is equivalent to editing a compliance policy without sign-off and discarding the previous version.

In most current deployments, an agent's operating instructions can be changed by anyone with platform access, with no audit trail for that change, no approval requirement, and no rollback capability. The result is a governance failure at the most important point in the AI's operating instructions.

## Audit trails that name people

The purpose of an audit trail is to answer a specific question: who took this action, with what authorisation, and when? An audit trail recording system names, service account identifiers, and API call sequences satisfies the letter of audit logging without answering the governance question.

An airport security log recording 'airside access event' without naming the individual would be found non-compliant by any investigator. An incident report describing what happened without identifying who was responsible would not survive a legal proceeding. AI audit trails that cannot trace to named, verified humans fail the same test.

**Organisations that can demonstrate their AI is governed are in a different position from those that can only demonstrate it is observed.**

## On Connectivity

Integration development overhead is one of the primary reasons AI programs stall between pilot and production. Production requires connecting to the real systems of record, with their authentication requirements, their data schemas, their exception handling logic, and their own governance obligations: permissions and boundaries that are inherited from the Control Plane. Building these connections from scratch for each AI deployment is expensive, slow, and creates a maintenance burden that compounds as the connected systems change over time. A strong Execution Layer replaces manual orchestration processes end-to-end, reduces error rates at every system boundary, satisfies governance requirements, and scales without a proportional increase in headcount. It should enable observability across AI agents and enterprise systems, while enforcing security and compliance policies. It should be ecosystem-agnostic, connecting to everything in your ‘airport’ so your organisation maintains flexibility and choice in where to direct new AI capabilities.

## What Senior Leaders Should Be Asking

Conversations about AI in leadership teams tend to start with capability questions: what can this AI do, how much will it cost, what are our competitors doing. These are legitimate starting points. But they are not sufficient.

The governance questions that will matter to audit committees, privacy officers, and regulators are different questions. They are worth raising now, while the infrastructure decisions are still being made. Governance must span AI in a way that enables the enterprise to leverage the power such governance provides, to scale AI in a safe and defensible manner to provide clear and measurable benefits realisation.

### On identity and accountability

When your AI takes an action in an enterprise system, whose identity is recorded in the audit trail? If the answer is a service account name, a system identifier, or anything other than a named, verified human who can be contacted, interviewed, or held responsible, you have an accountability gap that no amount of monitoring will close. In other words, the Control Plane must extend its reach into enforcement at the point of execution.

### On prompt governance

If you need to change the instructions your AI operates under, what is the process? Is there an approval workflow before the change takes effect? Is there a version history showing what changed, when, and who approved it? Can you revert to yesterday's instructions if today's change produces unintended consequences? If the answer to any of these is no, the most important governance document in your AI deployment is being managed at a standard that would not be acceptable for a routine operational procedure.

### On demonstrating compliance

If a regulator, an auditor, or a parliamentary committee asks you to demonstrate that your AI operated within defined parameters, that the people directing it had the authority to do so, and that you can account for every material decision it made, what evidence can you produce? A monitoring dashboard showing activity levels and error rates is an

operational record. Whereas an identity-linked audit trail tracing every AI action to an authorised human decision is compliance evidence. These are not the same thing, and regulators are developing the sophistication to tell them apart.

**The question worth asking your vendors**

Ask every AI governance vendor you are evaluating: 'When my AI agent takes an action in one of our enterprise systems, **what identity does that system's audit log record?**' The answer will tell you whether you are just being sold just a control tower or a full Control Plane.

## The Airport Needs More Than a Great Tower

A great control tower may seem like the right starting point. It is not the destination.

The airport analogy is useful because it separates capabilities that are routinely conflated in vendor presentations. The tower watches the air. It tells you where the planes are, whether they are on the right trajectory, and when something is about to go wrong. It is indispensable. But it does not screen passengers, govern contractors, or manage the identity of the people who fuel the aircraft. Those functions belong to the broader infrastructure.

The control plane is not a product you deploy. It is a set of enforced constraints that travel with every action your AI takes. Visibility tells you what happened. Policy tells you what should happen. But governance only exists at the point where the action plane cannot proceed without satisfying the control plane's requirements: where identity must be verified before a connection is bound, where an approval must be recorded before an AI-assisted workflow  is promoted, where an authorisation chain must be complete before an autonomous agent is permitted to act.

Organisations that have only the first (or have the first two and not the third), have documentation, not governance. The distinction will matter when something goes wrong, and it will matter to the people asking the questions.

Enterprise AI governance is in an early phase. The most visible capabilities, like dashboards, monitoring, and observability, are receiving quite a lot of attention. This is predictable. These capabilities are tangible and easy to show in a boardroom briefing. They look like governance because they look like control.

The Control Plane is harder to show and more important to build. It is the infrastructure through which accountability flows: identity enforced at every point where AI touches the enterprise, AI actions connected to the authorised humans who directed them, records created that can survive regulatory scrutiny. It is the difference between an organisation that can demonstrate its AI is governed and one that can only demonstrate its AI is observed.

That difference matters because the stakes of AI governance are not primarily technical. They are human. The people whose data is processed by AI systems, whose entitlements are shaped by AI recommendations, whose privacy is affected by AI-driven decisions, are not served by an organisation that can see its AI. They are served by an organisation that can account for it.

**The control tower tells you the planes are in the air.**

**The control plane tells you who authorised the flight, whether they had the right to do so, and exactly what happened from gate to gate.**

Both matter. Only one of them is governance.

—

Tristan Cox is Head of AI at Workato ANZ, advising enterprise and government organisations on AI strategy, governance, and implementation. The views in this paper draw on direct field experience and independent research. They are offered as a neutral framework for assessing AI governance. Readers are encouraged to test all claims against their own requirements and to hold every vendor, including Workato, to the standards set out.

**Appendix One**

# The Cost Paradigm: Governing AI Token Consumption

Token consumption is the most immediately measurable gap between having a control tower and having a Control Plane that extends into the Execution Plane. A tower shows you what AI is spending after the fact. A cost paradigm governs what it is permitted to spend, routes work to the right execution model before the bill is generated, and produces financial evidence comparable to any other operational cost category. Without it, AI budgets are not governed. They are hoped. In reality, there are components of cost control that span the control tower and the control plane.

In May 2026, Uber's COO acknowledged publicly that the company had burned through its entire annual AI budget in four months. The culprit was not a rogue system or a security breach. It was the mundane reality that AI agents consume tokens at a scale most organisations have not planned for, and almost none govern.

Uber's situation made headlines, but it was not unique. In the same period, Microsoft cancelled thousands of internal Claude Code licences, requiring developers to migrate to GitHub Copilot CLI by the end of its financial year on June 30. Per-engineer costs for AI coding tools had reached between $500 and $2,000 per month, with no framework for determining whether that spend was producing proportionate value. Separately, Peter Steinberger, the founder of the open-source AI project OpenClaw, ran 100 AI coding agents simultaneously for a month and generated a bill of $1.3 million in tokens. This was a deliberate experiment, run without budget constraints, to understand what agentic AI actually consumes at scale. It produced an unambiguous answer. Uber and Microsoft are among the most sophisticated technology organisations in the world, but demonstrate risks that even disciplined organisations face in the absence of cost governance. Steinberger shows the upper bound of what ungoverned agentic AI can spend when nothing stops it. If managing token costs requires deliberate governance even there, the challenge for organisations with less AI engineering infrastructure is proportionally larger.

Goldman Sachs projects that total AI token demand will multiply twenty-four times between 2026 and 2030, driven primarily by the growth of agentic deployments. The underlying arithmetic explains why: agentic systems do not make a single request and stop. They reason, query external tools, validate results, and loop on edge cases. Some agentic workflows already consume more than a thousand times the tokens of a simple

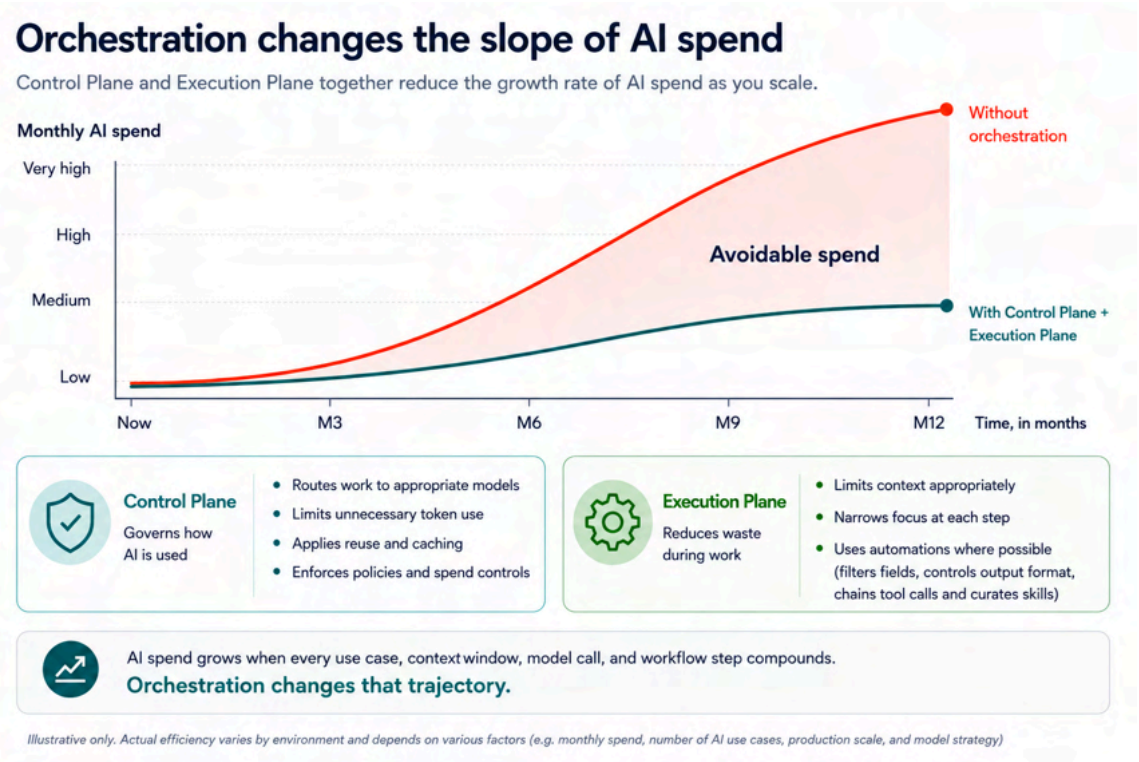
chatbot exchange. For organisations that have not built cost governance into their AI infrastructure from the start, the resulting bill is not incremental. It is structural.

## Why the architecture matters more than the model

There are a few parts to this problem. The answer lies partly in the Execution Plane, and partly in the Control Plane. The Control Plane provides the rigour and repeatable governance that enables the Execution Plane to be able to orchestrate workflows safely across multiple agents, applications, and human touchpoints.

The  instinctive response to a large token bill is to switch to a cheaper model. This addresses a symptom, not the cause. The structural driver of token waste in agentic systems is architectural, not model-related.

When an AI agent is given a task without a structured execution layer beneath it, the model must reason through the problem from first principles on every invocation. It must understand the context, decide which systems to query, interpret those responses, plan its next action, handle errors through further reasoning, and validate its own output. Every one of those steps consumes tokens. And the steps multiply with every additional connected system, every edge case encountered, and every ambiguous instruction in the original prompt.



Gartner's analysis of Model Context Protocol (MCP) architecture patterns confirms this directly. The aggregator pattern, where a single AI agent is exposed to all available tools and resources simultaneously, produces the highest token consumption of any MCP architecture because the model must reason through the full tool inventory on every call. Gartner states explicitly that poorly implemented MCP servers 'have caused big spikes in token use.' The composite pattern, where purpose-built MCP servers are scoped to specific use cases and expose only the tools and context relevant to that workflow, produces what Gartner describes as 'the most consistently low level of token use.' At enterprise scale, the difference between these two architectural choices is a budget decision, not a technical preference.

The principle that governs cost-effective agentic design is simple to state: the LLM should reason about *what* to do, not *how* to do it. Every 'how' that can be pre-encoded in a governed workflow, a pre-approved connection, or a policy rule should be. Every token spent on 'how' is a token that could have been eliminated without any loss to the quality of the outcome.

**Every 'how' that can be pre-encoded in a governed workflow, a pre-approved connection, or a policy rule, should be.**

## Drift: the hidden second cost

Token spend is the visible dimension of ungoverned AI. Drift is less visible, and in a government or regulated enterprise context it may carry heavier consequences.

Large language models are probabilistic systems. Given identical inputs on two separate occasions, a model will often produce slightly different outputs. For a consumer chatbot answering general questions this is inconsequential. For workflows where outputs carry legal weight, feed into official records, or determine entitlements, the picture is different.

Consider a document processing system operating at 97% accuracy. That sounds like a high bar. Across 10,000 documents in a year, a 3% error rate produces 300 incorrect outputs. In an FOI (Freedom of Information) context, each of those represents a potential statutory breach with its own investigation, remediation, and reporting obligations. The cost of the compliance response can substantially exceed the cost of the AI system that caused the problem.

Drift accumulates across three channels.

1. Compliance drift occurs when AI outputs diverge gradually from statutory or policy requirements as context shifts, prompts evolve, or model behaviour changes with updates. A system producing compliant output in February may produce subtly non-compliant output in June, with no single identifiable cause.
2. Data integrity drift enters when AI-mediated data entry or transformation introduces inconsistency into records systems over time, requiring expensive retrospective cleansing.
3. And rework loops compound both: when AI outputs require human correction, the corrected document is often sent back through the AI layer for formatting or finalisation, consuming additional tokens to address an error the AI caused in the first place.

## The three-tier answer

The architectural answer to both token cost and drift is the same. It is not to use less AI. It is to use the right execution model for each type of work.

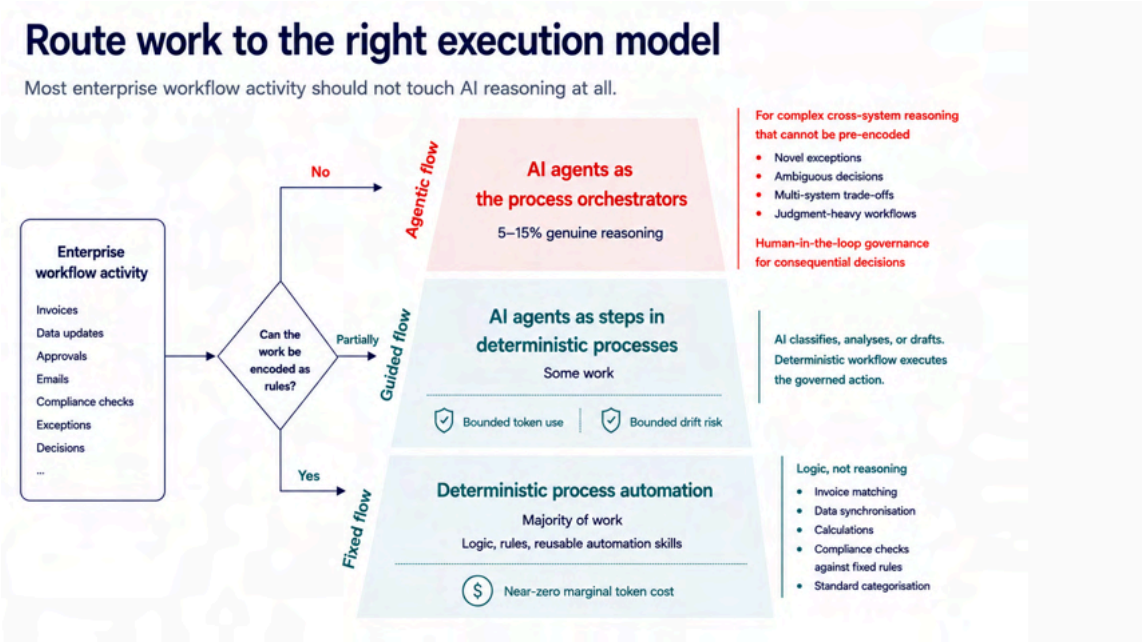
The majority of enterprise workflow activity is deterministic: invoice matching, data synchronisation, calculations, compliance checking against fixed rules, standard correspondence categorisation. These tasks do not require AI reasoning. They require logic, and logic executes deterministically at near-zero marginal cost, via reusable automation Skills, with the same output every time by design. AI reasoning should never touch these types of process steps..

The next most significant impact to token minimisation is hybrid: AI classifies, analyses, or drafts, but a deterministic workflow executes the governed action. The AI reasoning layer introduces some probabilistic variance, but consequential execution is deterministic and drift risk is bounded.

The remaining functions, which represent perhaps 5 to 15 percent of the orchestration, complex cross-system reasoning where the answer cannot be pre-encoded, genuinely requires full AI reasoning, with human-in-the-loop (HITL) governance at decision points that carry significant consequences. In a risk-based methodology, the appetite for human-in-the-loop control may shift over time as state trust is gained; levering opportunities for Human-On-The-Loop activities where a system acts more independently with batch or systematic human review. Some low-risk functions may traverse all the way to Human-Outside-The-Loop scenarios where documented rules are provisioned and approved, and the system behaviour proves as Deterministic as required for the process.

**The Three Tiers of Drift Control:**

- **Classic deterministic process automation**
- **AI agents as steps in deterministic processes**
- **AI agents as the process orchestrators**




Organisations that route work to the right tier report reductions in token consumption of 70 percent or more, with no measurable loss in output quality for the affected tasks. The savings come not from using inferior AI but from not using AI where a faster, cheaper, and more consistent alternative exists.

## Cost governance as Control Plane and Execution Plane functions

Token cost control is not a monitoring problem. A control tower can tell you how many tokens your agents consumed last month. A Control Plane determines, before execution, which model each task type should use, enforces that choice at the point of call, and routes tasks to smaller, faster models when the complexity does not justify the premium. At a minimum, multiple models must be available for connection so that an enterprise may actively choose to assign a specific agent to a specifically selected model, perhaps because of high performance alignment to a particular function (like mathematics), or where the required LLM function is adequately managed by a smaller, cheaper model.

Four principles matter here:

**Right model, right task.** A model appropriate for nuanced regulatory analysis is the wrong tool for extracting a postcode from a form field. Routing simple tasks to smaller, faster, cheaper models while reserving high-capability models for work that genuinely requires them is the most effective single lever for controlling AI spend at scale.

**Context management.** Passing the full conversation history into every model call is the most common source of token waste in production agentic systems. A Control Plane passes only the portion of prior interaction that the current step actually needs. In a ten-step workflow, the difference between accumulated context and managed context can be an order of magnitude in token consumption by the final step.

**Token budgets at the workflow level.** Individual task types carry a defined token allocation. When a task exceeds that allocation without producing a result, it escalates to human review rather than continuing to consume resources. This is the token equivalent of a purchase approval threshold: normal spend proceeds without friction, and unusual spend requires a decision.

**Finally, the choice of which controls to apply** is critical - AI agents were not ordered by a doctor, they should be used where they add value and not used where they do not. If a process can be automated, automate it! Leave the AI in the waiting room.

Without a control plane enforcing model selection, managing context, and tracking consumption against defined budgets by workflow type, AI spending is not governed. It is hoped for.

### What good looks like

A governed agentic system assigns each workflow a model tier, enforces that assignment at the API call level, manages context to pass only relevant prior state, and reports token consumption by workflow type against defined budgets. Variance above threshold triggers review before spend events compound.

This is not sophisticated technology. It is discipline applied consistently at the right layer. Select the right controls for the specific process, and don't be afraid *not* to use AI at all!

**Appendix Two**

# The Auditability Paradigm: Evidence, Not Logs

The second Control and Execution paradigm addresses auditability. The word 'auditable' appears in most AI governance frameworks. It rarely comes with a definition of what an audit trail must actually contain to satisfy a regulator, an auditor, or a legal proceeding. This gap matters because logging and auditability are not the same thing, and most organisations currently have one while believing they have both.

Logging records that something happened. Auditability records who was responsible, what they were permitted to do, and how that permission was exercised. An AI governance framework that delivers the first without the second has a monitoring system, not a compliance one. The distinction is not technical. It is the difference between a record that satisfies an operations team and evidence that survives a regulatory inquiry.

## The scale of exposure

The exposure is already measurable. Gartner's February 2026 research found that 60 percent of IT and security respondents already had evidence of, or suspected, unsanctioned AI agent activity within their own organisations. By 2028, Gartner projects that a quarter of all enterprise generative AI applications will experience at least five security incidents annually.

For organisations operating at government security classification levels, the stakes of those incidents are different. An undetected data exfiltration event through an ungoverned MCP server in an IRAP-assessed environment is a reportable breach with potentially serious consequences, handled under a different framework than a standard IT security incident.

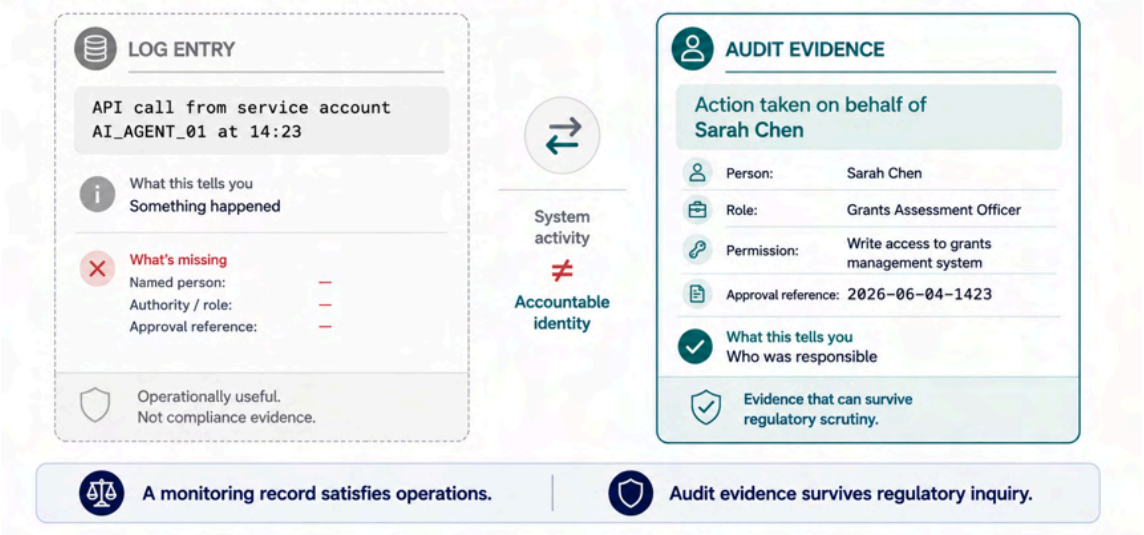
The Australian Government robodebt scheme ran for years before it could be unwound, partly because the automated debt calculation logic left no human-attributable decision trail. A Royal Commission investigation specifically called out the absence of records linking system outputs to accountable officials. That's exactly the "system account, not

a person" auditability failure that we need to prevent with any semi-autonomous system.

## The three requirements of a genuine audit trail

### Identity: the record must name a person, not a system

The entry 'API call from service account AI_AGENT_01 at 14:23' is a log entry. The entry 'Action taken on behalf of Sarah Chen, Grants Assessment Officer, role: write access to grants management system, approval reference 2026-06-04-1423' is audit evidence. The first tells you something happened. The second tells you who is responsible.




Human identity must pass through the AI layer to the system where the action is recorded. When an AI agent writes to an HR system, changes a configuration, or triggers a payment, the downstream system's audit log should record the human who authorised the action, not the AI platform's service account. An AI governance system that captures all activity internally but records a system identifier in downstream systems has half an audit trail. In a regulatory context, half an audit trail is functionally the same as none.

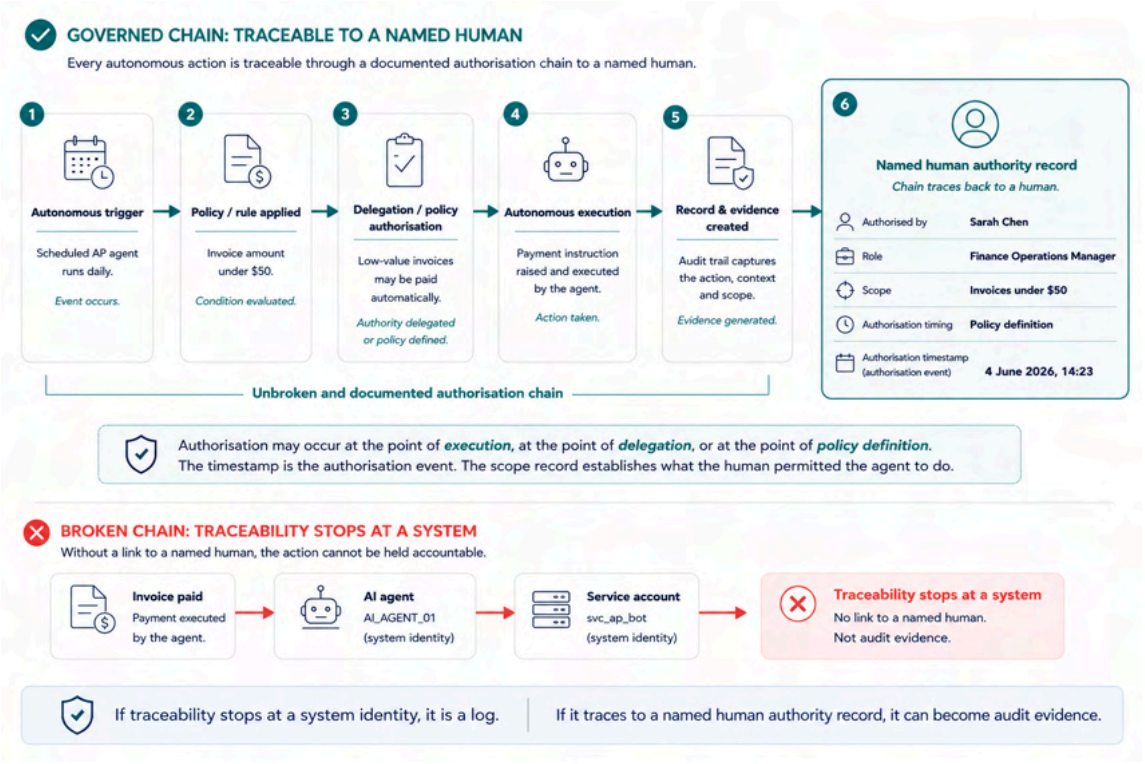
### Authorisation: the record must show what the actor was permitted to do

Recording who directed an action is not sufficient. The record must also show what the directing person was permitted to do, and that the action fell within those permissions. This requires the access rights of the human to be captured at the time of action, not reconstructed later.

Role-based access controls serve two distinct functions here. Access rights determine which systems and data the AI can reach on behalf of the authorising human. Action rights determine what the AI can do within those systems: read but not write, write to specific record types but not delete, initiate transactions below a defined threshold without further approval. Both must be recorded, and both must be enforced at the point where the AI touches the external system, not just at the platform boundary

## The authorisation chain

The identity requirement becomes more complex when AI agents act autonomously: on a scheduled trigger, in response to an environmental signal, or as a downstream step in a chain of inferences where no human issued an instruction at the moment of execution. For example, a scheduled agentic function that automatically processes accounts payable may have a rule that any invoice received that is for an amount under $50 will be paid without human scrutiny (but perhaps be subject to a later randomised audit). The correct governance model is to define an unbroken and documented authorisation chain traceable to a named human: Whether that authorisation occurred at the moment of execution, at the point of delegation, or at the point of policy definition. The name in the audit trail is that of the authorising human. The timestamp is the authorisation event. The scope record establishes what that human permitted the agent to do.




An airport smoke detector will trigger significant activity in an airport. The correct governance model is for an unbroken authorisation chain in the managed response, enforced by the Control Plane to the Execution Plane. Every action an AI agent takes must be traceable, through a documented chain, to a named human authorisation (in this case the accountable official on the Airport Emergency Plan). The name in the audit trail is the authorising human. The timestamp is the authorisation event. The scope record establishes what that human permitted the agent to do.

The test is whether the chain terminates in a human. If it does (i.e. if every autonomous action can be traced back to a documented human decision), the deployment is governed. If it does not (i.e. if an agent can take material actions that cannot be traced to any named human decision at any point in the chain), the deployment is ungoverned by definition, regardless of how well its runtime behaviour is observed. A monitoring dashboard cannot fix a broken authorisation chain. It can only confirm that an ungoverned action occurred.

The control plane enforces these constraints at the architecture level. The Execution Plane ensures AI does not carry credentials that would allow it to exceed the authorising human's permissions. Self-escalation is not a policy violation to be detected and remediated after the fact. It is structurally impossible.

### The downstream record as evidence

An AI governance platform that records a service account in the audit logs produces two disconnected records. One record names a system, the other names the agent platform. Neither names a human.

Governance has failed, because a regulator investigating an AI-assisted decision in a benefits system will go to the benefits system's audit log. If that log has recorded a service account as taking an action, the investigation stops there with a finding of non-compliance.

## The human-in-the-loop approval as an audit artefact

When a human approves an AI-recommended action as part of a governed workflow, that approval event is itself a compliance artefact. It is the point at which a human takes accountability for an AI recommendation, and the record of that moment requires the same treatment as any other significant decision in a regulated process.

A compliant approval record captures: the identity of the approving human; their role and the access rights associated with that role at the time of approval; the action they approved, including the information the AI presented to support the recommendation; the timestamp; and the unique identifier linking this approval to the downstream action taken as a result.

Financial services organisations apply this standard to human approvals of material transactions. Government agencies apply it to administrative decisions with significant consequences. AI-mediated actions that replace or support those decisions require the same standard. The involvement of an AI in generating the recommendation does not reduce the accountability requirement for the human who approved it.

## The airport standard

An airside access event at a major airport is not recorded as 'access occurred.' It records the badge identifier, the verified identity behind that badge, the specific gate, the time, and the system that authorised entry. That record is retained to a defined schedule, exportable on demand, and tamper-evident. When an incident occurs, investigators go directly to that record and it tells them who was where.

That is the standard. It is achievable for AI. Organisations that apply it will find that the question 'can you demonstrate accountability for this AI-assisted decision?' has a clear

answer. Those that have built a monitoring dashboard and called it governance will find the question considerably harder to answer than they expected.

**The audit test**

When evaluating an AI governance solution, ask the vendor to walk through a specific scenario: an AI agent has approved a transaction in your financial system. Show me the record in that financial system. Show me that it names the human who authorised it, their role, and the scope of their permissions at the time. Show me the approval event, including what information the approving human was shown before they agreed. If the vendor cannot demonstrate all three, the system meets logging standard, not audit standard.

**Appendix Three**

# The Connectivity Paradigm: Spanning Your Organisation for Measurable Value

The third paradigm is also the most commercially significant. The cost and auditability paradigms described in the preceding two appendices govern what AI spends and what it can prove. The Control Plane concept introduced in the main body (governing identity, access, and accountability ) has a broader expression at the infrastructure level: Orchestration is what makes that governance meaningful at scale. The connectivity paradigm is that broader expression in practice; *it determines what AI can actually do in runtime*. It is the extension of governance to the Execution Plane that unlocks power and scalability for transformational benefits realisation. Without it, the most well-governed AI deployment in the world is confined to giving answers about the systems it was pointed at during deployment.

Most of what is currently sold as AI transformation is, when examined carefully, AI acceleration. A faster way to draft a document. A smarter way to search a knowledge base. A more responsive interface to an existing process. These are real improvements. They are not transformative.

Transformation occurs when the outcome of a process changes, not just the speed at which a human completes one step within it. An AI agent that drafts a procurement contract faster than a human is useful. An AI agent connected to the procurement system, the supplier database, the approval workflow, the contract register, and the relevant stakeholders' notification channels, which can take a contract from initiation to signed archive without human hands on any step that does not require human judgment, is something different. The difference in this kind of multi-step orchestration is not the AI. It is the connectivity.

## Intersection of the Control Plane and the Execution Plane

The Control Plane defines blueprints, policies, semantic context, identity and optimisation constraints; the Execution Plan operationalises those definitions by assembling runtime context, executing task graphs and returning telemetry and outcomes to the Control Plane for audit, drift detection and continuous optimisation.

The Control Plane exposes agent-friendly interfaces and interoperability surfaces (e.g. MCP gateways) so that the Execution Plane can discover tools, fetch authenticated live data and perform grounded actions without bespoke integrations.

When an agent takes action, the Execution Plane emits telemetry (decisions, prompts, model/tool calls, cost/latency metrics) that the Control Plane consumes to produce lineage, explainability, compliance evidence and to enforce drift/exception handling and budget/policy constraints.

Human-in-the-loop interactions are coordinated across the two planes: the Control Plane sets which actions require human approval and the Execution Plane surfaces HITL checkpoints in runtime for review and final authorisation.


## What Execution through Enterprise Orchestration Actually Means

Enterprise orchestration is the coordination layer that makes genuine AI transformation possible. It sequences and governs the interactions between people, processes, and systems: the ERP that holds financial authorisation rules, the HR platform that manages employee records, the CRM that holds customer history, the document management system that holds the official record, the communication channels through which

approvals are sought and given, and the custom or legacy applications that store the data most specific to how that organisation actually works.

Large organisations typically run across hundreds of distinct applications. The data that defines how the business operates is fragmented across all of them. An AI agent with access to a slice of that landscape can produce outputs that are accurate about what it can see and wrong about everything it cannot. Governance rules, customer eligibility, financial limits, compliance obligations, and prior context are all attributes that live in systems the agent cannot reach unless the connectivity infrastructure is in place.

Orchestration is not simply a collection of API connections. It is the governed sequencing of interactions across those systems, with each step validated, each handoff logged, each exception handled, and each human decision point surfaced at the right moment to the right person. The orchestration can now be considerably more complex, including agent-to-agent interactions, AI mediation of enterprise systems, and composition of reusable skills that span different applications. An AI agent without orchestration can reason and recommend. It cannot act. And agents that cannot act do not produce operational outcomes.

## The Multiplier, not the Adder

The value of orchestration is multiplicative, not additive. A process that previously required three people to coordinate manually across four systems can run end-to-end when the right orchestration layer connects those systems. The time saving at any individual step is real but modest. The elimination of handoff friction, coordination overhead, error accumulation at transfer points, and the delays introduced by asynchronous human coordination is where the step-change in value lives.

Consider what happens to a standard HR onboarding workflow when every system involved is connected. The new employee record triggers access provisioning in identity management, creates the equipment request in the asset system, enrols the employee in the relevant training system, notifies payroll, opens the relevant project memberships, and sends the manager the day-one checklist, in sequence, governed, and with every action logged. Each individual automation step in this chain is straightforward. The orchestrated end-to-end sequence is transformative because it eliminates the coordination that previously required a human to manage the handoffs, chase the responses, and catch the gaps.

The same pattern at higher stakes reveals what is actually at issue. Consider a grants disbursement workflow in a government agency: An application triggers eligibility verification against the program rules, identity and entity checks on the applicant, screening for duplicate or fraudulent claims, and validation against the remaining appropriation in the financial system, before the complete picture is assembled for a grants assessment officer. The officer makes the assessment, per the legislation requirement for a named, accountable decision-maker. On their decision, the payment instruction is raised in the financial system, the applicant is notified, and the decision record, naming the officer, the delegation under which they acted, and the basis for the assessment, is filed where an auditor, an Ombudsman, or a review tribunal will look for it.

No single application vendor owns this chain. The grants management system, the identity service, the fraud-screening layer, the financial system that holds the appropriation, and the records system that holds the official decision are, in almost every agency, products from different vendors procured at different times under different mandates. An IT service management platform cannot reach the appropriation ledger. An ERP cannot own the eligibility-assessment layer. A productivity suite is not the system of record for the decision. The orchestration layer is the only thing positioned to sequence the chain, enforce that the assessing officer's identity and delegation flow through to the payment record, and produce the audit trail a public-sector review will demand. This is not the acceleration of a single step. It is a regulated, multi-system, publicly accountable outcome that does not exist without the connectivity layer beneath it. This is Execution of business objectives.

As Appendix 1 describes, the architectures that control AI token costs most effectively are those that route structured and repeatable work to deterministic execution rather than LLM reasoning. Orchestration is how that collage of human touchpoints, deterministic and non-deterministic execution is built. An organisation with strong orchestration infrastructure across its core processes is, almost by design, better positioned to control AI spend, because the high-volume work that drives token bills in ungoverned deployments is already running on workflow automation at near-zero marginal cost per execution. The two investments reinforce each other.

## The barrier between pilot and production

Integration development overhead is one of the primary reasons AI programs stall between pilot and production. A pilot that runs on clean sample data and a single connected system looks straightforward in a demonstration. Production requires connecting to the real systems of record, with their authentication requirements, their data schemas, their exception handling logic, and their own governance obligations. Building these connections from scratch for each AI deployment is expensive, slow, and creates a maintenance burden that compounds as the connected systems change over time.

IDC research, conducted in partnership with Lenovo, found that for every 33 enterprise AI proofs of concept a company launches, roughly four reach production deployment. Gartner states that around 60 percent of pilots stall. Token cost anxiety is one reason; the integration barrier is another. Organisations with pre-built, pre-governed connections to their enterprise application estate remove that second barrier before the AI program begins. The integration work is not eliminated, but it is not starting from zero on every deployment, which changes both the delivery timeline and the cost structure significantly.

This matters for the ROI calculation in a specific way. An AI program whose costs include significant bespoke integration development on each deployment is a program whose returns are eroded before the AI reasoning layer even runs. An AI program built on top of an existing orchestration foundation carries lower per-deployment costs and produces returns sooner, which changes the financial case for continued investment at every subsequent stage.

Gartner clients report up to 50% productivity improvement with iPaaS versus previous-generation integration technology; some vendors claim up to 80% productivity gains — a metric that can be used to approximate development/maintenance labor savings when scoped appropriately. And that iPaaS is a lever for all the connectivity your AI and AI orchestration requires.

## How the value is measured

The business value of AI-connected orchestration is measurable in operational terms, not in time-saving surveys or estimated productivity uplift per employee. The metrics that matter are end-to-end process cycle time, from request initiation to resolved outcome, across multi-step cross-system workflows; exception volumes, meaning how often processes require human intervention beyond the defined decision points; data integrity measures such as error rates at system boundaries and the frequency of retrospective correction; and integration delivery cost per new AI use case deployed.

These metrics appear in operational reviews and CFO reporting. They are auditable. They can be baselined before AI deployment and measured against after. They answer the question a finance function will always ask: compared to what we had before, and compared to what this cost, are we ahead?

The ROI case for a governed orchestration layer exceeds the sum of its component costs because of what the combination produces that none of the components can produce independently: An AI model that cannot access the systems of record produces text. An integration connector without AI reasoning automates a fixed, pre-defined process. The governed combination, sequenced across a structured workflow with appropriate human decision points, produces an outcome that replaces a manual process end-to-end, reduces error rates at every system boundary, satisfies governance requirements, and scales without a proportional increase in headcount. It enables observability across AI agents and enterprise systems, while enforcing security and compliance policies.That outcome is what justifies the investment. It is not available from any individual component in isolation.

### Deriving Business Value from Connectivity

The Orchestration layer replaces manual processes end-to-end, reduces error rates at every system boundary, satisfies governance requirements, and scales without a proportional increase in headcount. It enables observability across AI agents and enterprise systems, while enforcing security and compliance policies.

## The ground infrastructure argument

Return to the airport. The runway, the gate, the taxiway, the boarding bridge, the baggage belt, the fuel line: none of these is individually impressive. No airline purchases an aircraft fleet to use the boarding bridge. But without the boarding bridge, the passenger cannot get from the terminal to the aircraft. Without the taxiway, the aircraft cannot get from the gate to the runway. Without the baggage system, the luggage does not arrive where the passenger does.

The ground infrastructure of an airport is what makes the aircraft useful. The world's most advanced aircraft, sitting on a stand with no ground infrastructure connecting it to the terminal, the fuel supply, and the runway, produces no value at all. It is an extraordinary piece of engineering that cannot go anywhere. It’s just a plane on an empty plain.

Enterprise orchestration is the ground infrastructure for AI. The AI agents, the language models, the governance tools, and the human oversight mechanisms described throughout this paper all depend on it. They can each be excellent. Without the connectivity layer beneath them, their value is confined to the systems they can already reach. For most organisations, the systems they can already reach are a small fraction of the systems that matter.

The organisations that will realise durable, measurable returns from AI investment are not those that found the best model. They are those that built the connectivity infrastructure that makes any model useful across the full breadth of how their business actually operates.

### The value test

When assessing an AI program's business case, ask: does this investment include the orchestration layer that connects AI reasoning to the systems of record?

If the answer is no, the productivity numbers in the business case are overstated, because they assume the AI output becomes an outcome without accounting for the human coordination required to make that happen (often manual coding projects with technical debt and persistent maintenance issues).

The orchestration layer is not an optional component of AI transformation. It is the component that converts AI from a tool that helps people work faster into a system that produces outcomes independently.

### Notes and Sources

The following sources underpin the principal factual claims in this paper. Gartner research documents are available to subscribers via gartner.com. All other sources are publicly accessible at the URLs listed.

#### The Cost Paradigm — Sources


- **Uber burns through 2026 AI budget in four months** Fortune, May 2026: fortune.com/2026/05/26/uber-coo-ai-spending-tokens-claude-code/
- **Uber caps employee AI tool spending** Bloomberg, June 2026: bloomberg.com/news/articles/2026-06-02/uber-caps-usage-of-ai-tools-like-claude-code-to-cut-costs
- **Microsoft cancels internal Claude Code licences** Windows Central, 2026: windowscentral.com/microsoft/microsoft-cancels-claude-code-licenses-shifting-developers-to-github-copilot-cli-a-move-likely-driven-by-financial-motives
- **OpenClaw: $1.3 million in tokens in 30 days** The Next Web: thenextweb.com/news/openclaw-peter-steinberger-1-3-million-openai-token-bill | Tom's Hardware: tomshardware.com/tech-industry/artificial-intelligence/openclaw-creator-burns-through-1-3-million-in-openai-api-tokens-in-a-single-month
- **Goldman Sachs: AI token demand to multiply 24x by 2030** Goldman Sachs Insights: goldmansachs.com/insights/articles/ai-agents-forecast-to-boost-tech-cash-flow-as-usage-soars | Tom's Hardware summary: tomshardware.com/tech-industry/artificial-intelligence/ai-costs-begin-to-bite-as-agents-may-increase-token-demand-by-24-times-says-goldman-sachs-report
- **Gartner Technologies and Markets: ServiceNow's AI Control Tower: AI Governance Augmentation With a Complex TCO** Gartner (Subscription): March 2026 gartner.com/en/documents/G00845710
- **Gartner Delivery of Functional Responsibilities: Key Considerations for Adopting the ServiceNow AI Platform** Gartner (Subscription):10 September 2025 - gartner.com/en/documents/G00835375


#### MCP Architecture — Sources


- **Gartner Innovation Insight: MCP Gateways (G00839300), September 2025** Gartner (subscription): gartner.com/en/documents/6907866 | Secondary analysis: truefoundry.com/blog/truefoundry-and-the-mcp-gateway-revolution-insights-from-gartners-2025-report
- **Gartner Emerging Practices for MCP Servers and Tools** Gartner (subscription): gartner.com/en/documents/7233930

### The Auditability Paradigm — Sources

- **Gartner: 60% of organisations have evidence of unsanctioned AI agent activity** Gartner Best Practices to Counter MCP Security Risks (G00844301), February 2026 (subscription required)
- **Gartner: Top Cybersecurity Trends for 2026** Gartner newsroom: gartner.com/en/newsroom/press-releases/2026-02-05-gartner-identifies-the-top-cybersecurity-trends-for-2026
- **Asana MCP server data exposure incident, May–June 2025** BleepingComputer: bleepingcomputer.com/news/security/asana-warns-mcp-ai-feature-exposed-customer-data-to-other-orgs/ | Nudge Security analysis: nudgesecurity.com/post/asana-mcp-server-data-exposure-incident

### The Connectivity Paradigm — Sources

- **IDC research (with Lenovo): 88% of AI pilots do not reach production** CIO.com: cio.com/article/3850763/88-of-ai-pilots-fail-to-reach-production-but-thats-not-all-on-it.html

### Datadog IRAP Accreditation

- **Datadog achieves IRAP Protected status in Australia, October 2025** Datadog newsroom: datadoghq.com/about/latest-news/press-releases/datadog-achieves-irap-protected-status-in-australia/