

# NAVIGATING THE AI TRUST GAP: WHY GOVERNMENT CAN'T AFFORD TO WAIT

AUTHOR: AARON GREENMAN

Governments around the world including Australia have embraced responsible AI policies grounded in ethics, transparency, and public trust. But while caution is warranted, inaction also carries risk. As expectations grow and AI capabilities accelerate, slow adoption can erode government relevance, capability, and credibility. The goal for government shouldn't be to eliminate AI risks entirely, but to manage them responsibly while keeping pace with the community's needs. This includes reducing over-reliance on external technology vendors and building internal capability to steer and govern AI effectively.

## The principles might be appropriate, but are they enough?

Over the past two years in particular, the Australian Government has made real progress in setting strong foundations for responsible AI. Agencies have published transparency statements, appointed accountable officers, and committed to ethical principles such as human oversight, explainability, and fairness. The Digital Transformation Agency's AI policy and classification framework has provided helpful structure. Most federal agencies currently use AI in low-risk domains: automating internal workflows, summarising documents, analysing policy data, or powering staff-facing virtual assistants.

This measured and principled approach is commendable and necessary. However, principles without increased momentum may soon become a liability. The public sector faces an emerging challenge: *how to operationalise AI principles at scale and speed, without compromising trust*. The truth is, while governments are moving carefully, the world around them isn't slowing down.

## The emerging risk: what happens if we move too slowly?

Playing devil's advocate, it's worth confronting an uncomfortable truth: *being too risk-averse with AI can create new forms of risk*. Here's how:

First, there's the issue of service experience. Citizens increasingly expect the same level of responsiveness and personalisation from government that they get from the private sector. When accessing health, welfare, or tax services, people want accurate, timely, digitally enabled experiences. If government systems feel slow, disconnected, or hard to navigate, trust erodes, not just in the service, but in the institution itself.



Second, there's the missed opportunity for productivity. AI has the potential to automate low-value tasks, freeing up public servants to focus on strategic, high-impact work. Without it, inefficiencies compound. Staff remain bogged down in administrative burden, innovation slows, and the public sector struggles to keep pace with demand.

Third, regulatory and policy expertise is at risk. As AI becomes embedded in every sector, from finance to defence, governments need operational literacy to regulate, audit, and respond effectively. Agencies that haven't strategically implemented AI internally, may find themselves ill-equipped to govern its use externally.

Fourth, there's the talent challenge. Public servants want to do meaningful, future-focused work. If the public sector is perceived as behind the curve, it risks losing, or failing to attract skilled technologists, analysts, and innovators to industry or overseas markets.

Finally, and perhaps most urgently, there is growing concern, particularly following recent developments overseas, about third-party entities wielding disproportionate influence over public sector AI systems. As governments outsource capability to commercial providers, they may inadvertently cede control (or visibility) over core decision-making processes. This includes dependency on proprietary models, limited transparency into algorithmic behaviour, and constrained ability to explain or audit outcomes.

This underscores the need for robust in-house capability. Agencies must be able to assess, adapt, and govern AI tools, not just procure them. Otherwise, governments risk becoming passive users of someone else's technology, rather than active stewards of their own digital future.


### **How other governments are moving responsibly but assertively**

While current Australian government AI use largely remains assistive, supporting, rather than replacing human decision-making, there is a noticeable lack of adoption of agentic AI systems capable of autonomous planning and action. Yet, opportunities abound in well-governed deployments and globally, leading governments demonstrate that responsible yet assertive AI adoption, including agentic AI, is achievable and beneficial.

For example, New Zealand has implemented agentic AI with its SmartStart platform, where AI proactively registers births, schedules healthcare appointments, and coordinates associated social services automatically.

In Singapore, agentic AI systems coordinate real-time traffic management, dynamically responding to changing conditions and improving congestion outcomes.

Estonia uses an agentic AI that proactively helps citizens navigate and complete complex administrative processes across multiple government services.



These international experiences provide valuable insights into safely and effectively deploying AI that doesn't just assist human tasks but autonomously performs complex public-service functions, offering practical lessons for Australia's own strategic AI adoption.

These examples show that strategic, experimental, and iterative AI adoption is possible. The key is to be strategic and pair innovation with accountability, starting with modest, measurable pilots, applying proportionate controls, and building internal literacy alongside technical tools.

### **Bridging the trust gap: practical moves for government**

Governments don't have to go all-in on AI overnight. But they must start moving faster and smarter. That means:

- Begin with internal, assistive use cases such as content summarisation, translation, policy drafting, or document classification, which provide immediate productivity gains and build internal confidence.
- Pilot AI tools in controlled environments including sandboxes, trials, and internal settings, with clearly defined metrics, transparent oversight, and thorough evaluation processes to identify opportunities and challenges early.
- Strategically plan for AI adoption by clearly identifying and prioritising use cases that offer tangible value. This includes defining specific roles, boundaries, and oversight mechanisms for agentic AI systems to avoid unchecked autonomy, particularly in sensitive areas like welfare or healthcare decision-making.
- Clearly delineate the scope and autonomy boundaries for agentic AI deployments, ensuring these systems augment rather than replace human judgment in critical processes. For example, agentic AI could proactively streamline administrative services or environmental monitoring while explicitly leaving final decisions and sensitive judgments in human hands.
- Develop multi-disciplinary teams comprising technologists, policy experts, legal advisors, ethicists, and domain specialists to provide comprehensive governance across AI deployments, ensuring ethical considerations and transparency remain central at every stage.
- Adopt tiered risk frameworks that tailor oversight and governance to the level of potential risk and impact, enabling responsible but agile implementation rather than uniform, overly cautious approaches.
- Enhance AI procurement literacy, empowering agencies to critically evaluate third-party solutions, insist on transparency, and embed public-interest protections directly into contracts.
- Invest proactively in workforce training, transitioning staff from foundational AI awareness to advanced model risk assessment capabilities, ensuring public servants are equipped with both policy literacy and technical fluency.
- Establish dedicated AI assurance functions to rigorously review, continuously test, and audit AI systems, particularly agentic tools, maintaining accountability and public trust throughout the AI lifecycle.



To support these shifts, certain prerequisites are essential:

- Clear governance arrangements: Agencies must define enterprise-wide structures and assign AI-specific accountabilities and responsibilities, including at the model and system levels.
- Data ethics integration: A consistent data ethics framework needs to be embedded across AI lifecycles, with reproducibility, auditability, and transparency integrated into model design.
- Policy and procedural alignment: AI development and use must align with existing organisational policies, supported by targeted procedures for AI-specific risks.
- Performance measurement: Agencies should establish clear mechanisms to evaluate AI's effectiveness, impact, and compliance across time
- Risk and assurance frameworks: AI-related risks, including misuse of data and unintended outcomes, must be assessed through enterprise risk management processes, with appropriate controls in place

These recommendations, drawn from the Australian National Audit Office (ANAO) review of ATO's AI practices<sup>[1]</sup>, reinforce that effective adoption requires more than tools, it demands culture, capability, and control.

Most importantly, agencies must ensure that any AI system they use, whether developed in-house or by a vendor, remains within their control, explainability, and accountability.

[1]. Governance of Artificial Intelligence at the Australian Taxation Office | Australian National Audit Office (ANAO).

**CONSULTANCY THAT  
VALUES BIG THINKERS,  
WHO SOLVE UNIQUE  
PROBLEMS.**