Expertilise
Strategic Delivery • Technology • AI

EXECUTIVE BRIEFING 001

FROM AI MATURITY TO DELIVERED CAPABILITY

What personal AI reveals about delegated autonomy, organisational maturity and the delivery challenge for Programme, Portfolio and Project Management

When systems start making dynamic choices, the delivery challenge shifts to defining the boundaries within which those choices can safely be made.

Brian Kelly
Senior Programme and Project Delivery Leader
August 2026

Executive summary

Meta's Muse Glimmer and OpenAI's Computer History are used here as timely governance stress tests. Glimmer shows capable agentic AI moving onto personal hardware. Computer History shows AI gaining persistent context from a user's computer activity through an opt-in timeline of events. Alongside computer-use agents that can see a screen, click and type, personal AI is moving from answering questions toward observing, remembering and acting within an individual's digital environment.

That exposes a governance asymmetry. An individual may decide what an agent can access, what contextual history it may retain and what it may do, yet have no security team, architecture board, risk function, programme governance or formal operational acceptance process around that decision. It is unrealistic to expect a consumer to reproduce the multidisciplinary governance capability of a mature organisation.

For businesses, the issue is not merely AI adoption. Agentic AI introduces delegated autonomy: systems can increasingly choose steps, invoke tools, act across workflows and escalate selectively to humans. That changes decision rights, access boundaries, approval points, evidence requirements and operational accountability.

The enterprise control environment is developing quickly. Identity and authorisation, runtime orchestration, observability and evidence, data governance and broader AI governance remain distinct control concerns. Major platform vendors are nevertheless beginning to consolidate these capabilities inside their own ecosystems. The enterprise challenge is therefore shifting from simply identifying missing controls to integrating and assuring a heterogeneous control stack.

AI maturity models can help reveal whether the organisation has the strategy, governance, data, technology, workforce and value-management capability to manage this shift. For a typical mid-maturity commercial enterprise, I would start with the MITRE AI Maturity Model as the broad, vendor-neutral baseline, then add an agentic lens where autonomy is material to the portfolio.

But assessment is only diagnosis. Capability gaps have to become investment choices, sequencing decisions, shared enabling capabilities, programmes, workstreams, controls, operational readiness and measurable outcomes. Not every gap deserves a project; it first deserves an investment decision. That is where PPPM becomes strategically important: not as owner of every specialist discipline, but as the mechanism for turning cross-functional maturity gaps into coordinated delivery.

As agent authority expands, the evidence and organisational capability around it should expand as well. PPPM sequences the dependencies, delivery gates and evidence required to implement those decisions safely.

AI maturity tells us whether the organisation is ready.
Agentic delivery governance defines what the AI is actually allowed to do.

What this briefing covers

- Why personal AI moving onto personal hardware and gaining persistent context exposes a governance gap that individuals cannot reasonably be expected to close alone.
- Why delegated autonomy makes agentic AI different from a generic technology maturity problem.
- Which maturity model I would use as the baseline for a typical mid-maturity enterprise and why.
- A real-world delivery parallel from a live global network where dynamic decisions had to operate within defined business boundaries.
- How PPPM translates maturity gaps into decision rights, controls, readiness and measurable capability.
- Why shared foundations such as identity, data access, integration, monitoring and controls should be treated as portfolio assets rather than rebuilt by individual projects.
- How progressive autonomy can be governed through evidence gates, from assistance to constrained action and, where justified, broader authority.
- What poor AI maturity delivery looks like in practice and the actions senior leaders should take now.

1. What Muse Glimmer and Computer History reveal

Muse Glimmer is a compact, open-weight model designed to run agentic tasks on a Mac or PC with a single graphics card. OpenAI's Computer History introduces a different capability: an opt-in ChatGPT feature for macOS that uses activity events such as clicks and keystrokes to build a timeline that ChatGPT and Codex can reference. OpenAI says it does not capture images, video or audio.

Together, they expose two shifts in personal AI. Glimmer moves capable agentic processing onto personal hardware. Computer History gives AI persistent context across a user's digital activity; OpenAI's computer-use capability can then perceive a screen and use a virtual mouse and keyboard to complete multi-step actions. The governance issue is therefore no longer only where the model runs, but what the AI can observe, remember and do. Personal AI makes a hidden assumption in conventional governance visible: we often assume there is an organisation around the technology.

The personal governance gap

An individual can give an agent access to email, files, browsers, applications and connected devices, and increasingly allow AI to accumulate contextual history across those environments. Product safeguards, opt-in controls and operating-system permissions still matter, but there may be no formal security review, architecture decision, risk acceptance, change process or operational owner around the choice.

Just a person, a set of permissions and a decision.

2. The enterprise issue is delegated autonomy

Businesses face a related but different problem. Once an AI system can plan, select actions, use tools and execute multi-step workflows, the key governance question is no longer only whether the model is accurate or safe. It is how much authority the organisation is prepared to delegate to it.

That requires explicit answers to questions traditional maturity discussions can leave implicit: Which decisions may the agent make? Which credentials and systems may it use? Which actions require human approval? Which actions must be reversible? What triggers escalation? What evidence proves the controls work? Who owns the outcome when the agent acts within its authorised envelope?

I use the term autonomy envelope for the defined set of decisions, actions, systems, credentials and operating conditions within which an AI system is authorised to act without further human approval.

For CIOs, this is not simply a procurement decision. Microsoft now provides a recognisable agent control plane across Agent Framework, Agent 365 and Entra, bringing orchestration, identity, inventory, security and governance closer together within its ecosystem. Enterprise estates will remain heterogeneous, so the challenge is increasingly how to integrate, govern and assure controls across vendors, platforms and operating domains. An organisation should not increase agent autonomy faster than the capability needed to govern that autonomy can mature.

Specialist functions own risk appetite and control decisions. PPPM orchestrates how those decisions become sequenced, evidenced and delivered.

Agentic AI turns governance into a question of decision rights: what is the system authorised to decide and do, under what constraints, and who remains accountable?

3. Is the organisation mature enough?

Maturity models are useful because they force the enterprise to look beyond isolated pilots. They test whether the surrounding organisational system can support AI repeatedly, safely and at scale.

For a typical mid-maturity commercial enterprise - active pilots, some central AI capability and uneven governance - I would start with the MITRE AI Maturity Model. It is broad, vendor-neutral and designed around an organisation-specific target state rather than the assumption that every capability must reach maximum maturity.

My starting point: MITRE for the enterprise baseline. Add an agentic maturity lens where delegated autonomy is a material part of the strategy.

Lens	How I would use it	Why
MITRE AI Maturity Model	Default enterprise baseline. Assess the organisation across responsible use, strategy/resources, organisation, technology enablers, data, and performance/application.	Broad enough to expose system-wide capability gaps and flexible enough to set different target levels by business need.
Agentic AI maturity lens (e.g. Microsoft's Agentic AI Adoption Maturity Model)	Overlay where agents are expected to take actions, orchestrate workflows or operate with increasing autonomy.	Makes decision rights, governance, security, operating model, human-agent interaction and value explicit.

Sector-specific or internal models may be better where regulation, mission or enterprise architecture creates distinct requirements. The maturity model is not the point of view. The point of view begins with what the organisation does with the diagnosis, and whether the target state is sufficient to govern the autonomy it intends to deploy.

4. Operator example: governing dynamic decision-making in a live global network

A useful parallel from delivery experience

In a global satellite communications environment, I led a live-network transformation that moved from fixed bandwidth allocations towards dynamic allocation based on business priority and available capacity when the network was under contention.

The technology was not AI, but the delivery problem provides a useful parallel for agentic systems. Once allocation became dynamic, the challenge was no longer simply whether the technology worked. We had to define the business priorities the system should follow, the boundaries within which it could make decisions, how those decisions would be validated in live conditions, and how the change could be introduced safely across multiple business areas.

The programme used proof-of-concept validation, cross-business agreement, phased rollout, hypercare and data-led monitoring in production.

Operator lesson: when a system starts making dynamic choices, the difficult work moves from controlling every individual decision to defining the rules, boundaries, evidence and accountability within which those decisions can safely be made.

5. Autonomy should be earned through evidence

Autonomy does not need to be binary. An organisation can progressively increase what an AI system is authorised to do as operational evidence accumulates. Greater authority should require greater confidence in the surrounding controls, operational readiness and accountability - not simply greater confidence in the model.

Stage	Operating model	Evidence gate
ASSIST	AI supports; the human acts.	Constrained access, baseline quality and known limitations.
RECOMMEND	AI proposes; the human decides.	Recommendation quality, exception patterns and effective human oversight.
ACT WITHIN CONSTRAINTS	AI executes bounded actions inside a defined autonomy envelope.	Controls, monitoring, rollback and operational ownership demonstrated.
EXPAND AUTHORITY	Permissions or decision rights increase where justified.	Sustained value, control evidence, operational readiness and an approved risk decision.

PPPM's contribution is to make evidence requirements explicit, sequence dependencies and ensure the operating capability is ready before authority expands.

Autonomy should be earned through evidence.

6. Assessment is not delivery

A maturity assessment may reveal fragmented pilots, weak data foundations, unclear decision rights, limited workforce capability or poor value tracking. None of those gaps closes because a heat map exists. The assessment becomes useful only when it changes investment and delivery decisions. Not every maturity gap should automatically become a project; it must first become an investment decision.

Fund shared foundations as portfolio assets

Identity, data access, integration, evaluation, monitoring and reusable control patterns may support many AI initiatives. Some platform vendors are beginning to package several of these layers into integrated control planes, but enterprise estates will still span multiple vendors and operating domains. Funding each control independently inside every use case encourages duplication and inconsistent implementation. PPPM can expose the shared dependencies, treat reusable foundations as portfolio assets and make cross-platform integration and assurance explicit delivery work.

Sequence by value and dependency

The highest-profile AI use case is not always the right initiative to fund first. If several use cases depend on the same unresolved data, identity, integration or operating-model capability, portfolio sequencing should reflect both strategic value and the enabling dependencies that unlock the wider portfolio.

ASSESS	TARGET	PRIORITISE	MOBILISE	DELIVER	TRANSITION	MEASURE
Where are the gaps?	What capability is required?	What creates value and unlocks dependencies?	Who owns it, how is it funded, what must happen first?	Build and integrate	Move into live operation	Did value land?

The most expensive maturity assessment is the one that produces a score but no decisions.

7. Where PPPM adds value in an agentic environment

AI or maturity finding	Required organisational response	PPPM contribution
Agent decision rights are unclear	Define the autonomy envelope: what the agent may decide, what requires approval and what is prohibited.	Decision gates, dependency mapping, acceptance criteria and evidence requirements that implement the specialist risk decision.
Tool, credential and data access is fragmented	Create approved access patterns with named owners	Dependency mapping, environment readiness, cross-

	and explicit trust boundaries.	team sequencing and evidence that access controls work.
Autonomy increases without sufficient evidence	Expand authority progressively as control evidence, operational readiness and value justify the next step.	Stage the progression from assist to recommend to constrained action, with explicit evidence gates before authority increases.
Pilots do not scale	Move from bespoke experiments to reusable architecture, shared foundations, controls and delivery patterns.	Portfolio prioritisation, shared-capability investment, rollout waves, supplier coordination and lessons learned.
Operations are not ready for autonomous behaviour	Establish monitoring, incident response, rollback, service ownership and support before autonomy increases.	Operational transition, hypercare, service acceptance and go-live governance.
Value is unclear	Link AI investment to measurable operational, strategic or transformational outcomes.	Benefits baselines, KPIs, portfolio reviews and explicit stop, refine or scale investment decisions.

PPPM orchestrates how specialist decisions, controls and dependencies become one executable portfolio of change and a deliverable operational capability.

8. What this looks like when it fails

Poor AI maturity delivery does not usually fail because the assessment was technically wrong. It fails because the diagnosis never becomes a governed delivery system - or because agent autonomy increases without the surrounding controls maturing at the same pace.

Failure mode	What it looks like	Consequence
Assessment shelfware	A score and heat map exist, but priority gaps have no sponsor, funding path or delivery owner.	The organisation can describe its weakness but cannot show material improvement.
Governance overload	Boards, policies and committees multiply, but teams still lack usable decision rights, evidence requirements or approval thresholds for agents.	More process without clearer control of autonomous behaviour.
Pilot cemetery	Functions run disconnected agent pilots with different platforms, access patterns and success measures.	Duplication grows and few pilots survive the move into production.
Foundation starvation	Visible use cases are funded while shared data, identity, integration, monitoring and control capabilities remain underdeveloped.	Pilots compete for the same weak foundations, duplication grows and autonomy can outrun operational readiness.
Blurred accountability	Everyone approves a component, but nobody clearly owns the integrated decision envelope or operational consequence.	Issues surface late and escalation becomes reactive.
Benefits vacuum	Usage is reported as success without a baseline or stop, refine or scale criteria.	Low-value initiatives persist while higher-value foundations remain underfunded.

9. What senior programme and portfolio leaders should do now

- Define the autonomy envelope before approving an agentic use case: permitted decisions, prohibited actions, systems and credentials, mandatory approvals, escalation thresholds and rollback expectations.
- Choose a maturity model for the decision it must support. For a typical mid-maturity enterprise, use a broad baseline first and add an agentic lens where autonomy is material.
- Set target maturity by capability. The organisation does not need maximum maturity everywhere, but it does need sufficient maturity around the risks and value of the use cases it intends to scale.
- Separate use-case investment from enabling-capability investment. Treat reusable data, identity, integration, evaluation, monitoring and control foundations as portfolio assets where they unlock multiple initiatives.
- Sequence the portfolio by strategic value and dependency, not profile. The most visible AI use case may need to wait until the capabilities that unlock it - and several other use cases - are ready.
- Force every priority maturity gap into one of two states: a funded change item with an owner and evidence of completion, or an explicitly accepted gap. Diagnosis without a decision is not progress.
- Treat increased agent autonomy as a gated production change. Move deliberately from assist to recommend to constrained action and broader authority only when evidence and organisational capability justify it.
- Make operational ownership a scale condition: named service owner, support model, monitoring, incident response, rollback, user adoption and benefits baseline before an agent gains material authority.
- Use stop, refine or scale as real investment decisions. A pilot should progress because value, readiness and evidence justify the next investment, not because the demonstration was impressive.

10. Why this is a timely PPPM opportunity

The project profession is already formalising its role. PMI's 2026 Standard for Artificial Intelligence in Portfolio, Program and Project Management positions AI adoption as work that must move from ad hoc use toward structured and accountable practice.

The opportunity for PPPM leaders is therefore larger than using AI to improve project administration. As platform vendors consolidate agent identity, orchestration, observability and governance inside their own ecosystems, the delivery challenge moves up a level. The strategic role is helping organisations define the target control architecture and translate maturity findings, autonomy decisions and specialist controls into investment choices, dependency-led sequencing, shared portfolio capabilities, cross-vendor integration, approval thresholds, lifecycle controls, evidence requirements, coordinated change, operational readiness and measurable value.

AI for project management is only half the story. PPPM must also help deliver the AI-ready enterprise.

11. The leadership thesis

Personal AI exposes the governance gap. Agentic AI makes delegated autonomy an enterprise design decision. Maturity models help reveal whether the organisation has the surrounding capability to manage that decision repeatedly and at scale. Major platform vendors are beginning to consolidate control layers inside their own ecosystems, but enterprise fragmentation persists at the boundaries between them. AI maturity is therefore not simply a capability question; it is also an investment, sequencing, integration and decision-rights question.

Assessment does not close the gaps. Portfolio management decides what matters enough to fund, which shared capabilities unlock the portfolio and what should stop. Programme management coordinates interdependent change, delivery gates and the evidence required to increase authority safely. Projects build and validate the capabilities. Operations has to own and sustain the result.

Authority should increase only as evidence and organisational capability increase with it.

AI maturity is not achieved by adopting a framework. It has to be delivered.

Scope note

This briefing deliberately focuses on organisational maturity, delegated autonomy and delivery. A future Expertilise briefing will examine the emerging enterprise agent control plane in more detail: where identity, runtime orchestration, observability and governance are converging inside platform ecosystems, where fragmentation remains across vendors, and how PPPM can orchestrate the target architecture, integration, lifecycle controls and assurance required for governed production.

About the author

Brian Kelly is a senior programme and project delivery leader with more than 25 years of experience delivering complex technology, transformation and product change across global and regulated environments, including Accenture and Viasat/Inmarsat. His current focus is the intersection of AI, strategic delivery and organisational transformation: how enterprises move from promising AI capability to governed, operational and measurable outcomes.

Briefings archive: https://briefings.expertilise.com

More background and current commentary: linkedin.com/in/brianpjkelly/

Sources and further reading

MITRE: The MITRE AI Maturity Model and Organizational Assessment Tool Guide | https://www.mitre.org/news-insights/publication/mitre-ai-maturity-model-and-organizational-assessment-tool-guide

Microsoft Learn: Agentic AI adoption maturity model | https://learn.microsoft.com/en-us/agents/adoption-maturity-model/

Microsoft Learn: Microsoft Agent Framework overview | https://learn.microsoft.com/en-us/agent-framework/overview/

Microsoft: Microsoft Agent 365 - The Control Plane for Agents | https://www.microsoft.com/en-us/microsoft-agent-365

Microsoft Learn: Microsoft Entra Tenant Governance | https://learn.microsoft.com/en-us/entra/id-governance/tenant-governance/overview

Microsoft Learn: Microsoft Entra Agent ID - What's new | https://learn.microsoft.com/en-us/entra/agent-id/whats-new-agent-id

Microsoft Learn: Agent identity integration for Copilot Studio | https://learn.microsoft.com/en-us/microsoft-agent-365/builder/identity

OpenAI: Computer-Using Agent | https://openai.com/index/computer-using-agent/

The Verge: ChatGPT's Computer History tracks your clicks and keystrokes, 16 August 2026 | https://www.theverge.com/ai-artificial-intelligence/980742/chatgpts-computer-history-tracks-your-clicks-and-keystrokes

Okta: Okta for AI Agents - General availability and features | https://support.okta.com/help/s/article/okta-secures-ai?language=en_US

NIST NCCoE: Accelerating the Adoption of Software and AI Agent Identity and Authorization | https://csrc.nist.gov/pubs/other/2026/02/05/accelerating-the-adoption-of-software-and-ai-agent/ipd

PMI: The Standard for Artificial Intelligence in Portfolio, Program and Project Management | https://www.pmi.org/standards/artificial-intelligence

Reuters: Meta launches new AI model as Zuckerberg champions open-weight push, 10 August 2026 | https://www.reuters.com/world/china/meta-launches-new-ai-model-zuckerberg-champions-open-weight-push-2026-08-10/

Author note: the MITRE recommendation, autonomy-envelope formulation, progressive-autonomy model and operator parallel are the author's judgement. The live-network example was not AI. Microsoft and OpenAI are included as current illustrations of platform convergence and personal AI capability, not as product recommendations.