



G-Cloud 15

AI Factory

January 2026

Contents

PwC AI Factories 3

 What “AI Factory” means in practice 3

 How we deliver: standards-first, but delivery-led..... 3

 Core AI Factory principles we tailor to each department 3

 Accelerators and optional tools — without lock-in 4

 Risk, safety and trust built in 4

 Outcomes that matter 4

 Extended viewpoint: PwC’s generic experience delivering AI Factory services 5

 Depth of specialist capability and scalable resourcing..... 6

 Proven patterns for moving from PoC to production 6

 Cloud-native AI Factories and multi-cloud pragmatism 6

 Controls for safety, security, compliance and public trust..... 6

 Cost and value management built into the platform 7

 “Client Zero” learning and accelerators 7

 Outcome orientation: measurable user value, scalability and operational resilience 7

PwC AI Factories

PwC deploys AI Factories through secure-by-design, reusable “golden path” patterns, automated evaluation gates and operate-by-default delivery. We assess AI readiness, define the roadmap, build the production runway (CI/CD, IaC, observability), and release only when performance, security, cost and operability thresholds are evidenced and documented.

PwC’s AI Factory is a repeatable, secure, “production runway” for taking AI from idea through to live, user-facing services, tools and reusable components at pace, and with the controls public sector teams need to operate safely at scale. It combines a proven operating model, reusable technical patterns (“golden paths”), and delivery guardrails across architecture, data engineering, DevSecOps/ MLOps/ LLMOps, user-centred design, and performance evaluation so that teams can move beyond experimentation and consistently deploy AI into business-as-usual at a predictable cadence.

What “AI Factory” means in practice

Public sector organisations are increasingly expected to scale AI while maintaining compliance, security, transparency and value-for-money. PwC’s AI Factory approach is designed to integrate advanced AI with existing cloud infrastructures and systems and to help scale AI solutions by designing effective architectures across cloud platforms. We start by helping clients become “AI Ready”, then define an integration strategy and delivery roadmap, and deliver solutions through standardised patterns that reduce rework and improve handover into operations.

How we deliver: standards-first, but delivery-led

Our delivery approach embeds required standards into the delivery mechanics, so the “secure, standard route” is also the fastest route. We align to client governance and commonly adopted public sector practices (e.g., GDS-aligned delivery, Technology Code of Practice, architecture and service management standards), and we implement CI/CD guardrails, evidence packs, and a clear “definition of done” so assurance is continuous and rework is reduced.

Core AI Factory principles we tailor to each department

PwC’s AI Factory delivery is guided by principles that directly address public sector priorities (security, auditability, cost control, operability and reuse): secure-by-design; reuse-first “golden path” patterns; evidence-led AI quality with automated evaluation gates; operate-by-default (monitoring/ runbooks/ SLOs defined pre-live); and design-to-cost with FinOps guardrails (unit-cost metrics, budgets, alerts and rate limits).

Accelerators and optional tools — without lock-in

Where helpful, we bring proven accelerators refined through PwC’s “Client Zero” experience. Critically, all services, pipelines, controls and documentation are implemented using client-approved tooling and patterns so the client can operate, support and extend solutions without dependency on proprietary PwC platforms.

Risk, safety and trust built in

Moving from PoC to production introduces predictable risks (data quality/lineage, security/privacy, model drift, hallucinations/prompt injection, and cost volatility). We manage these through:

- Gated delivery and measurable controls
- Data contracts and automated validation
- Threat modelling and least-privilege controls
- Drift monitoring with rollback/retrain triggers
- Red-teaming and guardrails for GenAI
- Cost controls such as budgets, rate limits, caching/batching and spend alerting — to keep services safe, reliable and affordable.

Outcomes that matter

The AI Factory is designed to help departments:

- Scale AI into real services (not just pilots)
- Increase productivity
- Track benefits from business case to value realisation
- Upskill teams through co-delivery and knowledge transfer
- Create a scalable framework for future AI delivery across cloud platforms.

Features of PwC's AI Factory approach

- Implement AI readiness triage, building confidence in idea qualification and prioritisation
- Develop reusable reference architectures with reusable “golden path” patterns
- Create a scalable and robust AI production runway
- Evidence-led quality gates
- LLMOps/MLOps toolchain integration within your enterprise technology stack
- Responsible and secure AI controls
- Design-to-cost guardrails
- Co-delivery and knowledge transfer using SFIA-aligned specialists.

Benefits of PwC's AI Factory approach

- More AI solutions reach production via repeatable release gates and standard delivery patterns
- Reduced rework and faster delivery through paved-road templates and embedded standards
- Safer deployments with measurable compliance controls and auditable change
- Predictable operational cost through budgets, alerts, rate limiting, caching and model selection optimisation
- Improved adoption and usability by measuring user outcomes alongside technical performance
- Scalable foundation for future AI delivery across cloud platforms and departments
- Benefits realisation tracked from business case through implementation to value
- Upskilled teams and reduced supplier dependency via co-delivery, KT, and client-approved toolchains.

Extended viewpoint: PwC's generic experience delivering AI Factory services

PwC's experience in AI Factory delivery comes from repeatedly solving the same core public sector challenge: organisations can experiment quickly, but struggle to industrialise AI into secure, reliable, cost-controlled services that can be supported long-term. Our AI Factory approach is built to close this gap by combining:

- Operating model and governance
- A repeatable engineering “paved road”
- Delivery disciplines that make production readiness non-negotiable without slowing teams down.

Depth of specialist capability and scalable resourcing

We field SFIA-aligned specialists across architecture design, data engineering, DevSecOps/ MLOps/ LLMOps, user-centred design, performance evaluation, product delivery and change. Where clients need rapid access to niche skills, we can draw from PwC's Tech Catalyst community (~4,500 technology and AI specialists) to bring the right expertise at the right time, aligned to the client's governance and delivery processes.

Proven patterns for moving from PoC to production

In delivery, we focus on the practical “factory mechanics” that increase the throughput of production releases:

- Triage and clear acceptance criteria
- Reference architectures and ADRs
- CI/CD guardrails
- Infrastructure-as-code modules
- Production readiness packs that enable smooth handover.

This approach has been used to address environments where organisations were “experimenting often, but seeing too few solutions scaled into production”, by introducing repeatable release gates and operational readiness criteria to reduce rework and ensure consistent transition into operations.

Cloud-native AI Factories and multi-cloud pragmatism

Our AI Factory work typically sits on existing cloud platforms, integrates with enterprise systems and is designed to meet our clients where they are. We assess the current cloud set-up to identify integration points and what is required to be “AI Ready”, then deliver an integration strategy and roadmap, selecting and tailoring AI tooling to automate processes and enhance analytics without disruption to operations.

Controls for safety, security, compliance and public trust

AI Factory success in government depends on trust - demonstrable security, privacy, responsible AI and auditability. We therefore treat risk management as part of delivery:

- Threat modelling and assurance-led design
- Least-privilege access and secrets management

- Secure logging
- Drift monitoring with retraining/rollback triggers
- GenAI red-teaming with grounding and refusal guardrails and human-in-the-loop for higher-risk decisions.

Cost and value management built into the platform

Departments need AI services that are financially sustainable. We embed FinOps guardrails (unit-cost metrics, budgets, alerts, rate limits, and spend dashboards) and treat cost as a release gate alongside reliability, performance and security that are important considerations when improving predictability and reducing cost volatility as services scale.

“Client Zero” learning and accelerators

PwC has built its own AI Factory capability using reusable patterns, reference materials and accelerators (e.g., Terraform-based standard architectures, shared services for LLM provisioning/management, and optional agentic accelerators). Our “Client Zero” experience helps clients avoid common pitfalls and accelerate time-to-value; and we implement everything using client-approved tooling/patterns so the department can run and evolve the AI Factory independently.

Outcome orientation: measurable user value, scalability and operational resilience

To ensure AI Factory delivery is not just built, but demonstrably usable, supportable and scalable, we define success measures that reflect what matters in live public services such as:

- **User value and adoption:** Feedback loops and HITL evaluation
- **Cost effectiveness:** Reduced cost per additional use case through reuse
- **Scalability:** Shared production-ready architecture with ongoing monitoring and compliance evidence.



Confidential. This document does not constitute a Call-Off Contract for Services with PricewaterhouseCoopers LLP. Where we are engaged to provide Services, our Services will be governed by a subsequent Call-Off Contract that may be entered into between us. If you receive a request under freedom of information legislation to disclose any information we provided to you, you will consult with us promptly before any disclosure. All information contained in this document or otherwise provided or made available as part of any Award Procedure is confidential and may not be disclosed to anyone else without our prior consent.

© 2026 PricewaterhouseCoopers LLP. All rights reserved. 'PwC' refers to the UK member firm, and may sometimes refer to the PwC network. Each member firm is a separate legal entity. Please see www.pwc.com/structure for further details.