Datatonic

From flat-pack to factory: Scaling enterprise agents with platform-based AIOps

DATATONIC | WHITEPAPER | 2026

Datatonic

Table of contents

03 The flat-pack approach to building AI agents doesn't cut it anymore

05 Creating an agent factory with platform-based AIOps

07 3 ways the factory approach drives value
+ Accelerating agent deployment at scale
+ Security at scale
+ Evaluation and monitoring at scale

13 The factory approach in action: 3 case studies

16 Final thoughts: Factories can keep up, DIY can't

Chapter 1: The flat-pack approach to building AI agents doesn't cut it anymore

6 to 12 months ago, GenAI and agentic use cases were predominately experimentation around tactical problems

Use cases focused on internal automation and efficiency, with AI bringing together disparate data sources accessible with natural language. We're talking about back office, HR and financial planning, with invoice and contract management, enterprise search, and call transcriptions.
On the customer-facing side, use cases have focuse d on self-service agents and personalisation.

Across these (often shadow) POCs, there's a common denominator: the aim is to boost productivity, with 'hours saved' being the key metric – even if no one is actually measuring it.

Now, enterprises are moving towards full process transformation

You don't have to spend long online without hearing about how sectors face fundamental business model changes from AI. And that means agents are increasingly part of wider transformation projects that are reimagining end-to-end processes in areas like underwriting, creative campaign development, and support function management. Instead of bolting on to existing workflows, organisations are going back to basics, mapping out process steps and identifying where AI can augment and transform with humans in the loop, with multiple agents operating on different tasks. At the sharp end of this approach, we're seeing organisations leveraging AI agents able to automate 60-70% of tasks in a workflow.

But going from tactical, productivity-focused use cases to workflow transformation isn't an easy transition. Where organisations get stuck when they start scaling agents

Agents are easy to build individually but hard to scale – it's the difference between assembling flat-pack and a factory assembly line. Unless you have a factory approach, you come up against key blockers:



BLOCKER 1

You can't deliver fast enough

The business wants everything done now, but there are no clear blueprints for deployment, and UAT/evaluation can't move that fast – particularly with customer-facing agents.

Large backlogs build up, but the organisation lacks the in-house skills to confidently deploy agents into production, particularly for customer-facing agents in sensitive or regulated environments. It's one thing for a team of 2 people to maintain 3 agents; when you get to 50-70 use cases, it becomes impossible. The speed of tech evolution exacerbates this problem.

BLOCKER 2

Security and governance become blockers

Managing security and mitigating risks of jailbreaking and prompt injection become challenging as the agent ecosystem scales.

A use-case centric view of deploying AI agents risks a situation where there's no structure or accountability for broader guardrails. This is exacerbated by vendors pushing low- and no-code solutions, as well as their in-built AI functions.

BLOCKER 3

Agents don't perform reliably or deliver clear ROI

People already lack trust in GenAI outputs, and output consistency becomes a greater challenge when you introduce agents. When context grows, agents become less accurate and can become distracted, so it's hard to control as the organisation tries to scale.

You can't measure agent impact at scale, which means you can't attribute value – and therefore create the business case for investment. For example, it's hard to quantify how much time people spend using an agent, and even if you can calculate this individually, it's hard to roll it up to financial or operational metrics.

> *Agents are easy to build individually but hard to scale – it's the difference between assembling flat-pack and a factory assembly line.*

Chapter 2: Creating an agent factory with platform-based AIOps

To overcome these blockers, you need to move from a flat-pack approach where you build agents individually to a factory approach that meets enterprise demands.

That's what you get with **platform-based AIOps**, a best-practice approach for turning bespoke AI experiments into a high-throughput, well-governed production line.

There are 2 elements of platform-based AIOps:

Factory infrastructure
Platform that's an extension of internal developer platforms.

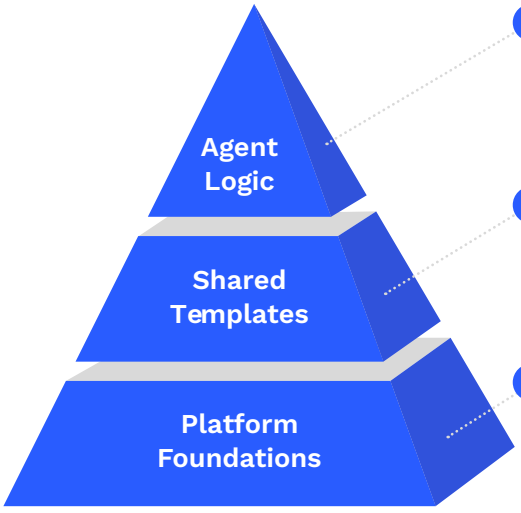
Factory mindset
The operating model that gives you control and reduces friction as you scale.

> *Platform-based AIOps is the factory approach you need to scale enterprise agents. It turns bespoke AI experiments into a high-throughput, well-governed production line that works.*

The factory infrastructure you need

What a high-performance AIOps platform looks like

Your AIOps platform should have a layered agent architecture that decouples shared enterprise foundations from unique agent logic.



	DESCRIPTION	EXAMPLES
3	**Custom Agent Logic** The unique configurations and specific reasoning rules that the agent needs in order to solve a distinct problem.	+ Tailored system prompts and personas + Orchestration workflows + Novel agent architectures
2	**Shared Templates** Modular, pre-built components that are standardised and reused across groups of similar agents.	+ RAG pipelines + Guardrails + Evaluation + Agent tools (e.g. connectors) + Skills
1	**Platform Foundations** The standardised, non-negotiable infrastructure, services and templates required to deploy, secure, and monitor all agents at scale.	+ Infrastructure as Code + CI/CD Pipelines + Evaluation & Observability components + Identity & Access + Platform guardrails (e.g. PII obfuscation, jailbreaking etc.) + MCP & A2A – Layer & Gateway

You need all 3 architectural layers to achieve speed, security, and performance at scale:

+ Standardising the foundational 80% means you can focus engineering efforts on customisation and fine-tuning. This approach also ensures scalability based on query volumes in production, meaning your underlying infrastructure always meets your needs and LLMs are rate-limited correctly.

+ The middle layer of shared technical clusters accelerates agent deployment on top of the 'standard' platform foundations. Creating these clusters generally requires expertise beyond traditional machine learning that's harder to DIY – particularly when it comes to defining patterns and templates, ensuring the correct security posture, and codifying effective evaluation and monitoring.

+ Agent logic is then easier to define and deploy because you have much less customisation for each use case.

3 benefits of platform-based AIOps

With its combination of factory infrastructure and factory mindset, platform-based AIOps helps you address the common blockers in Chapter 1 and achieve 3 key requirements for scaling enterprise agents:



Accelerating deployment:
Allowing agents to flow quickly from experimentation to productionisation while reducing duplicated effort by standardising tools, agent skills, data pipelines, and system access.



Improving security:
Balancing innovation with control. As agents work across solutions and data sources at scale, security can be enforced in many ways, including AI landing zones and a responsible AI layer/model armour.



Enabling continuous evaluation and monitoring:
Of agents in production, with more visibility on performance and a stronger link to metrics that matter.

Defining what success looks like

A blocker to scaling enterprise agents is difficulty measuring their impact. How can you build a business case without understanding the value? And to understand the value, you need something more meaningful than 'greater productivity'.

With a factory approach, the aim is to give each agent a business case that rolls up to financial and operational KPIs, with the ability to measure progress in real-time. This can be done by:

+ Creating a value framework: That incorporates both top-down and bottom-up KPIs.

+ Engaging stakeholders: To help them understand the process(s) you're optimising, quantify inefficiencies, and align with the relevant KPIs in the framework.

+ Building a trustworthy impact dashboard: That tracks the cost of running each agent and the KPIs it influences.

Chapter 3: Accelerating agent deployment at scale

Platform-based AIOps gives you the ability to build and deploy agents in a structured, productised way. There are 2 main ways it accelerates deployment at scale: by providing a marketplace of agent templates and by using agents to streamline governance cycles.

Agent marketplace

An agent marketplace reduces duplicated effort through templates that standardise tools, agent skills, data pipelines, and system access.

You essentially have a library with prescribed agentic patterns and how they apply to use cases, for example:

+ When do you use agents, subagents, and agents as tools?
+ When do you use agentic tools and how do you share those tools?
+ When do you use agentic skills and how do you share those skills?
+ When do you apply pre-validated technical clusters, e.g. for patterns like intent routing, looping, and ReAct?
+ When should you consider agent-to-agent?

Having these prescribed patterns is a fundamental benefit because MCP servers and agent 'cards' can be deployed with governance, security, and cost built in. Crucially, the templates come with 'pre-approved' data sources, with IAM baked in. You therefore provide sanctioned access to enterprise data, preventing access issues from slowing agent development. People across the organisation – and even third-party suppliers – can then see and reuse these agent templates.

We have extensive experience building agent marketplaces for enterprises, including in regulated industries. We've delivered more than 300 different agent templates and have a deep understanding of use case archetypes, standards, and context management – ensuring agents are as context-efficient as possible. Importantly, we design templates for maintainability as well as speed, and ensure you have a future-proof approach that works for MCP and can scale up for Agent-to-Agent Protocol when needed (which is key given how quickly this area is moving).
Our ultimate aim is to modernise the end-to-end SDLC with AI, combining coding agents with agentic and human skills to build even faster.

COMMON AGENT ARCHETYPES	CORE PATTERN	EXAMPLE USE CASES
Conversational Q&A	Conversational Q&A	Enterprise search, back-office support
Data extraction and insights	Conversational analytics across structured and unstructured data sources	Claims processing, invoice management
Commercial content generation	Multi-agent orchestration	Contracts, proposals, marketing campaign content
Technical content generation	Multi-agent orchestration	Augmenting the internal software lifecycle
Workflow planning and automation	Business process re-imagination, human-in-the-loop, agentic automation	Underwriting, network management

> *We've delivered 300+ agent templates and have a deep understanding of use case archetypes, standards, and context management.*

Agentic governance

Governance is the second biggest agent deployment bottleneck – and although it's critical, it shouldn't put the brakes on innovation the way it does. In regulated industries, the average deployment time from ideation to production is 6 to 12 months. The standard governance lifecycle includes 13 review steps that cut across at least 3 teams and more than 10 people, resulting in more than 40 hours of effort. This is too cumbersome and too slow.

The factory operating model accelerates and streamlines this process with an agentic governance approach. There are 2 elements to this. First, having an agent marketplace with established templates means you have **pre-approved use cases**, substantially reducing review requirements.

Second, a **suite of agents manages compliance and governance steps** – guiding each use case through the lifecycle from experimentation to monitoring with minimal human input. Importantly, this enables CI/CD for LLMs, so you can have automated pipelines moving agents from sandbox to production in days rather than weeks and months.

STANDARD GOVERNANCE CYCLE

Includes reviews from 3+ teams and 10+ people, taking 40+ hours

1. Define and ideate a use case
2. Present the use case in portfolio board
3. Obtain data and privacy approval
4. Estimate budget with cost calculators
5. Fill out use case catalogue form
6. Access and configure dev environment
7. Access data with service accounts
8. Kick off experimentation
9. Follow AI risk assessment
10. Pen test and get architecture sign-off
11. Deploy to production
12. Get final governance checks and approval
13. Review monitoring and compliance

FACTORY APPROACH WITH AGENTIC GOVERNANCE

Suite of agents that manage compliance and governance throughout the lifecycle

- Use case approval agent
- Use case synergy agent
- Policy-as-code
- Compliance review agent
- Use case documentation

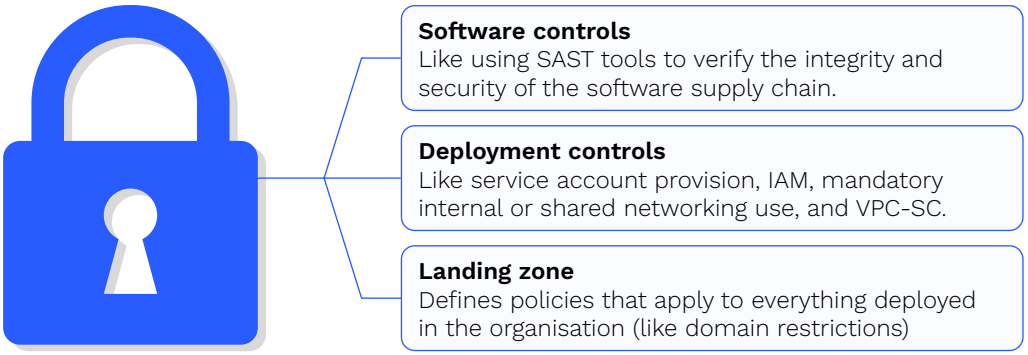
Using standard governance, average agent deployment time is 6 to 12 months in regulated industries. This is too slow.

Chapter 4: Security at scale

Expanding agent use requires a robust approach to managing security and mitigating risks, including around jailbreaking, prompt injection, personal identifiable information (PII) detection, and masking (both pre- and post-agent). With platform-based AIOps, controls are baked into the platform foundations and shared technical clusters – meaning you have the right levels from the start, not as an afterthought.

Layering security, with automation

We address security in layers through automation, ensuring each agent deployed meets robust requirements:



Additional controls that help you scale securely

With a factory approach, you can address more complex enterprise security requirements at scale. These include:

+ Data sovereignty: The architecture can be designed so multiple AIOps platform instances can be deployed into multiple physical environments to meet specific requirements, supporting compliance with regulations like GDPR, NIS2, DORA, and the EU AI Act.

+ Privacy detection and PII protection: LLM inputs and outputs are scanned for sensitive data and appropriate sanitisation is applied. This includes testing to validate effectiveness.

+ MCP Gateway: This provides a single management pane for agents to access all MCP sources, including external SaaS. This simplifies integration, unifies discovery and centralises governance.

+ Agent identities: You can easily assign unique, verifiable digital identities to autonomous AI agents, enabling them to authenticate, hold permissions and act securely (similar to human employee accounts). You therefore increase accountability and transparency in automated tasks.

Chapter 5: Evaluation and monitoring at scale

You need the ability to evaluate, test, and monitor agents to be able to track ROI, maintain agent efficiency, ensure consistency, and build trust. But when you're scaling at enterprise level, you can't spend hours in UAT for each use case. You therefore need continuous evaluation and monitoring that automates these processes while maintaining the right human input.

Continuous evaluation framework and performance monitoring

Evaluation is multifaceted, so you need a framework that sets out the appropriate performance criteria.

This involves answering questions like:

+ How do we choose the right LLM?
+ How do we evaluate multi-modal models?
+ How do we measure performance, ensuring outputs are of the right quality?
+ How do we ensure performance is maintained over time (regression testing, mitigating risks of semantic drift)?
+ How do we monitor and maintain performance when models upgrade?

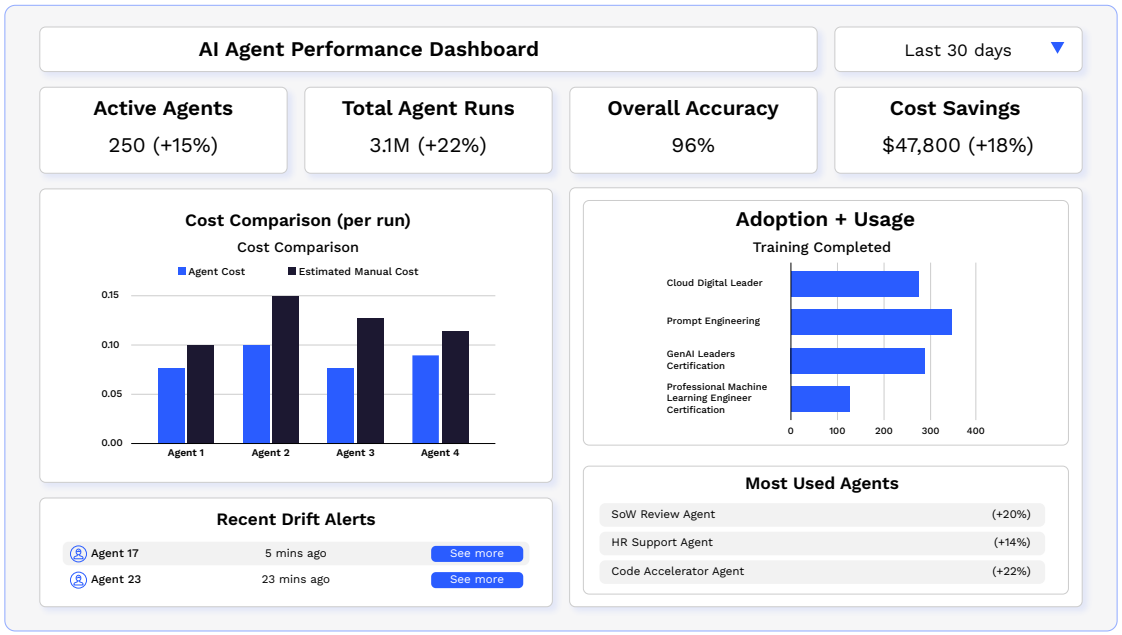
We offer unique IP here, with our Datatonic Evaluator Accelerator

This proprietary framework is designed to streamline the evaluation of LLM agents by integrating Phoenix tracing with DeepEval metrics. It provides the scaffolding to load datasets, generate synthetic tests, and track performance and safety across different project runs. We use Google stack tools to observe key metrics like latency (including between LLM calls) and the impact of guardrails. By default, key logs and metrics are included in provisioning.

Tracking agent performance with an impact dashboard

Through the AIOps platform, you can measure all agents against defined success criteria, including a FinOps element that gives cost visibility for each use case and relates it back to value delivered.

Here's an example dashboard:



Human-in-the-loop integration

Not everything should be automated, which is why you need an efficient way to integrate human-in-the-loop (HITL). This involves defining the escalation logic and metric thresholds for activating human involvement. After all, you don't want people to get pinged 100x a day, but you do want them to see relevant alerts and be able to evaluate. You also need to manage edge cases that need to bypass agents and go directly to a person, for example where it involves complex outcomes or a large business impact.

Alerting can be integrated into your existing tools or provided through new custom dashboards.

Managing model upgrades to prevent 'silent breakages'

LLMs evolve and deprecate quickly, meaning you don't get extended support the way you traditionally get with software. You may only have a couple of months to upgrade when a new version is announced, and agent performance often degrades afterwards. To scale agents, you therefore need a structured way to test post-upgrade effectiveness and ensure your current set-up works with the updated LLM.

This involves:

- **Evaluating LLM variants:** For example, if you're upgrading from Gemini 2.5 Pro, should you move to 3.0 Flash or 3.0 Pro?
- **Checking prompts still work:** A/B testing outputs with the same prompt.
- **Optimising prompts:** For the new model version.
- **Balancing cost, performance, and latency:** Based on agent task requirements, if you have a choice of model version.

> *Our proprietary Datatonic Evaluator Accelerator has templates wrapped in an (agentic) CLI, so agents build evaluation criteria on the platform itself.*

Chapter 6: The factory approach in action

So how does platform-based AIOps work? These case studies illustrate the approach and benefits for organisations starting with different objectives. Importantly, all are larger enterprises: other partners have built small AIOps platforms, but our USP is scaling agents in complex environments, even for the complicated requirements you see in creative and regulated industries.

MARKETING AGENCY

“AI is transforming our market; we need to be ahead of the curve”

CHALLENGE

“Our entire business model needs to shift because AI is changing the dynamics of customer demand and our cost base. How do we make it work *for us* rather than *against us*?”

SOLUTION

Datatonic delivered AIOps through:

Discovery:

+ Workshops to understand existing processes and where they could be streamlined and automated.
+ Process mapping to understand opportunities for optimisation and AI injection.
+ Redefining workflows and augmenting them with agents.

Implementation:

+ Secure AIOps platform designed to deploy agents at scale, integrating with existing systems, including the organisation’s ERP.
+ Defining agent archetypes with reusable patterns and in-built evaluation – market research, ideation and brainstorming, campaign strategy, content and creative, campaign management, and reporting.
+ Agent marketplace allowing users to discover and use the agents at scale.

Enablement: Training delivered through our award-winning Datatonic Academy, which empowers teams to build and use agents to maximum effect.

IMPACT

+ **Differentiation:** From robust, provable AI-enhanced processes that stand out in proposals and through enhanced client service.
+ **Increased profitability:** For service delivery.

VODAFONE

“Help us tap into new revenue streams”

CHALLENGE

“How can we leverage AI agents to help the business tap into next-gen revenue streams – at speed, with appropriate controls?”

SOLUTION

Datatonic developed Vodafone’s central AIOps platform, a first-to-market solution to accelerate AI product development.

+ **Discovery:** Designed the platform based on user story research.
+ **Implementation:** Built templates and security measures for a best-in-class AIOps platform with an agent marketplace.
+ **Expand:** Built new products for revenue growth, including a proposal assistant, marketing content creator, and network manager.
+ **Enablement:** Training and onboarding of the Vodafone team through ‘train-the-trainer’ sessions for adoption champions.

IMPACT

+ **Broad adoption across business units:** The AIOps platform is used across B2B, Commercial, HR, Customer Care, and Supply Chain Management.
+ **Supporting AI use at scale:** The platform has 600+ ML models in production, supports 3,000 data scientists, and is used in 18 markets.
+ **Accelerated time to deployment:** Time to deploy AI products reduced from 3 months to 2 weeks.

INSURANCE SECTOR

“Our KPI is to deploy 5 agents in H1”

CHALLENGE

“The business is putting pressure on us to inject AI into processes and deploy agents fast. We don’t know where to focus to meet our KPI, and don’t have the capacity or skills to deploy that fast.”

SOLUTION

+ **Discovery:** Took a step back and pivoted to an ROI-focused approach, supporting with business value mapping to understand the best opportunities for process optimisation and automation.

+ **Roadmap:** Based on the discovery, complaints and claims were prioritised because they involve large volumes of structured and unstructured data that currently requires human input for decisions.

+ **Implementation:** Built an AIOps platform with an agent marketplace for priority use cases, with HITL thresholds built in.

IMPACT

+ **Deployment bottlenecks removed:** The AIOps platform with templates means creating 5 agents is easy.

+ **Robust evaluation and HITL built in:** This is key given use cases are customer-facing. Agent performance is linked to business KPIs like time to resolution and number of claims resolved instead of focusing on an arbitrary number of agents as the KPI.

+ **Scalable and future-proof:** The agent marketplace is easy to expand with the next use cases in the roadmap.

Chapter 7: Final thoughts: Factories can keep up, DIY can't

Whether your C-suite is driving transformation, you've identified strategic opportunities for agentic AI to support new revenue streams, or the business simply wants to drive greater agentic adoption, there are similar challenges around speed, security, and performance.

To overcome those challenges, you need a factory approach

Platform-based AIOps gives you the infrastructure and operating model for deploying agents quickly, with the right controls, and in a way that tracks back to metrics that matter.

Importantly, platform-based AIOps has a low barrier to entry

Our experience and IP make it easy to get started. Although garbage in/garbage out holds true for any AI deployment, platform-based AIOps doesn't require the same degree of data structure, cleanliness and alignment as you need for MLOps. The focus is more on reducing risk of hallucination by ensuring a well-governed data landscape. You then provide sanctioned access to enterprise data because the agent templates available through the platform come with access to pre-approved data sources and in-built identity and access management controls.

As you move from flat-pack agent assembly to a high-throughput agent factory, keep these 3 considerations in mind



To evaluate agent ROI, you need to know what success looks like: Your project should start by defining a value framework so you know what business and operational KPIs are relevant to the processes where agents are being introduced.



Team enablement is essential: Ultimately, agents deliver value when they're used, and that requires building out AI skills and culture. Our award-winning Datatonic Academy can help here, with instructor-led courses that can be delivered virtually or in person. Also consider adoption champions and how 'train the trainer' sessions can be integrated into your implementation plan.



Future-proof your AIOps platform architecture: It should be built in a way that allows for reusability in new agentic AI use cases as well as non-agentic ones. Implementing the right foundations, security, and governance will pay dividends as your broader data and AI capabilities mature.

AI isn't slowing down – and neither are the opportunities it creates. To capitalise on them, you need the right infrastructure and operating model as your foundation: one that works today and is ready for whatever's next.

> *Implementing the right foundations, security, governance, and processes will pay dividends if you move to a unified data and AI platform.*

Datatonic



Go from friction to agentic AI at scale

Get in touch to discuss how we can help you transition to platform-based AIOps – helping you create a fast, secure agent production line that works. Learn more with our AIOps Discovery Workshop.

Request a workshop →

© 2026 Datatonic | All Rights Reserved