COOi
Co Open Innovation
studios

The Agentic Organogram & The Human-Agent Nexus

Restructuring Accountability and Operating Models for Autonomous Artificial Intelligence Agents

Contents

Abstract

Abstract

Enterprises are deploying autonomous AI agents into operating models designed for human-only decision-making, yet only a third have moved agentic workflows past the pilot stage, and up to 95% of pilots fail to show measurable business impact. This is not, at heart, a technology problem; it is an operating model problem, where manual governance oversees software that plans and acts at breakneck speed. It is a bit like clearing a plane for takeoff on autopilot, while the control tower still works off paper notes instead of radar. The autopilot does exactly what it was built to do, but without radar, standard protocols, and a named controller watching, no one is positioned to catch it before it drifts off course.

Three real incidents illustrate the cost of this gap. In 2012, Knight Capital lost $440 million in 45 minutes because its trading software had no automated circuit breaker to stop it. In 2023, a cascading, unmanaged failure in Optus's automated network defences caused a 14-hour outage affecting over 10 million people. In 2019, Vale's monitoring instruments flagged anomalous readings two days before its Brumadinho dam collapsed, killing 272 people, because no protocol routed that warning to a human in time. In each case the system performed exactly as built; it was the governance around it that failed.
In this point of view (POV) we propose a framework built on a foundational operating model assessment, demonstrated in a COOi Studios led analysis for a Data Office within a Corporate and Investment Bank. From that baseline, three pillars follow: People, where a Dynamic RACI (Responsible, Accountable, Consulted, Informed) shifts named accountability as an agent's risk tier changes; Process, where standardised handoff protocols pass full context between agent and human; and Technology, where governance is enforced through code, not policy documents. Ultimately, a sustainable competitive advantage will belong to enterprises that build the capability to direct, coordinate, and audit machine intelligence alongside human personnel.

Introduction

Introduction

The agentic AI market is projected to reach $45 billion by 2030, up sharply from $8.5 billion in 2026, a clear signal of how quickly enterprises are moving to adopt software that can reason independently and execute tasks dynamically.[1] Yet this rapid growth in investment often fails to translate into real business value, due to structural execution gaps.[2]

The real danger lies in businesses rolling out autonomous AI agents inside operating models that were only ever designed for people to run. These older operating models were built around stability, step-by-step workflows, manual sign-offs, and predictable hand-offs between people, all managed from the top down. Thus, when AI agents are introduced into a system built purely for humans, the line between "software as a passive tool" and "staff as the ones making decisions" breaks down. That breakdown is where the operational friction begins.[3]

Research shows that although 62% of corporate leaders are actively experimenting with autonomous agents, fewer than 15% have deployable solutions in live production environments.[4] This value-realisation deficit confirms that the software itself is rarely the primary barrier. Instead, successful digital transformation is dictated by how well the organisation executes its operating model and re-engineers its workflows.[5]

Enterprises are still running manual control towers (quarterly committees, human sign-off chains, retrospective audits) to manage software that behaves like an autonomous aircraft, continuously sensing its environment, replanning mid-flight, and executing thousands of micro-decisions a second. When that aircraft drifts off its flight path, there is often no controller watching, no protocol for handing it back to a human pilot, and no one willing to own the outcome once it lands badly.

The Control Tower

Who is watching?
Who takes control?
What intervenes automatically?

01 Named accountability

02 Standardised handoffs

03 Automated intervention

Problem Statement: Defining the Liability Vacuum

Deploying autonomous agents without clear operational boundaries introduces significant financial, operational, and regulatory risks. Traditional operating models assume that discretionary judgment, operational risk, and regulatory accountability belong exclusively to humans. Conversely, autonomous agents make real-time operational decisions, change multi-step action plans dynamically, and interact directly with business systems and external markets.[6]

When an agent makes an error, the lack of defined accountability creates a persistent liability vacuum: an institutional condition where no designated human role or department inherits the legal or financial accountability for machine-driven business errors.[2]
Legacy governance frameworks assume that choices are made by accountable humans, thus non-deterministic software operates in an unmanaged gap.[7] In these liability vacuums, IT divisions reject ownership of business outcomes, while Business Units refuse liability for IT written code. This delays incident resolution, increases compliance exposure, and ultimately prevents the organisation from seeing a return on its AI investments.

The Cost of Getting it Wrong: Real-World Business Challenges

We have selected real-world case studies across the Financial Services, Telecommunications, and the Resources sectors, where the execution gap caused substantial business damage, are reviewed.

Speed-Governance Divergence in Financial Services

In 2012, Knight Capital Group deployed a new algorithmic trading programme. Due to a deployment error, a dormant piece of code was activated, and the autonomous system began buying and selling shares erratically at high speed. Because the trading algorithm operated independently of any dynamic risk-escalation framework or real-time human supervisor (an "Agent Boss"), it executed millions of unintended trades in just 45 minutes. By the time human operators identified the problem and physically shut the system down, the company had lost $440 million, bringing the firm to the brink of bankruptcy.[8] This is a Technology gap that highlights the danger of deploying autonomous execution entities without corresponding automated circuit breakers.

Escalation and Accountability Vacuum in Telecommunications

In In November 2023, Australian telecommunications provider Optus experienced a massive network outage affecting over 10 million people and thousands of businesses. The failure occurred after a routine software upgrade at an international internet exchange sent unexpected routing changes into the Optus network. In response to this sudden overload of data, approximately 90 autonomous edge routers automatically disconnected themselves as a pre-programmed protective measure. Because these automated routers lacked a contextual hand-off protocol to alert human engineers or gracefully escalate the issue before shutting down, the default automated response cascaded, bringing down the entire national network for up to 14 hours.[9] This demonstrates what happens when automated systems reach their limits without a structured escalation pathway.

Telecommunications

Optus

14 hours

10m+ people

Lack of Process Handoff in Resources

In January 2019, a tailings dam at Vale's Córrego do Feijão iron ore mine in Brazil collapsed, killing 272 people.[10] Two days before the collapse, Vale's own automated monitoring instruments, piezometers installed to track the dam's structural stability, had flagged discrepant readings, and five of those instruments were not even functioning.[10] The information existed inside the system, but no protocol connected that automated warning to a human decision-maker in time to act on it. In the aftermath, Vale and its external safety certifier spent years disputing which of them was responsible for missing the risk.[11] An automated system detected a problem, but there was no standard handoff to route that signal to a person with the authority to stop operations, signifying a significant process gap.

Resources

Vale

272

Lives lost

Three different sectors, three different systems, yet the same underlying pattern: an autonomous or automated process was given a broad mandate and allowed to operate with no circuit breaker, no named human owner, and no protocol for handing control back. Nobody in the metaphoric air control tower knew there was a problem until the damage was already done.

Operating Model Framework

The Operating Model Framework for Agentic AI

Closing this gap does not mean grounding every autonomous agent and reverting to manual control. Modern airspace safely manages thousands of semi-autonomous aircraft at once, but only because it is built around harmonising the three pillars that every operating model rests upon: named accountability (people), standard protocols (process), and automated systems that catch problems before a human even needs to step in (technology). The same three pillars apply directly to agentic AI, and each one addresses the gap exposed by one of the incidents above. Before any of these pillars can be built, however, an organisation first has to know where it currently stands.

Foundational Assessment: Benchmarking the Operating Model

Before designing a new target operating model (TOM) for AI agents, a business must evaluate its current operational maturity to remove ambiguity. Skipping this step is the equivalent of trying to build a control tower's protocols without first surveying the airspace it is meant to govern: no controller can beassigned, no handoff standardised, and no automated safeguard configured correctly if nobody agrees on the current state of operations.

A practical example of this was recently seen within the Data Office of a Corporate and Investment Bank (CIB). The division lacked a standardised baseline to measure its data management operational maturity, which created ambiguity around existing gaps and stalled the design of an optimised TOM. To solve this, COOi Studios executed a comprehensive operating model maturity assessment, benchmarking its current state against industry-best practice using the Data Management Capability Assessment Model (DCAM) as the underlying benchmark.

The assessment evaluated the current state across the same three pillars this framework is built on:

Technology: the effectiveness of supporting systems and tools.

Process: current-state workflows and daily operating practices.

People: organisational design, roles, and skills.

The value of this exercise was not the benchmark itself, but what it produced: a clear, evidence-based roadmap for building the target operating model. Once that baseline is established, an organisation is in a position to apply the framework below to the specific challenge of integrating AI agents.

Pillar 1

People and Structural Accountability

Every aircraft in controlled airspace has a designated controller responsible for it at any given moment. Enterprises need the same discipline. Every AI agent cohort should be assigned an Ultimate AI Accountability Owner (UAAO): a senior person who carries legal, ethical, and regulatory responsibility for that agent throughout its lifecycle, supported by operational leads who track its day-to-day performance.

Ultimate AI Accountability Owner
Legal, ethical and regulatory responsibility for the agent lifecycle

Dynamic RACI
Roles shift automatically as risk tier confidence level or context changes

Risk-Based Escalation
From agent execution > human informed > human accountable > human approval

Real - World Lesson
The Optus outage showed what happens without a dynamic escalation path

This is precisely the mechanism that was missing at Optus: the autonomous routers had a pre-programmed response, but no dynamic escalation path that shifted accountability to a human engineer before the default action cascaded into a national outage.

Pillar 2

Process and Inter-Agent Orchestration Protocols

Unregulated and unstandardised workflows between multiple AI agents can lead to cascading errors and “delegation creep,” where agents take on tasks beyond their original mandate.[12] Standardised protocols are required to enforce safe coordination.

Contextual Handoffs
Ensure no information is lost when a task moves between agents and huans

Agent Boss
Monitors the workflow, resolves conflicts, handles exceptions and intervenes when needed

Executive Checkpoints
Halts high-risk actions for formal human approval with cryptographic sign-off

As-Is Process Mapping
The foundation for agentic information

The Agentic Opportunity Map
Identifies where to integrate, automate or redesign processes

That mapped, current-state baseline is precisely the raw material process reengineering needs: an organisation cannot redesign a workflow for agentic AI, decide where an agent should sit in it, or where a human handoff needs to be inserted, until the existing process has been documented and categorised well enough to redesign in the first place. As-Is mapping is therefore the foundation on which any process redesign for agentic inclusion must be built, not a parallel exercise to it.

Pillar 3

Technology and Deterministic Runtime Governance

Manual oversight of data governance and ethics policies is prone to human error, creates siloed text-based documents, and cannot scale with cloud infrastructure. This manual approach results in inconsistent enforcement of data glossaries, incorrect classifications, and high regulatory risk.[14] Organisational policy must be translated into hard, deterministic technical constraints.

Policy as Code
Use open policy agent (OPA) and rego with AWS to enforce governance automatically

Versioned Governance
Data glossaries and metadata models in machine-readable formats (JSON/YAML Schema), managed through Git workflows

Automated Enforcement
Checks run before deplaoyment or storage, bloacking violations and logging audit trail

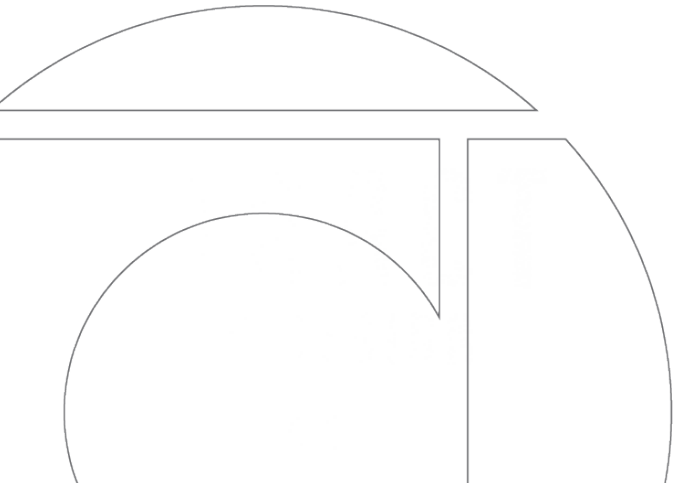
Continuous Compliance
Embedded in the pipeline, not an afterthought

SRE Circuit Breakers
Detect > Isolate > Stop > Escalate > Fail safety by design

Continuous Compliance and Glossary Automation embeds data governance checkpoints directly into deployment pipelines. For example, triggering a check on an AWS S3 document upload will extract the text and metadata into JSON format, run OPA compliance and privacy checks, and immediately block violations while logging a full audit trail. Deployments are automatically blocked if infrastructure configurations violate ethical baselines.
Alongside Policy-as-Code, systems require hard ceilings on resources, budget limits, and API usage. Automated Site Reliability Engineering (SRE) Circuit Breakers must monitor the agent's performance. If the agent's error rate spikes or it breaks safety rules, the circuit breaker instantly isolates the agent, cuts network access, and alerts the human Agent Boss.[14],[15]

Bringing the pillars together

The pillars work well if they operate together, and only once the foundational assessment has established what "normal" looks like. Named accountability with no automated monitoring still will not catch a fast-moving deviation in time. Automated monitoring with no standard handoff protocol will detect a problem correctly but leave no clear route back to a human decision-maker. And neither of those matters if the underlying processes were never mapped and centralised in the first place. The strength of this approach is that it closes the gap from all three sides at once, on a baseline the organisation actually understands.



Conclusion

Conclusion

None of the incidents examined in this paper happened because the underlying technology was defective in some exotic, unforeseeable way. Knight Capital's software worked exactly as coded; it simply had no governance protocol to prevent it from spiralling. Optus's edge routers executed their pre-programmed protective response exactly as designed; they simply had no dynamic escalation path to a human before that response cascaded into a national outage. Vale's monitoring instruments produced exactly the warning they were built to produce; there simply was no procedure guaranteeing a human would see it in time. In each case, the aircraft flew fine. It was the airspace around it that failed.

That is the core argument of this paper: the execution gap enterprises are experiencing with agentic AI is not an algorithmic problem. It is an operating model problem. An organisation cannot assign named accountability, design a dynamic RACI, standardise a handoff protocol, or codify automated enforcement for a landscape it has never actually mapped. The foundational assessment has to come first as it establishes an honest, benchmarked baseline of where people, process, and technology currently stand, so that the three pillars can be built on solid ground rather than assumption.

From that baseline, the work is threefold: naming accountability that shifts dynamically as an agent's risk changes, rather than sitting fixed on a document nobody revisits; reengineering processes for agentic inclusion using a mapped "As-Is" state as the foundation; and codifying governance into deterministic, automated checks, rather than trusting policy documents nobody really reads.

For executive leadership, the mandate is clear. The competitive advantage in the AI era will not belong to the organisations that buy and deploy autonomous models the fastest. It will belong to those that assess their current maturity baselines, establish total operational transparency through rigorous process engineering, and subsequently re-engineer the institutional architecture required to direct, coordinate, and audit machine intelligence.

By unifying human accountability, deterministic process orchestration, and technical controls like AWS Policy as Code, the Agentic Organogram Framework effectively closes the execution gap. It allows the modern enterprise to achieve machine-speed operational execution while maintaining robust, verifiable, and strictly human-anchored control.

Reference List

1. World Economic Forum (WEF). (2026). How Agentic, Physical and Sovereign AI Are Rewriting the Rules of Enterprise Innovation. World Economic Forum. Available at: https://www.weforum.org/stories/artificial-intelligence/how-agentic-physical-and-sovereign-ai-are-rewriting-the-rules-of-enterprise-innovation/
2. Bain & Company. (2026). How to Architect for Agentic AI: The Three-Layer Operating Model. Bain Insights. Available at: https://www.bain.com/insights/
3. Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., & Buckley, N. (2018). Coming of Age Digitally: Learning, Leadership, and Legacy. MIT Sloan Management Review and Deloitte Insights. Available at: https://sloanreview.mit.edu/
4. McKinsey & Company. (2025). The State of AI in 2025: Agents, Innovation, and Transformation. McKinsey Digital. Available at: https://www.mckinsey.com/
5. McKinsey & Company / Artic Sledge. (2025). Enterprise AI Strategy: The Printable Value–Risk Portfolio Builder. Available at: https://www.mckinsey.com/
6. Delinea Research / ResearchGate. (2026). Non-Human Identity Governance: A Governance Framework for Agentic AI Credentials in Regulated Enterprises. Available at: https://www.researchgate.net/
7. World Economic Forum & Dataiku. (2026). Your AI Governance Framework Was Built for a World That No Longer Exists. WEF Thought Leadership & Dataiku. Available at: https://www.weforum.org/
8. Securities and Exchange Commission (SEC). (2013). SEC Charges Knight Capital With Violations of Market Access Rule. SEC Press Release. Available at: https://www.sec.gov/
9. Australian Communications and Media Authority (ACMA). (2024). Optus Network Outage Investigation. ACMA Publications. Available at: https://www.acma.gov.au/
10. Robertson, P. K., de Melo, L., Williams, D., & Wilson, G. W. (2019). Report on the Technical Causes of the Failure of Feijão Dam I. Vale S.A. Expert Panel. Available at: http://www.b1technicalinvestigation.com/ and https://damfailures.org/case-study/feijao-dam-b-1-brazil-2019
11. TÜV SÜD AG / Wall Street Journal. (2019). Brumadinho Tailings Dam Certification and Liability Dispute. Available at: https://www.business-humanrights.org/en/latest-news/t%C3%BCv-s%C3%BCd-vale-criminal-lawsuit-re-brumadinho-dam-collapse-filed-in-brazil/ and https://verfassungsblog.de/brumadinho-dam-responsibility /
12. Bain & Company. (2026). Agentic AI Governance, Risk, and Controls for Business Leaders. Bain Insights. Available at: https://www.bain.com/insights/
13. Gartner. (2024). Hyperautomation. Gartner Information Technology Glossary. Available at: https://www.gartner.com/en/information-technology/glossary/hyperautomation
14. Forrester Research. (2026). AEGIS Framework: Agentic AI Enterprise Guardrails For Information Security. Forrester Technology Insights. Available at: https://www.forrester.com/
15. Microsoft Security Community. (2026). MCP Safety & Evaluation with the Agent Governance Toolkit (AGT). Microsoft Community Blog. Available at: https://techcommunity.microsoft.com/