



Guidance for AI proof of concept to scale

Benefits realisation for AI occurs through enterprise integration and not through disparate AI experimentation alone. Technical and organisational shifts are required to break down the biggest barriers of scaling up AI for business, enterprise or wider use.

Contents

Contents	1
Overview	3
Context	4
What is an AI proof of concept?	4
What does it mean to scale AI?	4
Principles	6
AI transition stages and dimensions	7
Stages: PoC vs pilot vs production.....	8
Dimension considerations	9
Common challenges and mitigation options	13
Common scenarios – lessons from the challenges	17
Scenario 1 – Co-design and transparency	18
Scenario 2 – Clear expectations	19
Scenario 3 – Cross-agency capabilities	20
Scenario 4 – Gap in user understanding	21
Scenario 5 – Gap in business buy-in.....	22
Scenario 6 – New AI model and legacy systems.....	23
Share your story	24
Appendix 1: AI readiness checklist for sponsors, managers and practitioners	25
1. Before starting a PoC.....	25
2. After PoC completion	26
3. Before scaling (pathway to production)	27
4. Pilot phase (controlled real-world testing and validation).....	27
5. After scaling (operationalisation).....	28
Appendix 2: AI evaluation	29
Appendix 3: Informing procurement	30

Appendix 4: Mapping of dimensions to the Technical standard for government’s use of artificial intelligence33

Overview

The purpose of this guidance is to enable Australian Government agencies to move beyond isolated artificial intelligence (AI) experiments and achieve meaningful business AI adoption, ensuring that AI delivers value and is embedded responsibly within government operations and services. It provides practical advice for successfully scaling AI proof of concepts (PoCs) into production solutions or business operations, bridging the gap between experimentation and enterprise-scale deployment.

By outlining key challenges, success factors and best practices, the guidance supports agencies in maturing their AI practices, aligning AI projects with strategic goals and integrating them into existing operational frameworks to realise lasting public outcomes.

This guidance is designed to complement existing Australian Government frameworks and standards, offering a flexible approach that agencies can adapt to their operational contexts for effective integration across diverse environments. It is designed to help agencies progress the roadmap they have set for AI adoption, applying the right level of component consideration and an appropriate balance of discipline and governance. The pathway to scale - from idea to an operational AI system - typically follows three stages: proof of concept, pilot, and then deployment into production.

Each of these stages involves systematic evaluation to ensure readiness for business integration.

This guidance covers:

- context – defines what it means to scale AI and what constitutes a successful AI PoC that scales up to enterprise deployment
- principles – for sustainable and responsible AI scale-up and adoption
- key considerations across the stages and dimensions of AI transition
- common challenges, mitigation options and project owners – advice to address challenges across business alignment, solution design, data management, technology, people, governance and operations
- common scenarios – illustrating both successes and failures in scaling up AI, offering actionable lessons learned
- tools and resources:
 - an AI readiness checklist
 - guidance for AI evaluation

- guidance for informing procurement decisions¹
- alignment to the Technical Standard for Government's use of artificial intelligence.

Context

What is an AI proof of concept?

An AI proof of concept (PoC) is a focused, small-scale experiment designed to demonstrate technical feasibility and potential business value of an AI use case. Its purpose is to test and validate that an AI technology or model works as intended before committing to full-scale business integration and deployment. A successful PoC involves a clearly defined problem statement, success criteria, appropriate data management and thorough testing of the AI technology and processes.

What does it mean to scale AI?

Scaling AI refers to the process of moving beyond isolated PoCs, pilots or experiments to specific business or enterprise-wide adoption where AI solutions are embedded into business operations. At scale, AI delivers sustained value, is repeatable and is adaptable to evolving needs. Successful scaling requires robust infrastructure, governance and change management, along with early assessment of risks, benefits, impacts, processes and integration pathways to ensure smooth and responsible implementation.

AI PoCs are excellent for early opportunity assessment, but without clear objectives, collaboration and defined pathways, most remain isolated experiments. PoCs often deliver promising early results like functional models or specific insights, and technical feasibility, however many fall short of operationalisation. Challenges include unresolved technical debt, weak sponsorship, resistance, inadequate risk controls or unclear ownership. Without a clear pathway to scale, even technically feasible PoCs risk becoming shelfware rather than business solutions.

At a global level and across industries, it has been observed that most organisations lead in AI PoCs but lag in scaling into business operations. Research shows that the majority of AI projects do not reach production or operations. In the public sector, many agencies remain in exploratory phases, with limited AI initiatives successfully scaled. Findings suggest the need

¹ [Buy products and services - BuyICT](#)

for AI to integrate with existing enterprise systems and continuous improvements made over time. High performing organisations consistently report measurable benefits, including cost savings and operational efficiencies, by embedding AI in business and aligning initiatives with strategic priorities.

Common challenges in AI PoCs that fail to scale include:

- unclear success criteria – no defined measures of value or impact
- technology-led approaches – lack of product thinking or user-centric design
- single-model or technique dependency – testing only one or insufficient number of AI models or techniques and declaring failure without exploring multiple options
- strategic misalignment – disconnected from agency or business priorities
- incomplete value framing – evaluations focus on investment vs outcome, ignoring current operating costs, existing service value, environmental impacts
- undefined cost-benefit expectations – no clear understanding of investment vs return
- duplication of effort and costs – siloed activities and lack of visibility
- weak data and privacy governance – poor practices undermine compliance and scalability
- process vs service confusion – automating flawed processes without rethinking the underlying service design for better opportunities will just speed up poor processes
- inflexible ICT processes – AI development is forced through traditional ICT pipelines, limiting agility and innovation
- insufficient stakeholder engagement – ignoring the importance of adequately engaging relevant stakeholders including business, technical and end-user groups
- risk avoidance mindset – excessive caution prevents experimentation, learning and iteration, stalling progress before value can be demonstrated.

To move from experimentation to implementation, agencies must embed AI initiatives within broader strategic, operational and governance frameworks. This includes defining ownership, aligning with business goals, and planning for long-term sustainability from the outset.

Agencies that scale AI from PoCs to sustained levels for business outcomes share key traits and requires both system-level coordination and evidence-based decision making. They include the ability to:

- align AI initiatives with strategic goals and public outcomes
- clear accountability and decision gates
- incorporate product thinking and user-centric design early
- use modular, reusable infrastructure and ensure robust governance

- utilise high-quality data across the initiative
- develop and retain their intellectual property
- mitigate vendor lock-in
- apply agile practices (e.g. ModelOps²) to support rapid iterative development with controls and safety embedded
- maintain continuous evaluation³ and monitoring from the outset
- implement change management processes for adaptability
- foster collaboration across functional teams
- invest in long-term capabilities and governance to support responsible innovation leading to business outcomes
- encourage learning, sharing, and reuse to support ongoing improvement and integration of innovation.

Principles

To successfully scale AI from PoC to enterprise-wide impact, organisations must go beyond technical experimentation and establish the right strategic, operational and governance foundations. The following principles are essential for enabling sustainable, scalable AI adoption:

- 1. Strong foundations:** Core capabilities (data, talent, tools and processes) are in place and actively maintained to support AI at scale. These foundations enable composability, allowing components to be reused and adapted. Agencies should ensure they retain their intellectual property ownership and maintain control over assets.
- 2. Enterprise-ready design and infrastructure:** Solutions are built with scalability, interoperability and operational resilience in mind, supporting sustained investment in technical architecture and design.
- 3. Robust governance and trust frameworks:** Clear policies, risk controls, accountability and ethical safeguards guide responsible AI use, with mechanisms for oversight and adaptation, and shared governance across domains. Ensure compliance with the AI in government policy.

² Definition of Artificial Intelligence Model Operationalization (ModelOps) - Gartner Information Technology

[Glossary](#)

³ Commonwealth Evaluation Policy | evaluation.treasury.gov.au

4. **Cross-functional collaboration and accountability:** Alignment across technical, business, legal and operational teams ensures shared ownership, accountability and long-term sustainability of AI initiatives.
5. **Strategic alignment and measurable outcomes:** AI initiatives are tied to business priorities, with defined success metrics and baselines, systematic evaluation methods and pathways to value.
6. **Culture of responsible innovation and business value:** Foster a culture that embraces continuous learning, ethical innovation and business impact. Encourage experimentation with measurable value upholding responsible AI practices.
7. **AI Literacy at all levels:** Promote AI literacy across the agency (from staff to leadership) ensuring a shared understanding of AI concepts, opportunities, risks and responsible practices.
8. **Right technology for the right problem:** Technology choices are driven by the business problem to be solved – not by novelty or trend. Solutions should be fit-for-purpose, cost-effective, aligned with agency goals and designed to mitigate vendor lock-in for interoperability and flexibility. Introducing new technology may prompt examining business processes to improve and modernise services.

AI transition stages and dimensions

The pathway from conceptualisation to AI system implementation is typically a three-stage process involving a PoC, pilot and then release into production or the operational environment. In the Australian Government context, this may involve (for example):

PoC: An agency builds a generative AI tool using historical case files to test whether AI can effectively summarise long policy documents, identify relevant points and provide decision support for staff. Depending on the objectives, the PoC boundaries must be set appropriately – whether in a constrained business environment or a constraint-free setting.

Pilot: The AI tool is deployed to a small group of policy officers in one division, using redacted real documents, with feedback collected on accuracy and usability.

Production: The AI tool is rolled out agency-wide, fully integrated with secure document repositories, supported by IT, and governed under AI Assurance and the Digital Service Standard.⁴

In practice, multiple PoCs may be conducted to experiment with distinct components of a broader initiative, forming a composite foundation for implementation. The components may include technical evaluation and business process fit or validation. Several PoCs may also be conducted in parallel to test various aspects of the evaluation area of focus i.e. open source versus AI proprietary technology type.

All activities that form part of the pathway to production need to comply with all relevant Australian laws and legislation.

Stages: PoC vs pilot vs production

The table below provides a comparative definition of each stage whilst also identifying the various aspects that could be considered in implementation design and ongoing solution sustainment.

	PoC	Pilot (transit to scale)	Production (scale)
Definition	Early, experimental build to test concepts, feasibility or specific technical components. Short time frame, budget constrained.	Limited-scale, real-world implementation to validate value, usability and readiness. Budget constrained and timeframe specified.	Fully deployed, enterprise-scale system integrated into business-as-usual operations
Purpose	Explore ideas, prove feasibility and learn quickly	Validate solution in controlled conditions before scaling	Deliver consistent, reliable, compliant service at scale
Users	Internal teams, stakeholders or test group	Real users (staff, citizens) in a limited or controlled setting	All intended end-users (public, agency-wide, cross-government)
Metrics	Technical feasibility, functional performance, scientific and empirical test metrics	User satisfaction and feedback, business operation impact	Operation performance, business operation measures, ongoing monitoring
Data	Use of synthetic data due to privacy and data governance risk	Live or near-live real data with governance and safeguards	Full live production data under governance frameworks
Security	Foundational, meets data, infrastructure and IP requirements	Controlled and increased oversight	Full compliance, robust and ongoing monitoring

⁴ [Digital Experience Policy](#)

Infrastructure	Light-weight or sandbox environment	Core processing and integration fully tested under high volumes	Full enterprise integration
Integration	Minimal or stand-alone; often mocked or simulated	Partial/controlled integration with legacy systems or cloud services	Full enterprise integration with other government systems and services
Investment	Low (time, cost, compliance)	Medium (infrastructure, governance, partial training)	High (enterprise infrastructure, full compliance, long-term sustainability)
Governance	Light-touch; risk-based exemptions allowed; compliance with PSPF, ⁵ ISM, ⁶ AI Assurance. ⁷ Data governance and data-type considerations must be addressed to meet security and privacy requirements.	Increased oversight, ethics and assurance processes; compliance with PSPF, ISM, AI Assurance. Strengthened data governance controls for increased scale of sensitive data and information exposure.	Full compliance with PSPF, ISM, AI Assurance, procurement and audit standards. Robust governance and incident management for operation-wide exposure to data and information.
Risk tolerance	High – failure is acceptable and expected. Expectation that PoC clearly states assumptions.	Moderate – failure is still possible but must be controlled.	Low – failures must be rare and managed through continuity planning.
Outcome	Proof of feasibility – decide if idea has merit	Proof of viability – decide whether to scale to production	Operational service – provides measurable, sustainable outcomes
Lifecycle	Disposable, often short-lived	Transitional – can be iterated or scaled	Long-term – maintained, supported and continuously improved
Procurement (if applicable)⁸	Short-term contract, vendor checks, performance evaluation, IP ownership, auditable purchase orders	Partial compliance meeting all mandatory requirements for the controlled setting, including formal purchase orders	Full compliance, long-term contract, open standards, formal exit planning

Dimension considerations

A critical part of ensuring PoC to scale success includes thorough and detailed planning. The framework laid out below is intended to assist agencies with a pro-active approach to planning. It considers the key dimensions of transition from system inception to ongoing sustainment.

⁵ [Protective Security Policy Framework](#)

⁶ [Information Security Manual](#)

⁷ [National framework for the assurance of artificial intelligence in government](#)

⁸ [Public Governance, Performance and Accountability Act 2013](#)

Dimension	Purpose	Activities	PoC vs pilot vs production
Business & strategic alignment	Ensures the PoC addresses real business problems, demonstrates value, ⁹ and aligns to agency and government priorities. Ensure alignment to the policy and legislative requirements of the agency.	<p>Define the problem statement and scope.</p> <p>Set measurable success criteria (KPIs, citizen outcomes).</p> <p>Secure executive sponsorship and budget alignment.</p> <p>Link to organisational strategy or policy outcomes.</p>	<p>PoC: Validate hypothesis quickly, success measured qualitatively.</p> <p>Pilot: Test impact on a controlled group with quantitative measures.</p> <p>Production: Embedded into BAU, outcomes tracked against agency strategy.</p>
Architecture & solution design	Strong design ensures scalability, security and interoperability. Poor design creates technical debt and lock-in.	<p>Define architecture boundaries and assumptions.</p> <p>Align to enterprise/government architecture frameworks.</p> <p>Assess open-source vs proprietary balance.</p> <p>Plan scalability and resilience pathways.</p>	<p>PoC: Lightweight, often disposable architecture.</p> <p>Pilot: Limited interoperability, partially hardened.</p> <p>Production: Enterprise-ready, compliant with reference architecture.</p>
Data & integration	Ensure data quality, sovereignty and compliance requirements are met. Identify and manage interoperability risks.	<p>Assess data quality, lineage and provenance.</p> <p>Implement metadata, cataloguing and governance controls.</p> <p>Ensure application programming interfaces (APIs) and integration pathways are reusable.</p> <p>Meet data and security requirements.</p>	<p>PoC: Basic integration.</p> <p>Pilot: Controlled data subsets, partial integration.</p> <p>Production: Full compliance, lineage, interoperability across systems.</p>

⁹ Benefits Management Policy | digital.gov.au

<p>Technology & tools</p>	<p>Ensure the right platform and management tools are selected and deployed. Choices influence speed, compliance and sustainability.</p>	<p>Select cloud/on-prem consistent with hosting strategies.</p> <p>Leverage containerisation.¹⁰ Streamlined software development process Operational support tooling.</p>	<p>PoC: Use lightweight environments.</p> <p>Pilot: Partially hardened environments.</p> <p>Production: Approved infrastructure and tooling, Meet all security requirements.</p>
<p>People & skills</p>	<p>Ensure that the PoC has the right mix of resources with technical, business, policy and change management skills.</p>	<p>Identify skill gaps and training needs.</p> <p>Involve cross-functional teams (policy, legal, ICT, data, end-users).</p> <p>Build stakeholder engagement plans.</p> <p>Foster capability uplift alongside delivery.</p>	<p>PoC: Small focussed team.</p> <p>Pilot: Cross-functional engagement expands.</p> <p>Production: Full capability planning and organisational ownership.</p>
<p>Governance & risk</p>	<p>Ensure the implementation has the right level of security, compliance and accountability. Skipping governance exposes agencies to reputational and operational risk.</p>	<p>Apply risk assessments early (security, ethics, privacy).</p> <p>Incorporate auditability and approvals.</p> <p>Reference AI ethics principles and assurance frameworks.</p>	<p>PoC: Risk register.</p> <p>Pilot: Compliance and all required approvals.</p> <p>Production: Full governance and accreditation.</p>
<p>Experimentation & validation</p>	<p>To build evidence for scaling decisions and manage risk identified through controlled learning.</p>	<p>Define clear goals and hypotheses.</p> <p>Use rapid experimentation and A/B testing.</p> <p>Run performance benchmarks and user feedback loops.</p> <p>Validate against business success criteria.</p>	<p>PoC: Experiments. Scientific testing. Foundational performance testing.</p> <p>Pilot: Structured user validation. Wide performance testing.</p> <p>Production: Controlled rollouts with full monitoring.</p>

¹⁰ [Guidelines for system hardening | Cyber.gov.au](https://www.cyber.gov.au/guidelines-for-system-hardening)

<p>Delivery & operations</p>	<p>Defines the processes and ensures efficiency when scaling.</p>	<p>Use agile methods with a clear sprint cadence.</p> <p>Track resourcing and budget constraints.</p> <p>Manage vendor/partner contributions.</p> <p>Implement DevOps¹¹ for operational readiness.</p>	<p>PoC: Delivery processes suitable to ensure adequate testing.</p> <p>Pilot: Agile delivery with a governance layer.</p> <p>Production: Full-scale operational delivery and service management. Ongoing monitoring and performance evaluation post deployment.</p>
<p>Scalability & transition to production</p>	<p>Ensures solutions will handle large-scale use, operational continuity and enterprise integration.</p>	<p>Conduct performance/load testing.</p> <p>Implement observability, monitoring and incident response.</p> <p>Prepare change management and training.</p> <p>Ensure business continuity and disaster recovery.</p>	<p>PoC: No scale requirements.</p> <p>Pilot: Limited scale and partial monitoring.</p> <p>Production: Fully scalable, supported, and monitored system.</p>
<p>Sustainment & exit strategy</p>	<p>Prevents stranded investment and ensures long-term maintainability.</p>	<p>Plan for ownership, handover and knowledge transfer.</p> <p>Define funding models and cost transparency.</p> <p>Document exit/decommissioning strategy.</p> <p>Ensure decision registers are retained and that lessons learned are recorded and stored as artefacts</p>	<p>PoC: Disposable or maybe adapted to a future initiative.</p> <p>Pilot: Transition planning begins.</p> <p>Production: Full lifecycle sustainability embedded.</p>
<p>Non-AI considerations</p>	<p>Ensures that agencies assess whether AI is the right solution by evaluating non-AI alternatives such as rules-based automation, process redesign or system optimisation. Prevents unnecessary AI use, reduces cost and</p>	<p>Assess whether the problem can be solved through simpler or established approaches (business process redesign, workflow optimisation, rule-based engines).</p> <p>Conduct root-cause analysis to ensure the problem is not due to</p>	<p>PoC: Validate whether AI is required by comparing outcomes of non-AI approaches.</p> <p>Pilot: Where non-AI elements are part of the solution, test blended/hybrid workflows alongside AI components.</p> <p>Production: Embed the chosen hybrid or non-AI solution into BAU if it outperforms AI for specific parts of the workflow; ensure AI is only</p>

¹¹ [What Is DevOps?](#)

	aligns with responsible innovation principles.	<p>poor processes or data quality.</p> <p>Evaluate potential benefits of non-AI interventions (cost, risk, time to value).</p> <p>Engage business, ICT and policy stakeholders to compare options before committing to AI.</p> <p>Document rationale for choosing AI, hybrid or non-AI approaches.</p>	applied where it adds measurable value.
AI types & fit-for-purpose selection	Ensures agencies select the most appropriate AI technique for the problem, improving effectiveness, cost-efficiency and scalability. Supports transparent decision-making, reduces over-engineering and aligns with the AI Technical Standard .	<p>Define the AI functional need (e.g. classification, summarisation, forecasting, search, anomaly detection, entity extraction, generation).</p> <p>Compare AI techniques: machine learning (ML), natural language processing, deep learning, knowledge-based systems, retrieval-augmented generation (RAG), fine-tuned LLMs, open-source vs proprietary models.</p> <p>Assess model complexity vs business need.</p> <p>Evaluate implications for data, risk, explainability, infrastructure, cost and performance.</p> <p>Document why a specific AI technique or stack was chosen and alternatives ruled out.</p>	<p>PoC: Experiment with multiple AI types/architectures to find the most viable (e.g. baseline ML vs LLM, rules vs hybrid).</p> <p>Pilot: Validate the selected AI type with real or near-real data, testing constraints like explainability, latency, reliability and performance variation across user groups. Test that it meets enterprise requirements for resilience, integration, governance and cost-effectiveness.</p> <p>Production: Maintain model versioning and benchmarking across different AI types over time.</p>

Common challenges and mitigation options

The table below identifies the common challenges for AI PoCs across various implementation dimensions, whilst also suggesting potential mitigation options and a nominal project owner. Incorporation of the mitigation suggestions will assist in managing project risk,

and allocating ownership will strengthen implementation accountability. In some agencies suggested owners may wear multiple project hats.

Dimension	Common challenge	Mitigation option	Potential owner
Business & strategic alignment	Chasing technology hype without business case	Require a problem statement and value hypothesis approved by business owners before funding the PoC.	Accountable Official/Chief AI Officer
	Undefined or unrealistic success measures	Establish SMART KPIs with baselines and success criteria tied to policy or service outcomes.	Program/Project Manager
	Lack of senior stakeholder buy-in	Secure executive sponsor and governance body endorsement before starting for ownership and accountability.	Accountable Official/Chief AI Officer
Architecture & solution design	Ignoring interoperability and integration standards	Align to the AI Technical Standard and reuse existing APIs where possible.	Enterprise Architect
	Over-engineering or under-engineering scope	Define scope boundaries and assumptions and identify the potential for scalability during design.	Solution Architect
	Accumulating technical debt without a transition plan	Maintain a technical debt register with planned remediation.	Technical Lead
Data & integration	Using poor quality data	Assess data quality, lineage, provenance and completeness early. Implement data validation checks, metadata standards, data cleansing processes and establish governance controls before use in PoCs.	Data Custodian/Data Architect
	Using real sensitive data prematurely	Use synthetic or anonymised datasets until governance approvals are in place.	Data Custodian
	Poor integration planning, duplication and silos	Conduct an integration impact assessment with data architects before building.	Enterprise Architect
	Ignoring privacy, retention and sovereignty obligations	Apply PSPF, ISM, and Privacy Act requirements from the outset and get data custodian sign-off.	Privacy Officer/Data Custodian

Technology & tools	Vendor lock-in due to short-term tooling	Choose open standards and modular architectures; and include exit clauses in procurement contracts.	Procurement Lead/Enterprise Architect
	Skipping security/compliance reviews for rapid setup	Run early “security by design” reviews and integrate ISM compliance testing into CI/CD.	Security Lead
	Over-investing in disposable PoC tech	Use low-cost sandbox or cloud credits and set “disposable tech” rules for PoCs.	Project Manager
	Gap between business needs and technology solution	Ensure all business requirements are validated against a requirements traceability matrix during implementation.	Business/Enterprise Architect
People & skills	Excluding key end-users or policy teams	Use co-design workshops and embed frontline staff in sprint reviews.	Product Owner/Project Manager
	Reliance on contractors without knowledge transfer	Require contractors to produce documentation, training and shadowing arrangements.	Program/Project Manager
	No workforce strategy for scaling	Align to relevant APS Workforce Strategy and develop a skills uplift plan early.	Human Resources/Workforce Planner
	Not considering workforce process redesign	Map the current and future business processes and redesign work function and roles to meet new requirements.	Business Analyst/Workforce Planner
Governance & risk	Treating governance as an afterthought	Set up a lightweight governance board with risk and compliance representation for the PoC.	Governance Lead
	Ignoring probity in vendor selection	Use Commonwealth Procurement Rules, AI Procurement Guideline and run probity checks even for small-scale PoCs.	Procurement Lead
	Overlooking mandatory standards (e.g. PSPF, ISM)	Integrate APS AI assurance checklists into project planning from the start. Conduct an impact assessment using the AI Impact Assessment Tool .	Risk & Compliance Officer
Experimentation & validation	Failing to define success/failure upfront	Document hypotheses and thresholds before running tests.	Product Owner
	Skipping user testing	Involve user representatives and frontline staff in validation cycles.	UX Lead/Product Owner

	Missing empirical and scientific testing	Utilise traditional and novel model evaluation techniques.	Technical/Test Lead
	Using PoCs as de facto production without validation	Create “no production use” guardrails in governance, with approval required before scaling.	Governance Board
	Scope of PoC is not valid once the solution is scaled for production	Set threshold performance levels to be met under scaling. Ensure the target levels are met during implementation testing.	Technical Lead/Project Manager
Delivery & operations	Overly rigid waterfall approaches	Adopt agile with sprint demos and incremental delivery.	Delivery Manager
	Ignoring resourcing realities	Prepare a realistic resourcing model including contingency staff.	Project Manager/ Workforce Planner
	Weak vendor management	Assign a vendor manager role and use performance-based milestones and Service Level Agreement (SLA) targets.	Vendor Manager
	Inadequate evaluation and monitoring	Define threshold service level expectations. Ensure best practice tools are deployed to capture measurement data.	Delivery Manager/ Operations Manager
	Unclear roles and responsibilities	Develop a RACI that clearly maps owners and accountabilities. Review and update RACI post-production release.	Delivery Manager/ Operations Manager
Scalability & transition to production	Underestimating performance needs	Run load testing, chaos testing, and scalability simulations before scaling.	Technical Lead/Test Manager
	Ignoring service continuity planning	Develop Incident Management and Business Continuity playbooks and Plans.	Ops Lead/Delivery Manager
	No ownership identified for long-term operations	Define operational handover and embed ownership in BAU teams.	Service Manager/ Agency Service Owner
Sustainability & exit strategy	No budget for scaling beyond PoC	Secure forward funding approvals tied to success metrics and ongoing evaluation. Ensure budget for continuous performance review and enhancement throughout the AI system’s lifecycle.	Agency Service Owner/Finance Lead

	Loss of knowledge when contractors leave	Require a knowledge transfer plan with embedded APS staff.	Project Manager/ Service Manager
	Ignoring lifecycle end-of-service planning	Include decommissioning/archiving as part of the PoC closure checklist.	Project Manager/ Governance Lead
Non-AI considerations	Assuming AI is required when simpler solutions exist	Assess whether the problem can be solved through non-AI approaches such as process redesign, workflow optimisation, business rules engines or system configuration before committing to AI.	Business Owner/Policy Lead
	Overlooking root-cause issues in current processes	Conduct root-cause analysis to determine whether the issue stems from poor processes, insufficient data quality or legacy system constraints rather than a need for AI.	Business Analyst/Service Designer
AI types & fit-for-purpose selection	Over-engineering with unnecessarily complex AI models	Start with the simplest viable model and progressively increase complexity only if required.	AI Scientist/ Technical Lead
	Ignoring explainability, risk and compliance requirements	Assess explainability requirements, risk tier, latency constraints, hosting needs and governance implications when selecting the AI type.	Risk & Compliance Officer/ Technical Lead
	Insufficient evaluation of alternative AI methods	Compare multiple AI types during PoC to identify suitable approach. Document reasons for selecting each.	Product Owner

Common scenarios – lessons from the challenges

The journey from AI experimentation to enterprise-scale adoption is rarely linear. Agencies across the public sector are navigating a range of challenges and opportunities as they explore the use of AI to improve services, decision-making, and operational efficiency. The following real-world scenarios illustrate both successes and setbacks - highlighting what works, what does not, and why. These examples offer practical insights into the importance of co-design, clear expectations, stakeholder engagement, and the integration of AI with broader organisational strategies and systems.

Scenario 1 – Co-design and transparency

An agency set out to develop an AI-powered virtual assistant to help citizens navigate complex government services. From the outset, the project team prioritised inclusive co-design, engaging a diverse group of stakeholders including frontline staff, citizens from different cultural and linguistic backgrounds, and accessibility advocates.

Through workshops and user testing, the team identified key needs – such as support for diverse accents, simplified language, and clear pathways for users with low digital literacy. These insights directly informed the design of the virtual assistant, which was built using a fine-tuned language model grounded in approved policy content.

To ensure the virtual assistant could move beyond a successful PoC and scale across the agency, the team embedded several critical practices from the start. The PoC was strategically aligned with a clear business priority - improving citizen access to government services - and measurable outcomes were defined in partnership with business owners. The solution was architected for scale, using modular components and open standards to support future integration and avoid vendor lock-in. Privacy, security, and ethical considerations were addressed early, with oversight from a governance board and compliance with the Australian Government's AI policy and technical standards. A clear plan was established for data governance and integration as the project moved towards production. Executive sponsorship was secured before the PoC commenced, and regular updates kept leadership engaged, ensuring ongoing support for scaling. The team also planned for workforce capability uplift and change management, recognising that successful scaling would require new skills and ongoing support for staff. Feedback loops and performance monitoring were built in from the pilot phase, enabling the team to adapt the solution based on real-world use and user feedback.

The result was a virtual assistant that improved service accessibility and user satisfaction through transparent, traceable responses. The project was recognised as a model for ethical and inclusive AI design, and the co-design approach has since been embedded into the agency's broader digital transformation strategy.

What worked well	Lessons learned
<ul style="list-style-type: none"> • Early and inclusive co-design with frontline staff, culturally diverse users, and accessibility advocates ensured the solution reflected real-world user needs • Clear strategic alignment with a business priority – improving access to government services – and jointly defined measurable outcomes • Solution engineered for scale using modular components, open standards, and early attention to privacy, security and ethical safeguards • Strong executive sponsorship and continuous leadership engagement sustained momentum • Investment in capability uplift and structured change management supported readiness for scaling • Embedded feedback and monitoring loops enabled iterative improvement during pilot testing 	<ul style="list-style-type: none"> • Without continued resourcing and capability uplift, scaling risks stalling – even with a strong PoC • Scaling requires sustained alignment across business, ICT, legal, policy and operational teams; fragmentation risks slow-downs • Even strong co-design efforts should be maintained through production stages to avoid drift away from user needs

Scenario 2 – Clear expectations

An agency launched an AI PoC to automate the classification of incoming correspondence. Before development began, the team worked closely with business leaders to define clear success criteria, including measurable improvements in processing time, accuracy, and staff workload reduction.

A shared understanding of the project’s goals, benefits, and limitations was established early. The PoC was framed as a technical experiment and a strategic initiative aligned with the agency’s broader digital transformation agenda.

The PoC delivered strong results, exceeding performance benchmarks and demonstrating clear return on investment. Because expectations were well-managed and outcomes were communicated in business terms, the leadership team quickly approved the transition to pilot. The transition to pilot included further development of the technology, rigorous validation, integration with existing enterprise systems, and the development of operational processes for ongoing monitoring and improvement.

Meeting sponsor expectations including results achieved, risk controls and governance requirements, the solution was approved for production. The solution is now being scaled across multiple departments, with strong executive sponsorship and a roadmap for continuous improvement.

What worked well	Lessons learned
<ul style="list-style-type: none"> • Success criteria were clearly defined upfront – processing time, accuracy, and workload reduction – providing a shared baseline • The PoC was positioned not just as a technical experiment but as a key enabler of broader digital transformation goals • Strong results were effectively communicated in business language, supporting fast sponsor approval to move to pilot • Rigorous pilot-stage validation included integration with enterprise systems, operational readiness processes, and ongoing monitoring • Meeting risk, governance and quality expectations gave confidence in decision-making and cleared the path to production 	<ul style="list-style-type: none"> • Without early and explicit framing of limitations (e.g., edge cases, data quality constraints), expectations could easily inflate • Business owners required sustained engagement to avoid misalignment between early technical promise and real-world operational complexity

Scenario 3 – Cross-agency capabilities

A cross-agency taskforce was established to drive improvements in government services, identifying AI technologies as an opportunity for real-time monitoring and feedback, such as triaging email responses and analysing social media sentiment. To ensure the solution’s feasibility, multiple PoC projects were conducted, rigorously testing large language models (LLMs), data integration approaches, infrastructure, and security requirements. Executive sponsors from across government were engaged from the outset, with transparent planning and upfront approval of expert involvement and time commitments.

Insights gained from the PoCs directly shaped the solution’s requirements, design, and procurement strategy. Once validated, the solution was developed and piloted with end-user business groups to refine its effectiveness. In production, the system incorporated robust evaluation and monitoring frameworks, seamlessly integrating with enterprise reporting systems. This enabled sponsors and operations teams to proactively identify and address issues, ensuring the solution’s ongoing success and responsiveness.

Strategic collaboration, PoC with clear objectives, and transparent planning with early engagement of executive sponsors were key facets for the cross-agency AI initiative. Real-world validation, robust evaluation and monitoring frameworks, integrated with enterprise reporting systems, ensured the solution was both technically sound and operationally effective.

What worked well	Lessons learned
<ul style="list-style-type: none"> • Clear objectives and cross-agency buy-in enabled shared ownership of the problem and its solution • Multiple PoCs were used to test LLMs, data integration, infrastructure, and security before committing to scale • Early engagement with executive sponsors secured consistent endorsement and resourcing across agencies • Integration with enterprise reporting systems allowed real-time monitoring, issue identification and proactive adjustment • Evaluation and monitoring frameworks ensured the final system was technically sound and operationally viable 	<ul style="list-style-type: none"> • Cross-agency coordination requires up-front planning of roles, responsibilities and resourcing commitments • Delays can arise when agencies operate at different levels of digital maturity or have different risk appetites • Shared infrastructure and governance models must be agreed early, otherwise integration challenges multiply

Scenario 4 – Gap in user understanding

An agency is building an AI Chatbot to route clients through to the right digital service. A PoC team is created with a small internal technical team and a vendor partner. Citizens, frontline staff and diversity advocates are not consulted in the early design. The resultant chatbot in the pilot phase, and before business deployment, does not recognise diverse accents and does not provide easy to use pathways for those with a low level of digital literacy.

The chatbot uses a LLM that is finetuned using the agency FAQs. When asked by a prospective end-user during the pilot stage “why do I need to provide this document to support my application for assistance?” the chatbot provides a vague answer rather than citing the underlying policy or legislation. Citizens could feel they are being given a “black box” answer without any clear traceability. Advocacy groups start to raise concerns about fairness and accountability.

Inclusive co-design is important for AI solutions to meet real-world needs, as excluding end-users leads to poor accessibility and usability. Early and ongoing engagement with stakeholders helps identify and address potential issues before deployment, preventing public criticism and reputational risk. Ultimately, embedding user-centric design principles throughout development ensures that AI solutions are usable and deliver value to all intended users, demonstrating that technical success alone is not enough for meaningful impact.

What worked well	Lessons learned
<ul style="list-style-type: none"> • The pilot highlighted accessibility and explainability issues before the system reached the public • The shortcomings prompted a shift toward user-centric design principles and stronger governance oversight 	<ul style="list-style-type: none"> • Lack of early engagement with citizens, frontline staff and diverse user groups resulted in a chatbot unable to recognise diverse accents or support users with low digital literacy • The model produced vague, non-traceable answers, leading to perceptions of a “black box” system • Failure to embed explainability from the start led to concerns from advocacy groups around fairness, transparency and accountability • Technical success is insufficient – human-centred design and policy traceability must be embedded from the outset

Scenario 5 – Gap in business buy-in

An agency’s data science team developed an AI solution to search the internet and identify corporations that have provided information that may indicate questionable business registration. The data science team presented the PoC to the leadership team and the project stalled.

The leadership team was sceptical. They did not fully grasp how the model worked and were hesitant to trust an AI to support decision-making. The traditional methods were well-understood and provided a sense of control.

The business leaders were worried, and this was showed through a lack of trust, risk aversion and misaligned goals. Ultimately, a lack of a strong business case combined with general lack of understanding and trust meant the PoC was never given the green light for production.

Agencies should engage stakeholders early and often, and support transitions with clear change management strategies. Alignment with strategic priorities for senior leaders to champion and a clear governance structure for oversight and risk controls established from the outset contribute to the key learnings.

What worked well	Lessons learned
<ul style="list-style-type: none"> • The PoC revealed a critical oversight: leadership did not understand the AI solution. This early discovery prevented costlier failures at scale. • It emphasised the need for strong business case alignment as a core component of AI delivery. 	<ul style="list-style-type: none"> • Leadership scepticism, lack of understanding, and misaligned goals halted progress despite technical capability • Traditional processes felt safer to the business, highlighting insufficient change management and communication about the AI solution's role • A compelling business case and early, deliberate trust-building are important for executive endorsement • Technical teams alone cannot carry an AI initiative – success requires business ownership and clear governance

Scenario 6 – New AI model and legacy systems

A service delivery function achieved an impressive 85% accuracy in triaging citizen complaints and feedback. The AI team was ready to deploy the model, but they hit a wall – the agency's IT existing infrastructure.

To move to production, the model needed a stable deployment environment that could be easily updated and monitored, but the traditional IT change management process required extensive documentation and manual approvals for even minor code changes, creating a bottleneck for a model that needed frequent training and updating.

The project stalled as the IT division, already strained with other priorities, could not dedicate the resource to modernise the infrastructure and change management processes to support the AI model.

Agencies should treat AI development and IT modernisation as a single, unified effort. Establish a governance framework from the start that includes both data science and IT teams. The framework should define clear pathways for data access and model deployment, allowing continuous feedback.

What worked well	Lessons learned
<ul style="list-style-type: none"> • The team achieved strong model performance, demonstrating clear value potential • Early identification of infrastructure constraints prevented rushed deployment into an unsuitable IT environment 	<ul style="list-style-type: none"> • The solution could not progress because IT infrastructure and change management processes could not support modern AI deployment cycles • Manual approvals and traditional processes created bottlenecks inconsistent with the need for frequent model updates • Competing IT priorities hindered timely progress, highlighting the need for coordinated planning • AI development and IT modernisation should be treated as a unified program, with shared governance, shared resourcing, and early alignment on deployment pathways

Share your story

Agencies across government are encountering similar opportunities, challenges and patterns as they explore how to responsibly embed AI into services and operations. Sharing lessons learned – successes and setbacks alike – strengthens the entire system and accelerates uplift across the APS.

If your agency has an experience, challenge or insight to contribute, we encourage you to share your story. Reflecting openly helps support better decision-making and enables agencies to navigate the complexity AI brings. To share your story, please contact ai@dta.gov.au

Appendix 1: AI readiness checklist for sponsors, managers and practitioners

This checklist is designed to guide agencies in maturing their AI practices. It supports confident, responsible adoption by outlining key considerations at each major stage of the AI journey – from initial proof of concept (PoC) to enterprise-scale deployment. By working through these questions, agencies can strengthen strategic alignment, reduce common risks, and establish the foundations for sustainable and impactful AI innovation.

1. Before starting a PoC

Strategic alignment

1. Is the PoC linked to a clear business priority or public outcome?
2. Have success criteria and measurable outcomes been defined?
3. Is there a clear understanding of expected costs and benefits?
4. Has the [Benefits Management Policy](#) been followed?
5. Have approvals been provided by the [Accountable Official](#) and the [Chief AI Officer](#)?
6. Has the proposed PoC been checked for reuse opportunities and reducing duplication via the [agency's internal use case register](#)?

Design & approach

7. Is AI the right solution? If so, is the selected type of AI fit-for-purpose?
8. Can AI realistically deliver the intended outcome?
9. Is this a novel use case or has it been done before?
10. Has product thinking and user-centric design been embedded from the outset, and is there a baseline understanding of user needs and current workflows?
11. Are stakeholders (business, IT, legal etc.) engaged early for co-design and transparency?
12. Has the business been consulted and is there a principal Business Owner involved that can ensure proposed outcomes are validated?
13. Does the approach align with the [Technical standard for government's use of artificial intelligence](#)?
14. Are there any workforce implications/considerations that need to be managed?
15. Has a decision register been established?
16. Has a lessons learned log been established?

Data & infrastructure readiness

17. Is the required data available, high-quality, and appropriately governed?
18. Are integration points with existing systems well understood?
19. Are automated pipelines in place to tracking experiments and results?

Governance, risk & impact

20. Have ethical, privacy, and security considerations been addressed? These are outlined in the AI Policy, especially [AI Policy Appendix C](#) depending on Use Case.
21. Is there a governance framework in place for oversight and accountability?
22. Has an impact assessment been conducted for the appropriate use case using the [AI Impact Assessment Tool](#)?

2. After PoC completion

Value assessment

1. Did the PoC meet the defined success criteria?
2. What tangible benefits were demonstrated (e.g. efficiency, cost savings, accuracy)?
3. Is there buy-in from prospective user groups?
4. Should progress be made to scale-up the AI?
5. Were the Benefits Management outcomes met?
6. Is there clear pathway to deploy to production and adequate funding available?

Stakeholder engagement

7. Is there leadership buy-in and trust in the solution?
8. Have all privacy considerations been factored in? Considering the Office of the Australian Information Commissioner (OAIC) as a foundational stakeholder.
9. Is there business sponsorship?
10. Have results been communicated in business terms, not just technical metrics?
11. Are there any workforce impacts that need to be supported and engaged through the implementation process?

Technical feasibility

12. Have any technical debt or integration challenges been identified?
13. Can the solution operate within existing ICT constraints?

3. Before scaling (pathway to production)

Enterprise readiness

1. Are there clear steps for further development, validation, integration, and testing (including regression testing and user testing)?
2. Has adequate funding been sought and approved to cover the pathway to Production? Is there a provision for top up funding if needed?
3. Is the solution designed for scalability, interoperability, and resilience?
4. Are modular, reusable components and a composable architecture in place?
5. Is the data breach response plan still applicable for the solution being assessed?
6. Has planning commenced to support any potential workforce impacts where relevant?
7. Has the Technical standard for government's use of artificial intelligence been considered?

Operational processes

8. Are automated pipelines in place for continuous deployment and monitoring?
9. Are natural change management processes defined to handle updates smoothly?

Governance & compliance

10. Are policies, risk controls, and ethical safeguards embedded for scaled use? Agencies need to evaluate against the AI Policy, especially AI Policy Appendix C dependent on use case.
11. Is there a clear ownership and accountability model across teams?
12. Has an impact assessment been conducted?

4. Pilot phase (controlled real-world testing and validation)

Real-world testing

1. Is the solution being tested in a controlled, representative environment?
2. Are diverse user groups and subject matter experts involved in testing and validating the AI solution?

Equity, inclusion and diversity

3. Are equity criteria embedded in pilot evaluation?
4. Are digital literacy supports and accessible feedback channels in place?

5. Has the Digital Experience Policy and Digital Inclusion Standard been reviewed?

Impact monitoring

6. Are feedback loops established to refine the AI solution before scaling?

Operational readiness

7. Are operational processes to manage AI trialled (e.g. support, monitoring, updates)?
8. Is the relevant workforce supported through potential changes (e.g. training and upskilling)?

5. After scaling (operationalisation)

Monitoring & adaptation

1. Are continuous monitoring and performance tracking mechanisms in place?
2. Are feedback loops for model retraining and improvement working effectively?

Change management & culture

3. Are change management processes working effectively to support adoption?
4. Is there a culture of learning, sharing, and reuse across teams?

Sustainability

5. Are long-term capability-building and talent development plans in place?
6. Is governance evolving to support responsible innovation at scale?
7. Are mechanisms in place to monitor compliance with the [AI in government policy](#)?

Appendix 2: AI evaluation

Scaling AI from PoC to full deployment involves continuous evaluation across all key aspects:

- technical feasibility and system performance – robust AI models, infrastructure and data management
- business alignment – AI objectives and business goals alignment, and change management
- monitoring and maintenance – AI system monitoring, retraining and feedback loops
- risk and compliance – addressing ethical issues, compliance and security
- finance and planning – ensuring adequate staffing and resources are available at each stage of the AI initiative

Evaluation practices should align with the [Commonwealth Evaluation Policy](#).

Below is a summary of the key evaluation options for scaling AI:

Category	Evaluation options
Technical feasibility and system performance	<ul style="list-style-type: none"> • Conduct thorough testing on edge cases, noisy data and real-world scenarios • Use various AI model metrics to test quality and reliability • A/B testing • Test how the AI will scale in terms of computation and data • Test performance for bottlenecks and system limitations • Test latency, throughput and response times • User and business feedback • Ensure smooth updates for deployment of updates • Ensure disaster recovery mechanisms are in place • See the AI Technical Standard for further details
Business alignment	<ul style="list-style-type: none"> • Ensure the AI solution continues to meet business objectives • Perform cost-benefit analysis and ROI projections in line with the Benefits Management Policy • Gather feedback from business • Assess change impact • A roadmap for training end-users
Monitoring and maintenance	<ul style="list-style-type: none"> • Tracking mechanism for performance of AI over time and changes • Ensure feedback loops allow to address AI drifts and degradation • Ensure AI explainability is embedded
Risk and compliance	<ul style="list-style-type: none"> • Assess the AI for ethical issues including bias • Ensure regular audits and review mechanisms are in place • Ensure the AI is compliant with legal frameworks
Finance and planning	<ul style="list-style-type: none"> • Cost analysis for staff and infrastructure, and each AI stage • Evaluation embedded in each AI stage

Appendix 3: Informing procurement

AI PoCs are essential for informing procurement decisions for AI initiatives because they transform a theoretical investment into a validated business case. They provide measurable evidence that helps mitigate the financial, technical and operational risks associated with adopting new, and usually complex, AI technologies.

PoCs inform procurement decision for AI initiatives across key areas:

- Risk mitigation – a PoC can identify risks at a small scale before significant funding, contracts and resources are committed
- Validation of business value – a PoC can prove whether the AI technology can solve the specific business problem
- Data readiness assessment – a PoC can reveal early on if necessary, data exists and good enough for the AI or determine gaps in the data holdings and governance
- Technical feasibility – a PoC can provide confidence on the feasibility of the technical solution
- Scope and requirements – PoCs can dictate technical requirements for the final solution; sets measurable threshold for vendor selection; identifies necessary pre-work to inform total budget
- Enables vendor and solution comparisons – PoCs allow for more direct and controlled competition (e.g. PoC-validated metrics, cost-benefit analysis)
- Provide negotiation leverage – PoCs inform risk-adjusted pricing and service level agreements

Below is a summary of the key procurement considerations for PoC implementation:

Category	Procurement consideration	Rationale/APS context	Practical actions
Problem definition & scope	Clearly define the business problem and desired learning outcomes of the PoC.	Commonwealth Procurement Rules (CPRs) require value for money and a clear purpose – AI PoCs must show measurable learning, not 'AI for AI's sake.'	Draft a concise statement of objectives (problem to solve, expected insights, duration, exit criteria). Tie to agency strategy or policy objectives.
Market approach	Use flexible procurement mechanisms suited to experimentation (e.g. limited tender, innovation panels).	PoCs require agility and rapid iteration; rigid procurement may delay innovation.	Use the DTA's Digital Marketplace or AI panels. Consider the ICT Procurement Framework Streamlined Approach for short-term engagements.

Vendor requalification & ethical assessment	Vet AI suppliers for ethical, security and data governance maturity.	Under the AI Technical Standard and AI Ethics Principles, vendors must demonstrate explainability, fairness and data governance capabilities.	Include mandatory criteria or evaluation questions covering responsible AI, security compliance (ISM) and alignment with AI Ethics Principles.
Intellectual property (IP) ownership	Define ownership of models, datasets and code generated during the PoC.	AI PoCs often produce reusable assets – the APS must retain rights to adapt or scale.	Specify IP terms upfront: government retains IP or receives a perpetual licence for all deliverables. Reference CPR Section 4 on long-term value.
Data access & sovereignty	Ensure compliance with PSPF, ISM and Privacy Act for data handling and storage.	Sensitive or citizen data may be used; data must stay within Australian jurisdictions and meet hosting certifications.	Require data to be processed and stored on Certified Hosting Framework (PROTECTED) providers or Australian sovereign cloud.
Security & privacy obligations	Integrate ISM, PSPF and Privacy Impact Assessments (PIA) into procurement conditions. ¹²	AI solutions must not compromise citizen trust or government integrity.	Include security and privacy compliance clauses; vendors must commence risk assessments and PIAs before PoC commencement.
Explainability & transparency requirements	Require transparency of AI models, training data and logic. Note: Vendors may push back claiming difficulty revealing corporate IP.	Aligns with the National AI Assurance Framework and AI Ethics Principles (Transparency, Accountability). ¹³	Mandate that vendors document model logic, limitations and decision paths; require plain-language explanations suitable for non-technical users.
Risk & assurance controls	Apply lightweight but effective risk and assurance frameworks for PoCs.	Even small PoCs should align to an AI impact assessment’s risk tier.	Require vendors to align with an AI impact assessment checklist, including bias testing and model performance validation.
Contract structure & duration	Keep contracts short, with defined exit or pivot points.	CPRs encourage proportionality – PoCs should not commit long-term funding before evidence of value.	Use staged contracts with clear deliverables, short timelines (3–6 months), and evaluation gates for scaling.

¹² See Footnote 6 and 7. [PIA assessment Tool - Privacy impact assessment tool | OAIC](#)

¹³ [National framework for the assurance of artificial intelligence in government and Australia’s AI Ethics Principles | Department of Industry Science and Resources](#)

Value for money & cost transparency	Require vendors to demonstrate cost breakdowns and reuse potential.	Prevents overspend and ensures learning outcomes justify investment.	Ask for transparent pricing by work package (data preparation, modelling, evaluation). Evaluate based on cost/benefit learning potential.
Open standards & interoperability	Prefer open architectures and reusable APIs.	Supports the Whole-of-Government ICT Strategy and avoids vendor lock-in.	Include mandatory criteria for open data formats, APIs and interoperability with the Australian Government Architecture (AGA).
Vendor lock-in mitigation	Ensure ability to transition, replace or re-host the solution after the PoC.	Lock-in risks long-term cost escalation.	Require export of model artefacts and documentation; ensure government retains access to underlying datasets and configurations.
Ethical AI & fairness clauses	Mandate adherence to the Australian Government AI Ethics Principles and relevant state policies.	APS agencies must demonstrate responsible AI use in line with ethics guidelines.	Include contract clauses requiring ethical assurance reports and bias testing results.
Performance measurement & evaluation	Define success metrics for assessing the PoC outcome and vendor performance.	Supports accountability and informs whether to scale to production.	Include performance indicators such as accuracy, fairness, explainability and user satisfaction. Link payments to evidence of learning outcomes.
Exit & transition planning	Require exit criteria and knowledge transfer deliverables.	Prevents stranded investment and supports scaling if successful.	Define handover deliverables (documentation, model artefacts, lessons learned report). Include government right to reuse learnings.

Appendix 4: Mapping of dimensions to the Technical standard for government’s use of artificial intelligence

To assist implementation teams with alignment questions and to provide greater technical depth for each implementation dimension, we have provided a mapping to the [Technical standard for government’s use of artificial intelligence](#) below:

Dimension	Relevant AI Technical Standard statements	Practical alignment notes
Business & strategic alignment	Statement 9: Conduct pre-work (design), Statement 12: Define success criteria (design), Whole-of-lifecycle Statement 1: Operational model	Problem framing, define measurable outcomes, operational model governance
Architecture & solution design	Statement 2: Reference architecture (whole-of-lifecycle), Statement 10: Human-centred design, Statement 11: Design safety, Statements 7 & 8: Version control & watermarking	Align to reference architecture, embed design safety, maintain traceability
Data & integration	Statements 13–19: Data supply chain, orchestration, quality, fusion/integration, establish context datasets; Whole-of-lifecycle Statement 6: Manage bias	Ensure data quality, governance, sovereignty, integration and bias controls
Technology & tools	Statements 20–25: Training & modelling, Whole-of-lifecycle Statements 2, 4, 7, 8 (Reference architecture, auditing, version control, watermarking)	Ensure tooling supports traceability, auditing, secure model training
People & skills	Statement 3: Build people capability (whole-of-lifecycle), Statement 10: Human-centred design, Statement 1: Operational model roles, Statement 4: Auditing capability	Develop AI skills, define roles, embed co-design, build auditing capacity
Governance & risk	Statement 1: Operational model, Statement 4: Auditing, Statement 5: Explainability, Statement 6: Manage bias, Statement 11: Design safety, Statement 9: Pre-work	Define governance structures, enable auditing, embed explainability and bias management
Experimentation & validation	Statement 11: Design safety, Statement 12: Success criteria, Statements 22–25: Evaluation & continuous improvement, Statements 4 & 5: Auditing & explainability	Embed success thresholds, user validation, auditing, continuous learning

<p>Delivery & operations</p>	<p>Operate phase: Integrate, deploy, monitor; Whole-of-lifecycle Statements 4, 5, 7 (Auditing, Explainability, Version control)</p>	<p>Integrate securely, deploy responsibly, monitor performance & maintain auditability</p>
<p>Scalability & transition to production</p>	<p>Operate phase: Deploy & monitor, Whole-of-lifecycle Statements 4, 7, 8 (Auditing, Versioning, Watermarking)</p>	<p>Plan scalability, embed operational monitoring, maintain version control</p>
<p>Sustainment & exit strategy</p>	<p>Retire phase: Decommission, Whole-of-lifecycle Statement 4: Audit, Statement 7: Version control</p>	<p>Plan for decommissioning, preserve audit trails, archive responsibly</p>