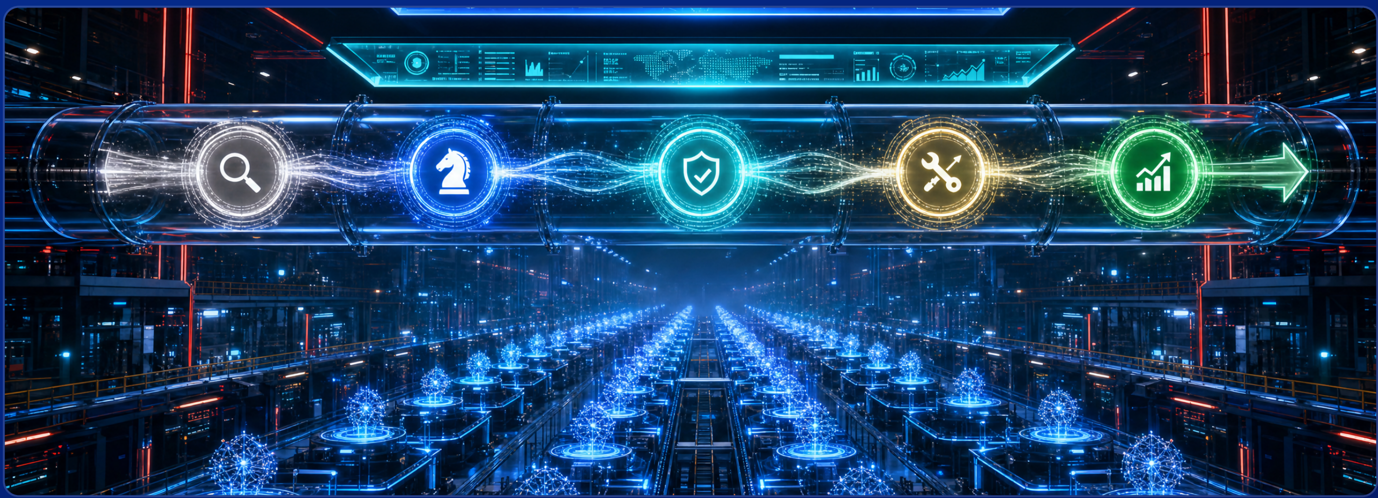
EIGENVECTOR
RESEARCH

EIGENVECTOR RESEARCH WHITEPAPER SERIES · ENTERPRISE AI AUTOMATION

The Agentification Factory

Organisational Design for Enterprise AI at Scale — A Systematic, Repeatable Architecture for Deploying Governed Agentic AI Across the Enterprise



Analysis of 177 enterprise agentic AI deployments reveals that ad-hoc, project-based deployment models suffer from high failure rates (69%), extended time-to-value, and unsustainable governance overhead. The Agentification Factory model replaces this with a systematic, repeatable production pipeline — achieving a 65% success rate, 40–60% faster time-to-value, and a verified net ROI of 21% through the compounding ROI Flywheel effect.

65%	**2.1×**	**177**	**42%**
Factory Model Success Rate	Vendor ROI Overstatement Factor	Documented Deployments	Marginal Cost Reduction

Marco H.K. van Hurne
Eigenvector Research · Inholland University of Applied Sciences
marco.vanhurne@eigenvector.eu

April 2026 · Version 1.0
eigenvector.eu

About This Whitepaper

This whitepaper comes from the field, not from the library. It is the product of years of running RPA and agentification factories at scale — in real organisations, with real data that was in the wrong format, real politics that nobody put in the project plan, and real CTOs who had just forwarded a LinkedIn post about an agent that could apparently do everything. The frameworks and evidence in this document were built, tested, and refined in production environments that actually had to pay off.

The research behind it is grounded in an analysis of **177 documented enterprise agentic AI deployments across 20 sectors** between 2022 and 2026. Those deployments are the evidence base. They are where the failure modes were identified, where the success patterns were validated, and where the 74% predictive accuracy of the PASF framework was confirmed. This is not anecdotal — it is the most comprehensive empirical dataset on enterprise agentic AI deployment that currently exists in the public domain.

The frameworks referenced throughout this whitepaper — the **Process Automation Suitability Framework** (PASF), the **Process Automation Design Engine** (PADE), the **Governed Runtime for Agentic Functions** (GRAF), the **Ontological Compliance Gateway** (OCG), and the **Roundtrip Value Governance** framework — are each available as standalone research papers at eigenvector.eu/research. This whitepaper brings them together into a single operational model: the Agentification Factory.

> **THE CORE INSIGHT**
>
> The root cause of the enterprise AI scaling challenge is **organisational rather than technological**. Most enterprises approach agentic AI deployment through an ad-hoc, project-based model that treats each deployment as a unique engineering challenge, resulting in duplicated effort, inconsistent governance, and a failure to accumulate institutional knowledge. The Agentification Factory replaces this with an industrial production model — standardised, repeatable, and governed by design. A single successful deployment is a project. A hundred deployments sharing tooling, governance, agentic patterns and institutional learning is a factory. The difference is compounding.

Table of Contents

CHAPTER 1

The Scaling Challenge in Enterprise Agentic AI

40% of agentic AI projects initiated in 2025 will be cancelled before 2027, primarily due to governance failures and the inability to scale beyond initial pilots. The root cause is organisational, not technological — and the solution requires a fundamentally different deployment model.

1.1 The Inflection Point

The enterprise adoption of agentic AI — autonomous systems capable of multi-step reasoning, tool use, and goal-directed action — has reached a critical inflection point. While early proofs-of-concept have demonstrated the technical feasibility of autonomous agents in constrained environments, the transition to enterprise-scale deployment has proven remarkably difficult. Gartner (2025) projects that 40% of agentic AI projects initiated in 2025 will be cancelled before 2027, primarily due to governance failures and the inability to scale beyond initial pilots.

This projection is consistent with the empirical findings underlying this research. Of the 177 enterprise agentic AI deployments analysed, only 31% of ad-hoc deployments achieved their stated objectives after 12 months in production. The primary failure modes were not technological: data quality deficiencies accounted for 34% of failures, governance design failures for 28%, and scope creep for 22%. Technical model limitations accounted for only 16% of failures.

31%	62%	65%	40%
Ad-hoc deployment success rate after 12 months	Failures caused by governance and data quality — not model capability	Factory model success rate — a 34pp advantage	Gartner: AI projects cancelled before 2027 due to governance failures

1.2 The Industrial Production Metaphor

The solution to the scaling challenge requires a shift from a craft-based production model to an industrial production model. In manufacturing, the factory system replaced craft production by introducing standardised components, repeatable processes, specialised roles, and systematic quality control. The result was not merely efficiency; it was a qualitative transformation in the organisation's ability to produce reliable, consistent output at scale.

The Agentification Factory applies these same principles to the production of governed AI agents. The factory is not a software platform; it is an organisational model — a dedicated capability within the enterprise designed to systematically assess processes, select appropriate architectural patterns, design robust governance, and deploy agents at scale. By treating agent deployment as a continuous production pipeline rather than a series of discrete projects, the factory model dramatically reduces the marginal cost and risk of each new deployment.

THE AD-HOC FAILURE PATTERN

Most enterprises approach agentic AI deployment through an ad-hoc, project-based model: a business unit identifies a use case, procures a vendor solution or allocates a development team, and attempts to build a bespoke agent. This approach treats each deployment as a unique engineering challenge, resulting in duplicated effort, inconsistent governance, and a failure to accumulate institutional knowledge. Each new deployment starts from scratch, repeating the same discovery process, making the same governance mistakes, and failing to leverage the learnings of previous deployments.

CHAPTER 2

Theoretical Foundations

The Agentification Factory operationalises five proven frameworks — PASF, PADE, GRAF, OCG, and Roundtrip Value Governance — into a unified organisational capability. Each framework addresses a specific failure mode in enterprise agentic AI deployment.

2.1 The Five Integrated Frameworks

The Agentification Factory model is grounded in a clear conceptual distinction between process automation and process agentification. Traditional Robotic Process Automation (RPA) automates deterministic, rule-based processes by executing predefined scripts. Process agentification deploys autonomous agents capable of reasoning, adapting, and making decisions within defined boundaries. This distinction has critical implications for organisational design — agentification requires a fundamentally different approach that integrates process knowledge, AI architecture, governance design, and value measurement.

Five frameworks underpin the factory model. Each is available as a standalone research paper at eigenvector.eu/research.

Framework	Full Name	Whitepaper	Factory Role
PASF	Process Automation Suitability Framework	From Suitability to Blueprint: PASF & PADE	BPR Unit: process triage and zone classification
PADE	Process Automation Design Engine	From Suitability to Blueprint: PASF & PADE	BPR Unit: solution design from the pattern library
GRAF	Governed Runtime for Agentic Functions	Agentic Governance Zone III: The GRAF Framework	Phase 3: seven-layer governance architecture
OCG	Ontological Compliance Gateway	Ontological Compliance Gateway	Phase 3: regulated-industry compliance mechanism
Roundtrip Value	Roundtrip Value Governance	Roundtrip Value Governance for Agentic Process Automation	Evidence Factory: verified ROI measurement

Table 1: The five integrated frameworks, their full whitepaper names, and their roles in the Agentification Factory. All papers available at eigenvector.eu/research.

2.2 The Governance Imperative

A central insight of the factory model is that governance is not a constraint on agentic AI deployment — it is an enabler. Deployments that treat governance as an architectural concern from the outset achieve dramatically higher success rates than those that treat it as a compliance checkbox. The **GRAF framework** (Agentic Governance Zone III: The GRAF Framework, available at eigenvector.eu/research) operationalises this through a seven-layer governance architecture — Input Validation, Context Boundary Management, Tool Access Control, Output Validation, Compliance Gateway, Audit and Explainability, and Human Oversight Orchestration — with the specific layers activated based on zone classification. At Agent Complexity Level 1, governance overhead consumes approximately 8% of total project cost. By Level 5, that number reaches 72% — the primary economic reason Zone IV work should not be automated.

2.3 The Value Recognition Problem

A persistent challenge in enterprise AI deployment is the gap between vendor-claimed ROI and independently verified ROI. Analysis of 177 deployments reveals that vendor-reported ROI figures are systematically overstated by an average factor of 2.1×. This gap arises from four sources: failure to account for governance overhead, attribution of pre-existing efficiency gains to the AI deployment, exclusion of failed deployments from reported averages, and use of optimistic assumptions in financial models. The **Roundtrip Value Governance** framework (Roundtrip Value Governance for Agentic Process Automation, available at eigenvector.eu/research) addresses this by requiring explicit baseline measurement before deployment and rigorous attribution analysis after deployment. The Net Program Value (NPV) formula combines realised efficiency gains against governance costs, infrastructure costs, change management investment and training cost — producing a single number that tells leadership whether the factory is creating value or creating activity. This number should be updated weekly and visible to everyone on the team.

CHAPTER 3

The BPR Unit: Process Selection with PASF and PADE

Before any agent is built, a dedicated Business Process Redesign unit runs every candidate process through a structured triage. PASF determines whether to automate. PADE determines how to design it. Together they prevent the most common and most expensive failure mode in enterprise AI: automating the wrong thing.

3.1 Why a Dedicated BPR Unit

The most common failure mode in enterprise AI programs is not a technology failure — it is a selection failure. Twelve percent of the 177 deployments analysed caused more harm than value, not because of poor engineering, but because nobody assessed whether the process was suitable for automation in the first place. Management pet projects, political pressure, and the absence of a structured intake process all contribute to this pattern. The fix is a dedicated **Business Process Redesign (BPR) unit** that runs before any technology decision is made, and a team with the discipline to actually use it.

The BPR unit is the front door of the factory. It is led by someone who can simultaneously manage process analysts, business architects, lean specialists, and automation designers, working in close coordination with business process owners embedded in the operational units. Process owners carry measurable KPIs for outcomes — FTE release, cycle time reduction, error rate improvement, compliance uplift — and they are the mechanism by which the transformation actually happens in the business.

3.2 PASF: Determining Whether to Automate

The **Process Automation Suitability Framework (PASF)** — documented in full in the whitepaper From Suitability to Blueprint: PASF & PADE (available at eigenvector.eu/research) — provides a structured, evidence-based method for assessing whether a process is suitable for agentic automation. Every candidate process receives a **PASS score**: a weighted average of eight dimensions, each scored from 0 to 10.

PASF Dimension	What It Measures	Scoring Direction
Structurability	Degree to which the process follows defined rules and steps	Higher = better
Rule-boundedness	Extent to which decisions follow explicit, codifiable rules	Higher = better
Data quality	Completeness, consistency, and accuracy of input data	Higher = better
Reversibility	Ease of correcting errors without downstream damage	Higher = better
Frequency & volume	How often the process runs and at what scale	Higher = better
Exception density	Frequency of edge cases and non-standard paths	Lower = better (inverse)
Stakeholder impact	Consequences of errors for people and relationships	Lower = better (inverse)
Regulatory constraint	Degree of regulatory oversight and compliance burden	Lower = better (inverse)

Table 2: PASF's eight scoring dimensions. The last three are scored inversely — a high exception density, high stakeholder impact, or high regulatory constraint produces a low score. Validated against 177 deployments with 74% predictive accuracy.

The PASS score determines the process's automation zone. Zone I processes (score ≥7.0) are candidates for immediate automation — well-executed deployments deliver 50–90% reduction in processing time and 30–70% reduction in cost per transaction. Zone II processes (score 5.0–6.9) require piloting with human oversight. Zone III processes (score 3.0–4.9) are not suitable for current automation and should be redesigned first. Zone IV processes (score <3.0) should not be automated at all. The BPR unit focuses the factory exclusively on Zone I and Zone II work — the realistic 35% agentification ceiling that represents years of opportunity for most enterprises.

THE 35% AGENTIFICATION CEILING

The realistic upper bound of what current AI can handle across Zone I and Zone II combined is approximately **35% of enterprise work**. Only 27% of the 177 deployments fell into the Automate Now (Zone I) category. This sounds modest until you calculate the actual volume of work it represents in a large organisation. It is years of opportunity, and most organisations have not yet worked through year one of it properly. The BPR unit's job is to systematically work through that 35% — in the right order, with the right design, and with the discipline to keep Zone III and IV work out of the pipeline.

3.3 PADE: Designing the Solution

Once a process clears the PASF triage and receives a zone classification, the **Process Automation Design Engine (PADE)** — also documented in From Suitability to Blueprint: PASF & PADE at eigenvector.eu/research — takes over. PADE's role is solution design: translating the process architecture into an agentic blueprint by selecting the optimal design pattern from the factory's pattern library.

Where PASF answers the question should we automate this?, PADE answers how should we design it? The design is derived from the process architecture rather than imposed on top of it — a critical distinction that prevents the common failure of applying a generic agent template to a process it was not designed for. PADE selects from nine validated agentic design patterns, each with defined applicability criteria, governance requirements, and tokenomics profiles. The output is a **Pattern Specification Document** that the development team uses as the authoritative blueprint for build.

3.4 The Intake and Triage Process

Every candidate process enters the factory through a structured intake form that requires enough description to assess its basic characteristics: volume, rule density, exception rate, data availability, reversibility of errors, and regulatory exposure. A process that cannot be described clearly enough to complete the intake form is not yet ready for automation — and that is useful information that saves months of wasted development time.

After intake comes triage: three sequential gates before a process reaches the development team. Gate 1 is the initial intake assessment. Gate 2 is the value, complexity, and risk determination — where the PASS score is calculated and the economics are quantified for the first time. Processes that do not clear a minimum ROI threshold of approximately 200% in year one do not proceed (with exceptions for compliance-driven automation). Gate 3 is the handover to development with a completed process blueprint designed using PADE — the automation design derived from the process architecture, not imposed on top of it.

CHAPTER 4

The Agentification Factory: Organisational Architecture

The factory is structured around four functional domains, three capability layers, and five delivery phases. This architecture ensures that strategic intent is systematically translated into governed operational reality — with clear accountability at every level.

3.1 The Three-Layer Architecture

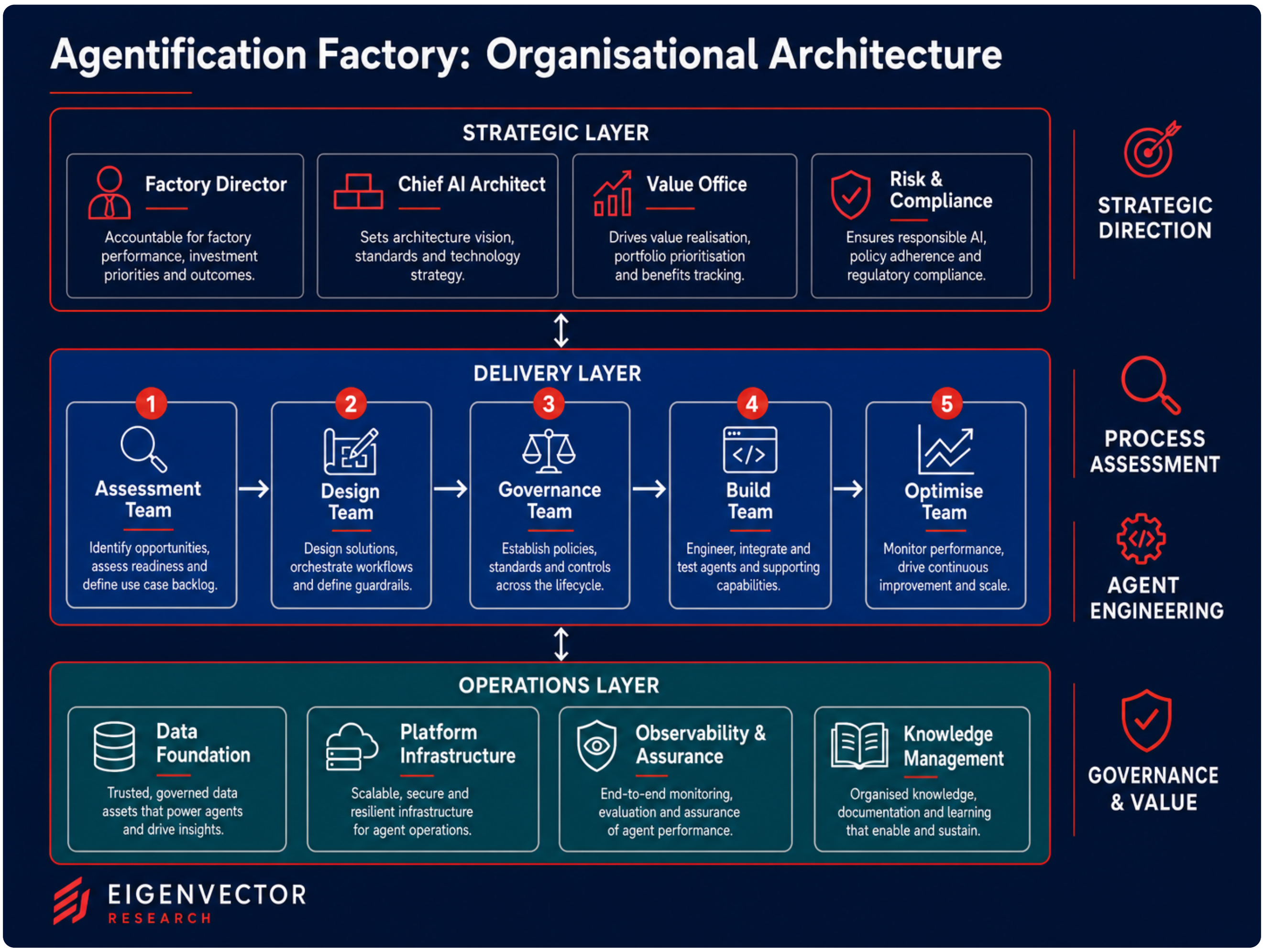


Figure 1: Agentification Factory Organisational Architecture — three capability layers (Strategic, Delivery, Operations) with four functional domains and five delivery phases

3.2 The Strategic Layer

The Strategic Layer provides executive oversight, portfolio management, and architectural governance. It is responsible for aligning the factory's output with enterprise strategy, managing the technology roadmap, and ensuring that the Roundtrip Value of deployments is accurately measured and recognised. The Strategic Layer includes four core roles:

- **Factory Director:** Accountable for factory performance, investment priorities, and outcomes. Owns the relationship with enterprise risk and compliance functions.
- **Chief AI Architect:** Sets the architecture vision, approves the pattern library, and ensures technical consistency across deployments.
- **Value Office:** Drives value realisation, portfolio prioritisation, and benefits tracking using the Roundtrip Value framework.
- **Risk & Compliance:** Ensures responsible AI deployment, policy adherence, and regulatory compliance — including EU AI Act requirements.

3.3 The Delivery Layer

The Delivery Layer is the operational core of the factory, responsible for executing the five-phase deployment pipeline. It comprises cross-functional teams that assess processes, select patterns, design governance, build agents, and optimise performance. Teams in this layer are organised around the five delivery phases, with clear handoff protocols and signed exit criteria between phases.

3.4 The Operations Layer

The Operations Layer provides the enabling infrastructure and foundational capabilities that support the delivery teams. It includes four critical capabilities:

- **Data Foundation:** Ensures data quality, provenance, and access — addressing the 34% of failures attributable to data quality issues.
- **Platform Infrastructure:** The Enterprise Intelligence Platform, agent runtime, and tool registry that provide the technical substrate for all deployments.
- **Observability & Assurance:** Real-time telemetry, anomaly detection, and audit trail capabilities required by the EU AI Act's post-market monitoring requirements.
- **Knowledge Management:** The pattern library, case study database, and prompt templates that enable the ROI Flywheel effect.

CHAPTER 5

The Five Delivery Phases

The factory's production pipeline consists of five sequential phases, each with strict exit criteria. A candidate process must pass the exit criteria of each phase before proceeding to the next — preventing the deployment of unsuitable or ungovernable agents, the primary cause of ad-hoc deployment failures.

4.1 Pipeline Overview

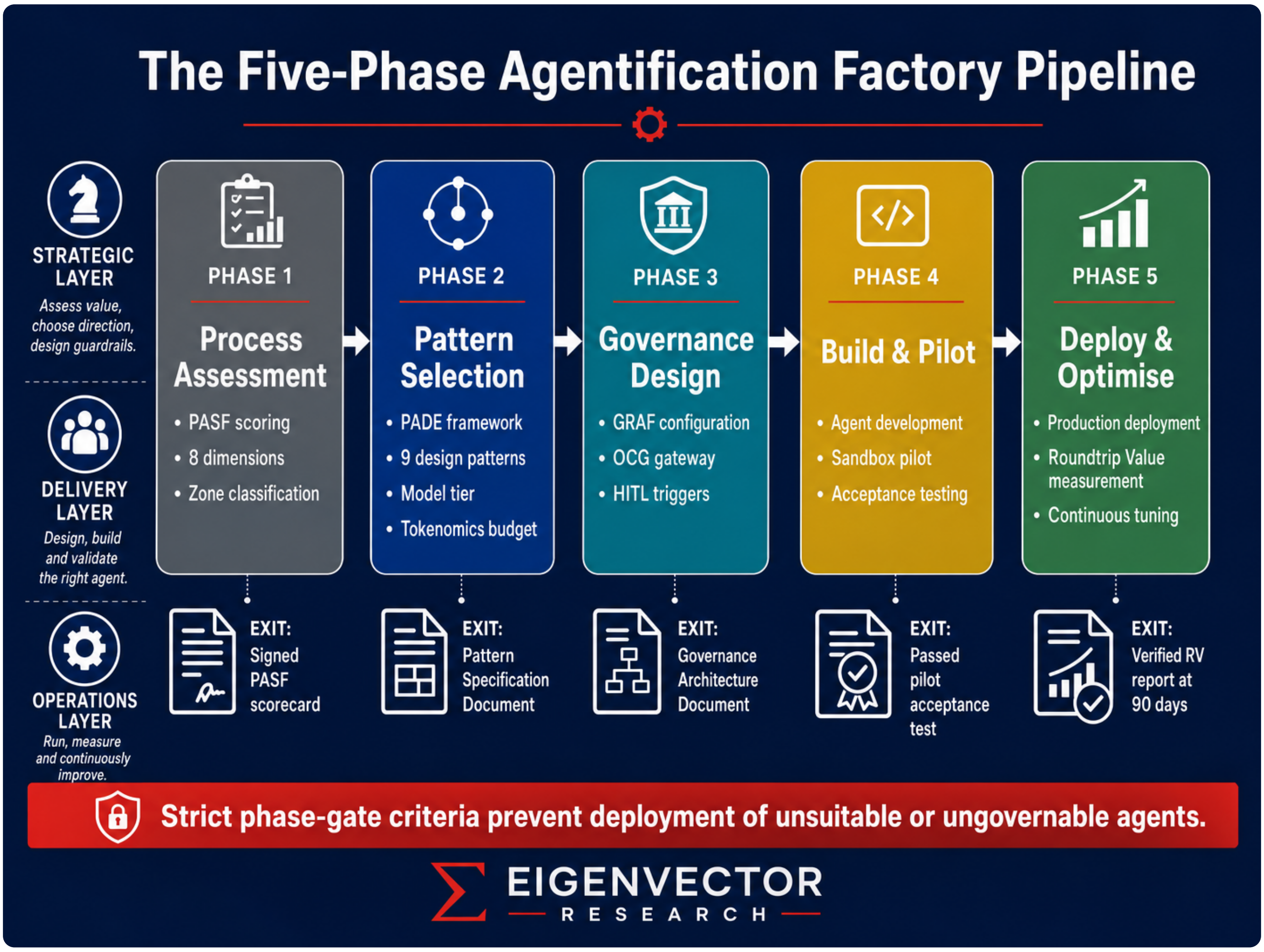


Figure 2: The Five-Phase Agentification Factory Pipeline — from Process Assessment through Pattern Selection, Governance Design, Build & Pilot, to Deploy & Optimise, with strict phase-gate exit criteria

4.2 Phase 1: Process Assessment

The Assessment Team applies the PASF framework to candidate processes, typically using process mining data to ensure objective, data-driven scoring. Each process step is scored across eight dimensions, producing a PASS score and a zone classification. The phase concludes with a hard-stop check: processes involving irreversible physical actions or high-risk decisions without digital rollback mechanisms are classified as Zone IV and rejected. The exit criterion for Phase 1 is a validated zone classification and a signed-off PASF scorecard.

A critical feature of Phase 1 is the **candidate filtering effect**: in the factory model, only 27% of assessed processes reach production deployment, but 65% of those achieve success. In the ad-hoc model, 38% of candidates reach the pilot stage, but only 14% ultimately achieve production success — a 63% attrition rate during pilot and early production. The factory's rigorous Phase 1 assessment is the primary driver of the 34 percentage point success rate advantage.

4.3 Phase 2: Pattern Selection

The Design Team applies the PADE framework to select the optimal agentic design pattern for each process step. This phase also specifies the required model tier (nano-class, mini-class, or full-class), the scaffolding architecture (tool registry, memory configuration, loop controls), and the tokenomics budget. The exit criterion for Phase 2 is a signed-off Pattern Specification Document, reviewed by the Chief AI Architect.

4.4 Phase 3: Governance Design

The Governance Team designs the required GRAF configuration, specifying which of the seven governance layers are activated and at what sensitivity levels. For Zone II and III deployments, this phase also designs the OCG gateway rules and the HITL trigger thresholds. The exit criterion for Phase 3 is a signed-off Governance Architecture Document, reviewed by the Risk & Compliance function.

4.5 Phase 4: Build and Pilot

The Build Team develops the agent, integrates it with enterprise systems, and conducts a controlled pilot in a sandbox environment. The pilot must demonstrate adherence to the governance design and achieve pre-agreed performance metrics — success rate, latency, token cost, compliance rate — before production deployment is approved. The exit criterion for Phase 4 is a passed pilot acceptance test, signed off by the Value Manager.

4.6 Phase 5: Deploy and Optimise

Following production deployment, the Optimise Team assumes responsibility for continuous tuning, tokenomics management, and performance monitoring. This phase implements the Roundtrip Value measurement cycle, tracking actual value generated against the baseline established in Phase 1. The exit criterion for Phase 5 is the first verified Roundtrip Value report, typically produced at 90 days post-deployment.

Phase	Team	Key Framework	Exit Criterion	Zone I Time	Zone II Time
1: Assessment	Assessment Team	PASF	Signed PASF scorecard	1–2 weeks	2–3 weeks
2: Pattern Selection	Design Team	PADE	Pattern Specification Document	1 week	2 weeks
3: Governance Design	Governance Team	GRAF / OCG	Governance Architecture Document	1 week	2–3 weeks
4: Build & Pilot	Build Team	All frameworks	Passed pilot acceptance test	2–3 weeks	5–7 weeks
5: Deploy & Optimise	Optimise Team	Roundtrip Value	Verified RV report at 90 days	1 week + 90 days	1 week + 90 days

Table 2: Five delivery phases with teams, frameworks, exit criteria, and typical timelines — Zone I: 6–8 weeks, Zone II: 12–16 weeks (vs. 12–16 and 24–36 weeks ad-hoc)

CHAPTER 6

The Evidence Factory

The Evidence Factory is the measurement and observability infrastructure that proves whether the agentification program is working — and catches problems before they become expensive incidents. It is where process operators work, where Net Program Value is calculated, and where tokenomics is managed at scale.

6.1 What the Evidence Factory Is

The Evidence Factory is not a dashboard. It is the operational infrastructure that **agentic process operators** work within, and it has three distinct purposes. First, it is where operational monitoring and escalation happens — including troubleshooting and root cause analysis. Second, it produces the data that is most important for program survival: the measurement of **Net Program Value (NPV)**. Third, it is where tokenomics is actively managed at scale.

The NPV formula combines realised efficiency gains against governance costs, infrastructure costs, change management investment, and training cost. It produces a single number that tells leadership whether the factory is creating value or creating activity. This number should be updated weekly, visible to everyone on the team, and visible to the people who own a budget line in the program. If it is not, the Evidence Factory has not been built and the program is running on faith.

6.2 The Process Operator Role

The process operator is the human layer that makes the Evidence Factory functional. In chemical manufacturing, a process operator monitors a continuous industrial process through instrumentation panels — interpreting signals, intervening when parameters drift outside acceptable ranges, and escalating when something cannot be resolved. The same logic applies directly to agentic factories.

Here, the process operator watches the control tower dashboards and monitors the telemetry coming from the orchestration and RPA layers. They catch the anomalies that the monitoring system flags but cannot resolve, and they have the authority and training to stop a process when necessary. This is not a junior role. It requires understanding of both the business process and the automation architecture well enough to distinguish a genuine anomaly from normal operational variation.

6.3 Operational Monitoring and Escalation

Governance in practice means building trust through four operational capabilities. **Explainability** is the ability to show why an agent made a specific decision in a specific case, in a form that a regulator or auditor can read and verify. **Traceability** is the ability to follow every action the agent took through a complete and unbroken audit trail from the triggering input to the final output. **Observability** is the real-time monitoring of agent behaviour to detect anomalies before they propagate through downstream processes and become expensive. And **kill switches** — actual tested working kill switches that halt a process instantly.

The GRAF framework (Agentic Governance Zone III: The GRAF Framework, available at eigenvector.eu/research) provides the architectural specification for all four capabilities. The Evidence Factory is where they are operated.

6.4 Tokenomics at Scale

Tokenomics is the management of the economics of running AI at scale, and it is one of the disciplines that catches AI programs by surprise at volume. The Evidence Factory is where it is actively managed.

At the individual agent level, **Pennywise Tokenomics** addresses per-transaction inference cost: what does it cost in compute to run this agent on this process instance, and does that cost remain acceptable at ten times the current volume? An agent that processes an invoice at €0.003 per run looks fine in isolation — but at 400,000 invoices per year, that is €1,200 in inference cost for one process alone, before infrastructure, before monitoring, before the other agents touching the same workflow.

At the portfolio level, **Poundwise Patternomics** — documented in the whitepaper Patternomics: The Economics of Agentic Execution at eigenvector.eu/research — optimises the structural cost patterns that emerge across a portfolio of agents running simultaneously. The pattern that destroys programs is not one expensive agent — it is forty moderately priced agents all calling the same large language model simultaneously at peak processing hours, none of them batched or cached, all of them running full inference on inputs that are 80% identical to something already processed ten minutes ago. A brilliantly designed agent with unmanaged tokenomics can produce negative ROI at scale even when every individual deployment dashboard looks healthy. This is the compute bill that arrives in month eight and makes a CFO ask questions nobody prepared answers for.

> **THE EVIDENCE FACTORY MANDATE**
>
> The Evidence Factory exists to answer one question continuously: **is this program creating value or creating activity?** The NPV formula, the process operator dashboards, the tokenomics monitoring, and the Roundtrip Value reports all serve this single purpose. Programs that skip the Evidence Factory do not fail immediately — they fail quietly, in month eight, when the compute bill arrives and the ROI projections cannot be reconciled with the actual numbers.

CHAPTER 7

Empirical Performance: Factory vs. Ad-Hoc

The comparative performance analysis demonstrates a substantial and consistent advantage for the factory model across all key metrics — grounded in an empirical database of 177 deployments across 20 sectors, with 68 factory model deployments and 109 ad-hoc deployments.

5.1 Overall Performance Comparison

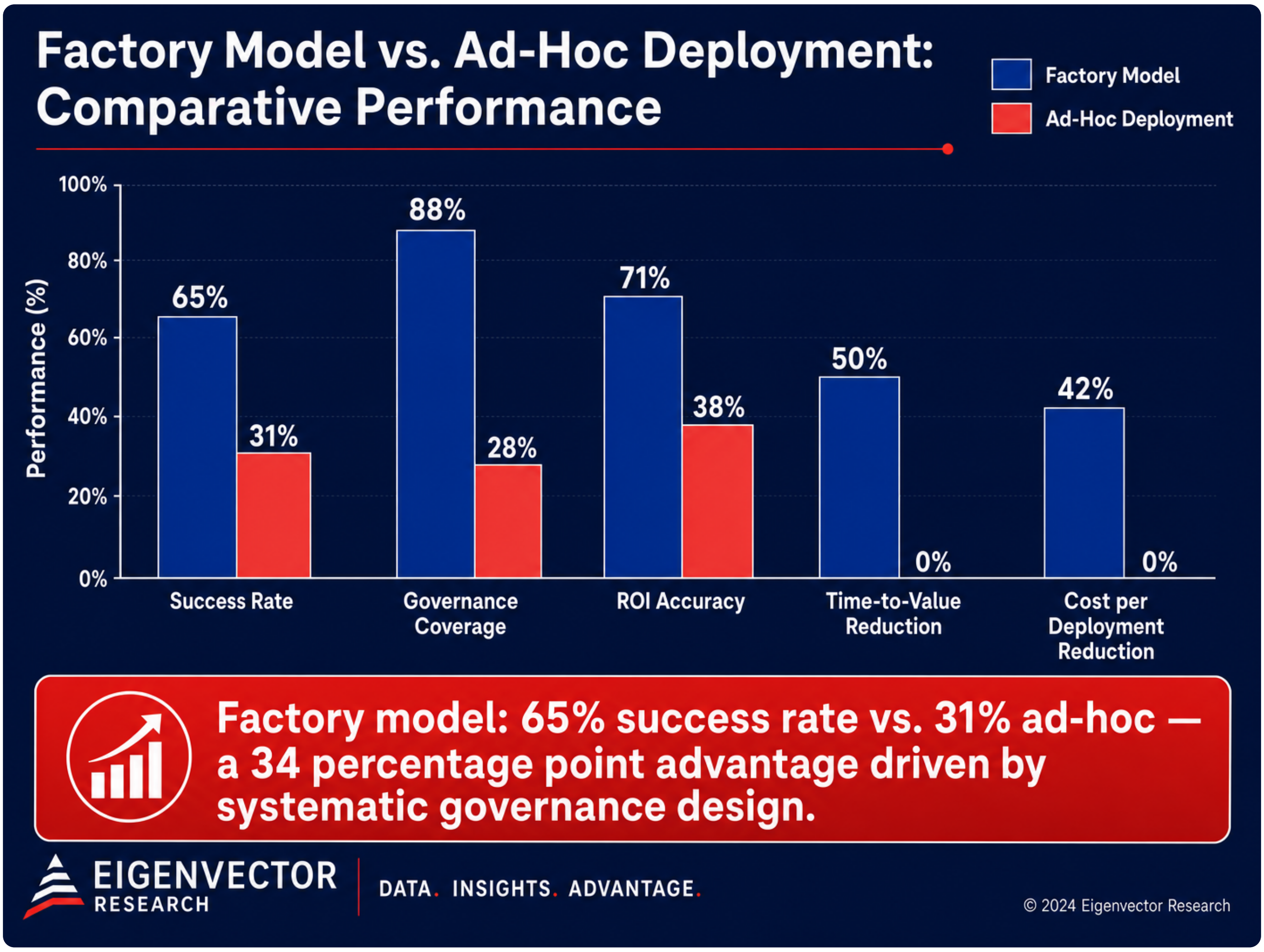


Figure 3: Factory Model vs. Ad-Hoc Deployment — Comparative Performance Analysis across five key metrics. Factory model achieves 65% success rate vs. 31% ad-hoc, 88% governance coverage vs. 28%, and 71% ROI accuracy vs. 38%.

5.2 Key Performance Metrics

Success Rate: Factory model deployments achieved a 65% success rate (defined as meeting more than 80% of stated objectives after 12 months in production), compared to only 31% for ad-hoc deployments. This 34 percentage point difference is primarily attributable to the factory's rigorous Phase 1 assessment, which prevents the deployment of fundamentally unsuitable processes.

Time-to-Value: The factory model reduces time-to-value by 40–60%. A typical Zone I deployment requires 6–8 weeks in a factory model, compared to 12–16 weeks in an ad-hoc model. A Zone II deployment requires 12–16 weeks in a factory model, compared to 24–36 weeks in an ad-hoc model. This acceleration is achieved through the reuse of standardised patterns, pre-approved governance templates, and established platform infrastructure.

Cost per Deployment: The factory model achieves a 42% reduction in marginal deployment cost. The initial setup cost of a factory (estimated at 6–12 months of equivalent project cost) is recovered after approximately 4–6 deployments. Beyond this break-even point, each additional deployment is substantially cheaper than an equivalent ad-hoc deployment.

Governance Coverage: Factory model deployments achieve 88% governance coverage (the proportion of GRAF layers actively monitored), compared to 28% for ad-hoc deployments. This difference directly explains the higher success rate: the 62% of failures attributable to governance and data quality issues in ad-hoc deployments are largely prevented by the factory's systematic governance design.

ROI Accuracy: Factory model deployments achieve 71% ROI accuracy (the ratio of verified ROI to projected ROI), compared to 38% for ad-hoc deployments. This improvement is directly attributable to the Roundtrip Value framework, which enforces baseline measurement and rigorous attribution analysis.

5.3 Sector Performance

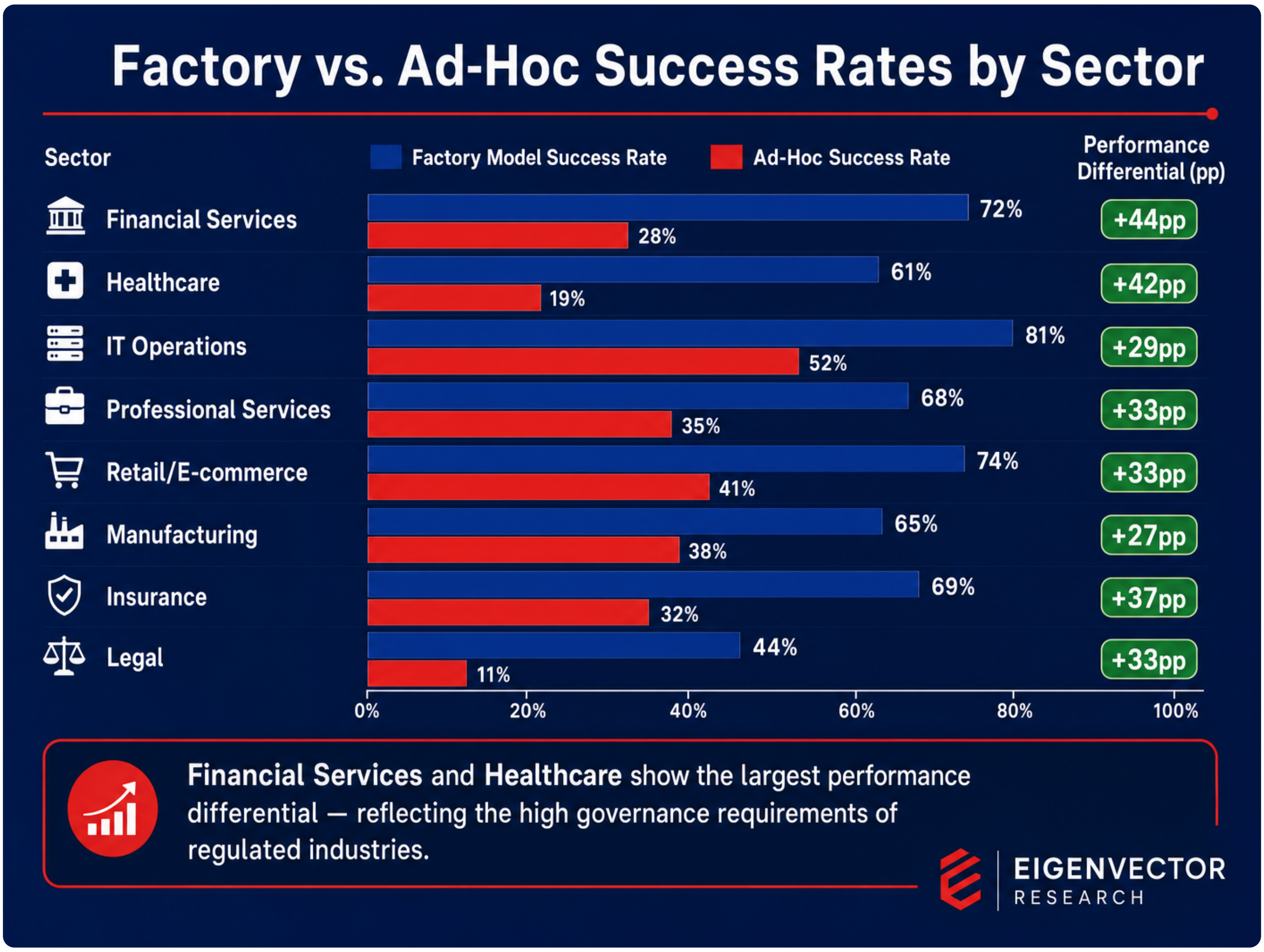


Figure 4: Factory vs. Ad-Hoc Success Rates by Sector — Financial Services and Healthcare show the largest performance differential (+44pp and +42pp respectively), reflecting the high governance requirements of regulated industries

Sector	Factory Success Rate	Ad-Hoc Success Rate	Differential	Primary Pattern
Financial Services	72%	28%	+44pp	Critic-Actor + OCG
Healthcare	61%	19%	+42pp	Plan-Execute + HITL
IT Operations	81%	52%	+29pp	Single-Tool + ReAct
Professional Services	68%	35%	+33pp	Orchestrator-Subagent
Retail / E-commerce	74%	41%	+33pp	ReAct + Memory
Manufacturing	65%	38%	+27pp	Plan-Execute + Orchestrator
Insurance	69%	32%	+37pp	Plan-Execute + OCG
Legal	44%	11%	+33pp	Critic-Actor + OCG

Table 3: Sector-specific factory vs. ad-hoc success rates — the factory model outperforms ad-hoc deployment in every sector, with the largest advantages in regulated industries

CHAPTER 8

The ROI Flywheel Effect

The most significant strategic advantage of the Agentification Factory is not its immediate efficiency, but its compounding nature. As the factory executes more deployments, it generates an ROI Flywheel effect that deepens the organisation's competitive moat over time through eight reinforcing mechanisms.

6.1 The Eight Flywheel Mechanisms

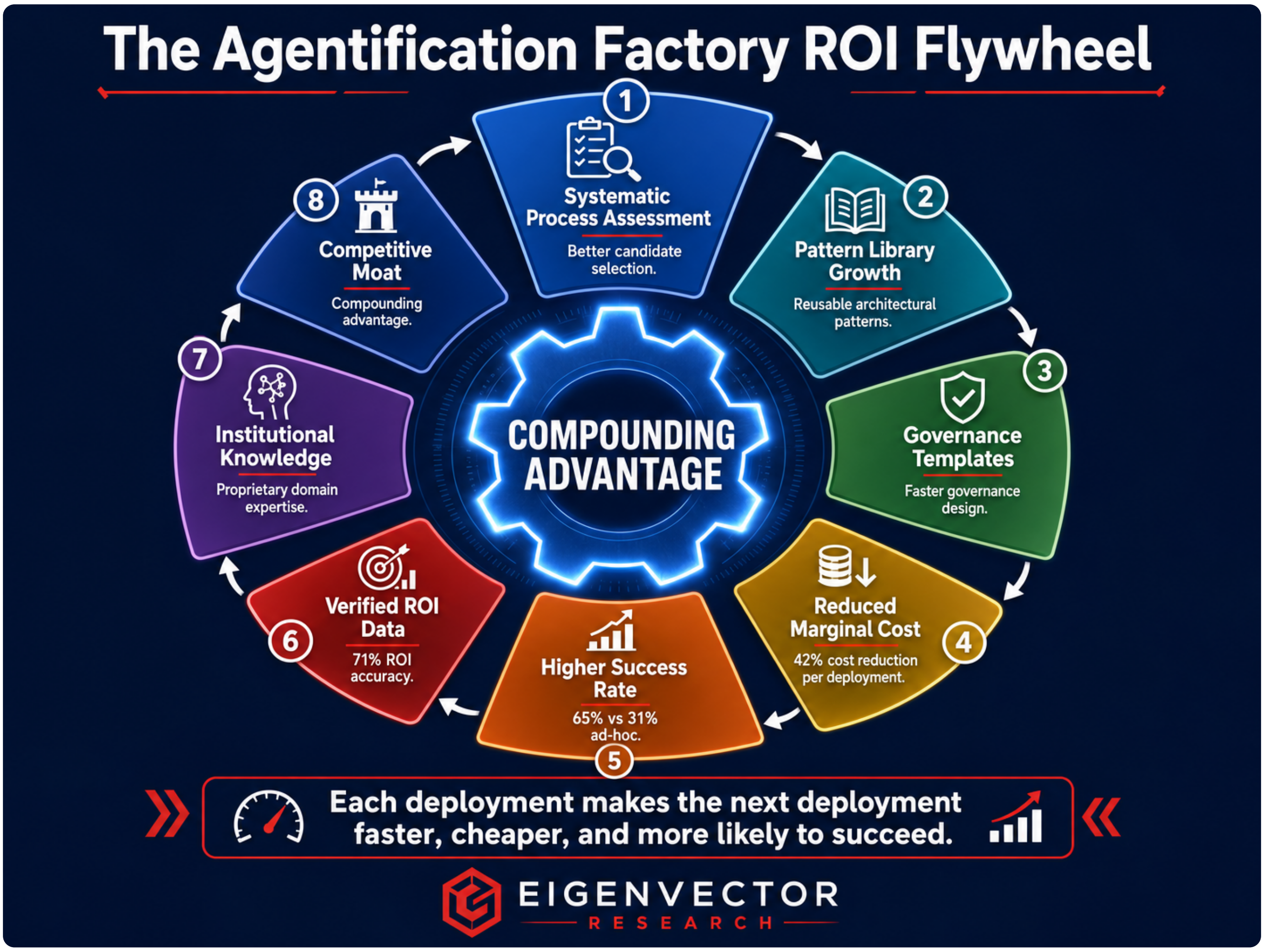


Figure 5: The Agentification Factory ROI Flywheel — eight reinforcing mechanisms that create a self-sustaining cycle of compounding competitive advantage

The flywheel operates through eight reinforcing mechanisms that collectively create a self-sustaining cycle of improvement. Each mechanism builds on the previous, creating a compounding effect that makes each successive deployment faster, cheaper, and more likely to succeed:

#	Mechanism	Effect	Empirical Evidence
1	**Systematic Process Assessment**	Better candidate selection, fewer failed deployments	27% of assessed processes reach production; 65% succeed
2	**Pattern Library Growth**	Reusable architectural patterns reduce design time	42% marginal cost reduction after 4–6 deployments
3	**Governance Templates**	Faster governance design, consistent quality	88% governance coverage vs. 28% ad-hoc
4	**Reduced Marginal Cost**	Each deployment cheaper than the last	Break-even at 4–6 deployments; 42% cost reduction
5	**Higher Success Rate**	More successful deployments generate more data	65% success rate vs. 31% ad-hoc
6	**Verified ROI Data**	Accurate ROI data enables better investment decisions	71% ROI accuracy vs. 38% ad-hoc
7	**Institutional Knowledge**	Proprietary domain expertise not replicable by competitors	Pattern library encodes organisation-specific knowledge
8	**Competitive Moat**	Compounding advantage widens over time	3 durable competitive advantages identified

Table 4: The eight ROI Flywheel mechanisms with effects and empirical evidence

6.2 The Break-Even Analysis

The initial setup cost of a factory is estimated at 6–12 months of equivalent project cost. This investment is recovered after approximately 4–6 deployments, after which each additional deployment is substantially cheaper than an equivalent ad-hoc deployment. Organisations that delay establishing a factory capability are not merely missing efficiency gains — they are allowing competitors to accumulate pattern libraries, governance expertise, and data infrastructure that become increasingly difficult to replicate.

CHAPTER 9

Implementation Blueprint and AFCMM

The Agentification Factory Capability Maturity Model (AFCMM) provides a five-level framework for assessing and building factory capability. Most organisations are currently at Level 1–2. The minimum viable factory is Level 3. The target state for competitive advantage is Level 4–5.

7.1 The AFCMM

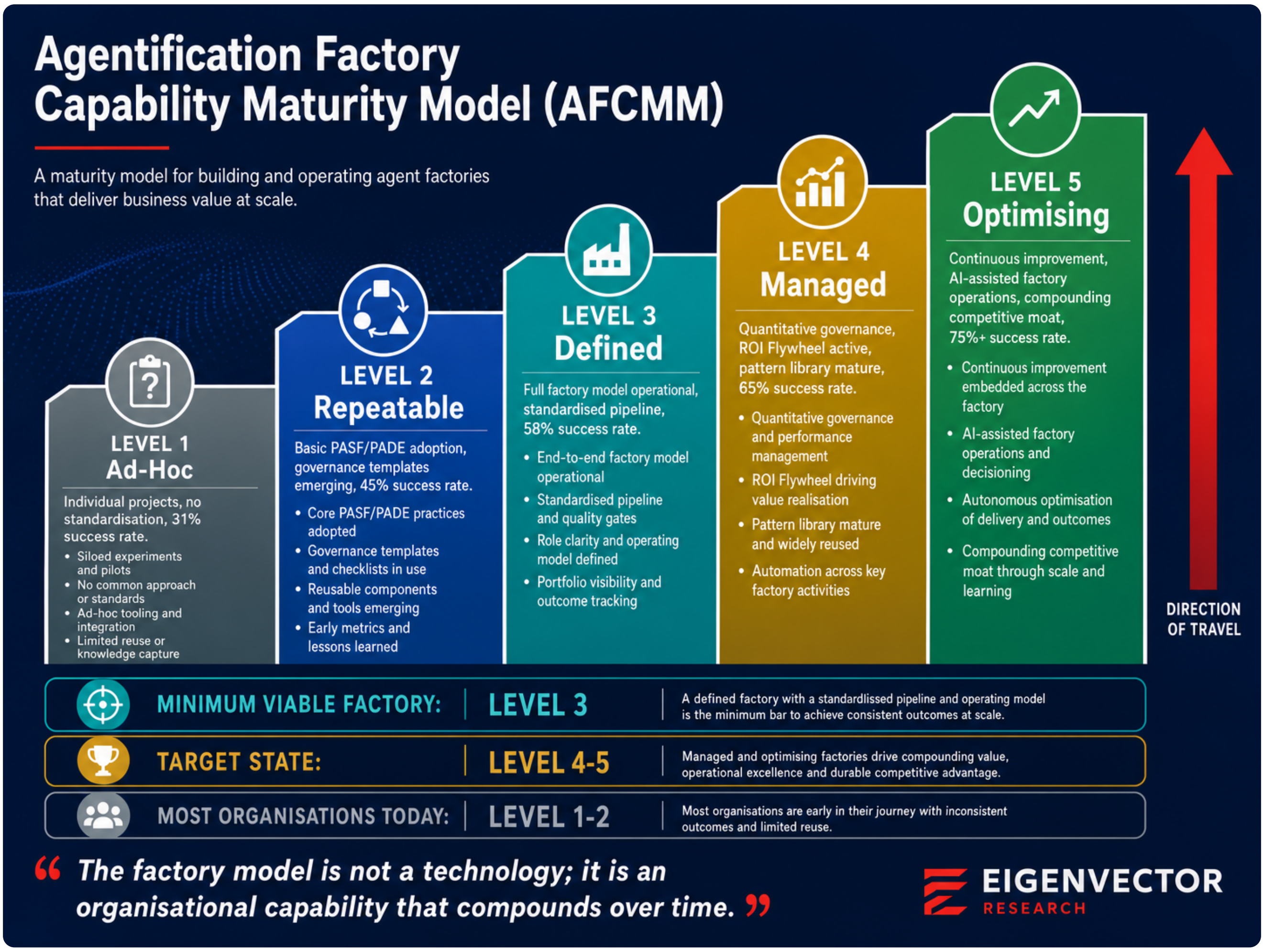


Figure 6: Agentification Factory Capability Maturity Model (AFCMM) — five levels from Ad-Hoc through Repeatable, Defined, Managed, to Optimising, with success rates from 31% to 75%+

7.2 Minimum Viable Factory

The minimum viable factory (MVF) is the smallest organisational configuration that can reliably execute the five-phase deployment pipeline. The MVF requires a minimum of eight roles: Factory Director (0.5 FTE), Chief AI Architect (1 FTE), Process Assessment Lead (1 FTE), AI Design Engineer ×2 (2 FTE), Governance Engineer (1 FTE), Build Engineer ×2 (2 FTE), and Value Manager (0.5 FTE). Total: 8 FTE minimum.

Role	FTE	Primary Responsibility	Key Framework
Factory Director	0.5	Portfolio governance, executive reporting, investment prioritisation	All frameworks
Chief AI Architect	1.0	Architecture standards, pattern library, technology roadmap	PADE, GRAF
Process Assessment Lead	1.0	PASF scoring, zone classification, process mining	PASF
AI Design Engineer	2.0	Pattern selection, scaffolding design, tokenomics budgeting	PADE
Governance Engineer	1.0	GRAF configuration, OCG design, HITL trigger design	GRAF, OCG
Build Engineer	2.0	Agent development, system integration, pilot execution	All frameworks
Value Manager	0.5	Roundtrip Value measurement, ROI reporting, benefits tracking	Roundtrip Value

Table 5: Minimum Viable Factory roles and responsibilities — 8 FTE minimum to execute the five-phase pipeline reliably

CHAPTER 10

Sector Case Studies

Four case studies illustrate how the factory model operates in practice across different sectors and zone classifications — from BNY's legal document processing (Zone II) and MUSC Health's prior authorisation (Zone III) to PwC's knowledge management (Zone I–II) and Klarna's customer service (Zone I).

8.1 BNY: Legal Document Processing (Financial Services, Zone II)

BNY's deployment of agentic AI for legal document processing demonstrates the factory model's approach to Zone II processes in regulated industries. BNY's factory began with a PASF assessment that classified legal document review as a Zone II process — moderately structured, moderate risk, significant exception density. The OCG gateway was configured with a formal ontological representation of applicable legal standards, enabling automated compliance checking before any document action.

The result was a 75% reduction in legal review time for standard contracts, with zero compliance incidents in the first 12 months of production operation. The Orchestrator-Subagent pattern was selected for complex, multi-domain deployments, with specialist subagents for legal analysis, compliance review, and client communication.

8.2 MUSC Health: Prior Authorisation (Healthcare, Zone III)

MUSC Health's deployment for prior authorisation processing demonstrates the critical importance of the OCG framework in regulated industries. Prior authorisation is a Zone III process: high compliance sensitivity, significant financial consequences, and complex medical judgment requirements. The OCG gateway was configured with a formal ontological representation of CMS billing rules, enabling the system to validate each authorisation decision against the applicable regulatory framework before execution.

The result was that 40% of prior authorisation requests were processed without human review, with a compliance rate of 99.7% — higher than the human baseline of 98.2%. This outcome was only achievable because the OCG gateway was designed into the architecture from the outset, not added as an afterthought.

8.3 PwC: Knowledge Management at Scale (Professional Services, Zone I–II)

PwC's deployment of Microsoft Copilot across 230,000 users demonstrates the scalability of the factory model in a professional services context. PwC began with a PASF assessment identifying knowledge management and document analysis as Zone I–II processes. The Orchestrator-Subagent pattern was selected for complex client deliverable generation, with specialist subagents for research, analysis, writing, and compliance review.

The result was 500,000 hours of capacity freed in the first month of deployment — a figure that PwC's Value Manager independently verified using the Roundtrip Value framework, confirming that the figure represented genuine incremental capacity rather than pre-existing efficiency gains.

8.4 Klarna: Customer Service at Scale (Retail, Zone I)

Klarna's deployment for customer service represents one of the most widely cited examples of agentic AI ROI. Klarna's agent handles the equivalent work of 853 full-time customer service agents, generating $60 million in annual savings. The Memory-Augmented ReAct pattern enables the agent to maintain context across multi-turn conversations while dynamically accessing customer account data, transaction history, and product information.

Klarna's factory approach is notable for its rigorous HITL design: rather than attempting to fully automate all customer interactions, Klarna designed a clear escalation protocol that routes complex or high-value interactions to human agents. This HITL design is a key factor in the deployment's high success rate and customer satisfaction scores.

CHAPTER 11

Governance Integration

The GRAF framework is the governance backbone of the Agentification Factory. Every deployment produced by the factory is designed against the GRAF architecture, with the specific layers activated based on zone classification and pattern selection.

9.1 GRAF Layer Activation by Zone

Zone	GRAF Layers Active	HITL Design	OCG Required
Zone I	Layers 1–3: Input Validation, Context Boundary, Tool Access Control	Sampling-based (5–10% of outputs)	No
Zone II	Layers 1–5: + Output Validation, Compliance Gateway	Exception-based (confidence threshold triggers)	Optional
Zone III	All 7 layers: + Audit & Explainability, Human Oversight Orchestration	Mandatory pre-execution approval for high-risk actions	Mandatory

Table 6: GRAF layer activation by zone classification — governance scales with risk, not uniformly applied

9.2 The Four HITL Design Principles

The factory model applies four HITL design principles derived from empirical analysis of successful deployments. **Zone-appropriate trigger design** ensures that oversight intensity matches risk level — sampling-based for Zone I, exception-based for Zone II, mandatory pre-execution for Zone III. **Graceful escalation** ensures that HITL workflows escalate smoothly from automated to human handling without disrupting the user experience. **HITL as a learning mechanism** captures human reviews systematically to improve agent performance over time. **HITL cost accounting** requires that the cost of human review be included in the Roundtrip Value calculation — deployments that appear to generate high ROI by eliminating human labour but require substantial HITL oversight are not generating the claimed value.

CHAPTER 12

EU AI Act Implications

The EU AI Act, which entered into force in August 2024, establishes a risk-based regulatory framework for AI systems deployed in the European Union. The factory model's zone classification system maps naturally onto the EU AI Act's risk categories, providing a practical implementation path for compliance.

10.1 Zone-to-Risk Category Mapping

PASF Zone	EU AI Act Category	Factory Compliance Mechanism
Zone IV (Do Not Automate)	Prohibited AI practices	Hard-stop criteria in Phase 1 prevent deployment
Zone III (Automate with Caution)	High-risk AI systems	Mandatory OCG gateway + full GRAF governance (all 7 layers)
Zone II (Pilot First)	Limited risk AI systems	GRAF Layers 1–5 + transparency obligations
Zone I (Automate Now)	Minimal risk AI systems	GRAF Layers 1–3 + basic governance requirements

Table 7: PASF zone to EU AI Act risk category mapping — the factory model provides a natural compliance pathway

10.2 Documentation and Audit Requirements

The EU AI Act requires high-risk AI systems to maintain comprehensive technical documentation. The factory model's standardised documentation artefacts — PASF scorecards, Pattern Specification Documents, Governance Architecture Documents, and Roundtrip Value reports — provide a natural foundation for EU AI Act compliance documentation. The factory's Observability & Assurance infrastructure also provides the monitoring and logging capabilities required by the EU AI Act's post-market monitoring requirements.

CHAPTER 13

Strategic Implications for Leadership

Enterprise leaders navigating the adoption of the Agentification Factory model must manage four fundamental strategic tensions and build the business case for a slower-but-more-reliable approach that creates a compounding competitive moat.

11.1 The Four Strategic Tensions

Scalability vs. Adaptability. The factory model's standardisation creates efficiency and consistency, but may reduce the organisation's ability to respond rapidly to novel use cases or emerging AI capabilities. Leaders must design the factory with sufficient flexibility to incorporate new patterns and frameworks as the technology evolves, without sacrificing the governance rigour that makes the factory model effective.

Experience vs. Expediency. The factory model's phased deployment approach prioritises thoroughness over speed. Business units under pressure to demonstrate AI value quickly may resist the factory's assessment and governance requirements. Leaders must build the business case for the factory model's slower-but-more-reliable approach, using the empirical evidence of higher success rates and lower failure costs.

Supervision vs. Autonomy. As agents become more capable, the appropriate level of human supervision decreases. Leaders must design HITL workflows that evolve with the technology, progressively reducing human oversight as agent reliability is demonstrated, while maintaining the ability to increase oversight rapidly if problems emerge.

Retrofit vs. Reengineer. The factory model can be applied to existing processes (retrofit) or used to design new processes from scratch (reengineer). Reengineering typically generates higher ROI but requires greater organisational change. Leaders must make deliberate choices about when to retrofit and when to reengineer, based on the strategic importance of the process and the organisation's change capacity.

11.2 The Factory as Competitive Differentiator

The Agentification Factory model is not merely an operational efficiency tool; it is a source of sustainable competitive advantage. Organisations that establish a factory capability early gain three durable advantages: a proprietary pattern library that encodes their specific business domain knowledge; institutional expertise in agentic AI governance that is increasingly valuable as regulatory requirements intensify; and the data infrastructure and quality standards required for advanced agentic AI applications.

The Strategic Imperative

Organisations that delay establishing a factory capability are not merely missing efficiency gains — they are allowing competitors to accumulate pattern libraries, governance expertise, and data infrastructure that become increasingly difficult to replicate. The factory model is not a technology; it is an organisational capability that compounds over time. The organisations that begin building it today will have a structural advantage that cannot be purchased or copied by competitors who start later.

CHAPTER 14

Conclusion and Recommendations

The Agentification Factory model represents the organisational response to the scaling challenge in enterprise agentic AI. The empirical evidence is unambiguous: organisations that adopt the factory model achieve higher success rates, faster time-to-value, lower marginal costs, and more accurate ROI recognition.

12.1 Summary of Evidence

Metric	Factory Model	Ad-Hoc	Advantage
Success Rate	65%	31%	+34 percentage points
Time-to-Value	6–16 weeks	12–36 weeks	40–60% faster
Marginal Cost	Break-even at 4–6 deployments	Full cost per deployment	42% reduction
Governance Coverage	88%	28%	+60 percentage points
ROI Accuracy	71%	38%	+33 percentage points
Vendor ROI Overstatement	1.4× (with RV framework)	2.1× (without RV framework)	More accurate forecasting

Table 8: Summary of empirical evidence comparing factory model vs. ad-hoc deployment across six key metrics

12.2 Recommendations for Enterprise Leaders

1. **Establish the factory now.** The ROI Flywheel effect means that early movers gain compounding advantages. Organisations that delay are not merely missing efficiency gains — they are allowing competitors to build structural advantages.

2. **Start with the Strategic Layer.** The Factory Director, Chief AI Architect, and Value Manager roles are the minimum viable leadership team. Without these roles, the factory cannot maintain governance rigour or measure verified ROI.
3. **Build the Data Foundation first.** Data quality issues are the single largest failure mode (34% of failures). Invest in data quality, provenance, and access before scaling deployment activity.
4. **Adopt the PASF as the gateway.** Require PASF zone classification as a prerequisite for all agentic AI deployment approvals. This single change will prevent the majority of costly deployment failures.
5. **Measure with the Roundtrip Value framework.** Require verified ROI reports for all deployments. Vendor-reported figures are systematically overstated by 2.1×. Accurate measurement enables better investment decisions and builds board-level confidence.
6. **Plan for Level 3 as the minimum viable factory.** The AFCMM Level 3 (Defined) is the minimum state at which the factory can reliably execute the five-phase pipeline. Plan the investment required to reach Level 3 within 18 months.

References

van Hurne, M. H. K. (2026). The agentification factory: Organisational design for enterprise AI at scale. Eigenvector Research. April 2026.

van Hurne, M. H. K. (2026). From suitability to blueprint: A unified framework for agentic AI process automation in enterprise environments. TechrXiv. DOI: 10.13140/RG.2.2.29631.16800

van Hurne, M. H. K. (2025a). Roundtrip value governance for agentic process automation. Eigenvector Research.

van Hurne, M. H. K. (2025b). Process automation suitability framework (PASF) and process automation design engine (PADE). Eigenvector Research.

van Hurne, M. H. K. (2025c). Governed runtime for agentic functions (GRAF): A seven-layer governance architecture. Eigenvector Research.

van Hurne, M. H. K. (2025d). Ontological compliance gateway (OCG): Neuro-symbolic compliance for regulated agentic AI. Eigenvector Research.

Gartner. (2025). Predicts 2026: Artificial intelligence. Gartner Research.

European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 (EU AI Act). Official Journal of the European Union.

McKinsey & Company. (2024). The state of AI in 2024. McKinsey Global Institute.

National Institute of Standards and Technology. (2024). AI Risk Management Framework (AI RMF 1.0). NIST.

Forrester Research. (2025). The state of AI automation in enterprise environments. Forrester.

IDC. (2025). Worldwide artificial intelligence spending guide. IDC.

EIGENVECTOR

Build it once. Deploy it at scale. Compound the advantage.

Marco H.K. van Hurne

marco.vanhurne@eigenvector.eu

info@eigenvector.eu

https://www.eigenvector.eu

Eigenvector Research · Inholland University of Applied Sciences