Article │ Received 23 February 2026; Revised 22 May 2026; Accepted 3 June 2026; Published 22 July 2026
https://doi.org/10.55092/let20260010

Auditable accountability without an AI act: Australia's public-sector AI assurance stack and the minimum reviewable trace

Grace Li

Faculty of Law, University of Technology Sydney, Sydney, Australia; Email: grace.li@uts.edu.au

Highlights:

- Procurement can operate as a quasi-regulatory lever for AI assurance in government.
- Minimum reviewable trace translates public-law reviewability into deployable evidence duties.
- The thin administrative record is the decisive governance failure mode for public-sector AI.
- Selective hardening should be triggered by use-case risk, opacity, and rights impact.
- Auditors, tribunals, ombudsmen, and contract managers need capability to assess AI artefacts.

Abstract: This article argues that Australia can pursue auditable accountability for public-sector artificial intelligence without enacting a single comprehensive AI statute, provided that existing legal duties, policy frameworks, standards, and procurement mechanisms are organised into an explicit assurance stack. An "AI Act" is used here to mean a unified statute that imposes system-wide ex ante obligations on providers and deployers. Australia currently relies on a more distributed model. The article's claim is that this model can still be made defensible, but only if abstract norms are translated into evidence disciplines that preserve contestability and reviewability. Using a doctrinal and functional method, the article shows how legality, procedural fairness, record-making, reason-giving, procurement discipline, privacy obligations, and audit practices can be aligned to produce a minimum reviewable trace for AI-influenced public decisions. The original contribution is twofold. First, the article conceptualises Australia's public-sector AI governance arrangements as an assurance stack whose layers only matter if they generate reviewable artefacts. Second, it proposes a bounded minimum reviewable trace that preserves the system configuration, material inputs and outputs, evaluation basis, reliance statement, and contestability pathway for a particular decision. The article also reframes selective hardening as a practical governance response for higher-risk uses, and presents an Assurance Requirement Level as a qualitative policy heuristic rather than a quantitative model. Concrete illustrations drawn from welfare eligibility and emergency-care triage demonstrate how the trace would work in practice. The article concludes that procurement is the principal operational lever for pushing evidence duties upstream, but that the stack will only succeed if audit offices, tribunals, ombudsmen, and agencies are equipped to interpret and test the artefacts they require.

Keywords: AI governance; AI assurance; administrative law; procurement; auditability; public sector; contestability; Australia

1. Introduction

Australia's public-sector AI landscape is often described as one of "many levers, but no AI Act". That description is broadly right, but it can mislead in two opposite directions. It may imply, first, that Australia lacks relevant law and governance infrastructure. It may also imply, secondly, that the existing mix of ethics principles, administrative law, procurement, privacy, audit, and soft-law guidance is already sufficient for contemporary AI deployment. Neither proposition is convincing. Australia does possess a dense set of legal and institutional levers relevant to public-sector AI, including ethics principles, government assurance policy, administrative law duties, privacy obligations, procurement rules, audit practice, and parliamentary and integrity oversight [1–3]. Yet these levers do not produce accountability merely by existing. They produce accountability only when they generate and preserve evidence that can be inspected, challenged, and acted upon.

This article argues that Australia can achieve auditable accountability without adopting a single comprehensive AI statute, but only if its current arrangements are treated as an assurance stack rather than as a loose bundle of principles. The expression "AI Act" is used here in a narrow sense to refer to a single comprehensive statutory regime imposing a general set of ex ante obligations on AI providers and deployers, exemplified most clearly by the EU AI Act [4]. Australia has not chosen that route. Instead, it governs through a distributed mix of public law, procurement, audit, privacy, sectoral obligations, and policy frameworks [1,2,5]. The central claim of this article is that such a distributed model can still be robust, but only if it is redesigned as an evidence-producing system.

The assurance stack proposed here is both descriptive and normative. Descriptively, it organises Australia's existing governance arrangements into a layered architecture. Normatively, it insists that each layer must produce artefacts that make AI-mediated official action reviewable. Those artefacts matter because public-law accountability in Australia remains deeply document-centred. Judicial review, merits review, ombudsman investigation, internal review, freedom of information processes, audit, and parliamentary scrutiny all depend heavily on reasons, records, and identifiable chains of responsibility [3,6–9]. If AI systems materially shape triage, ranking, drafting, recommendations, or final outcomes while leaving only a superficial administrative record, then legal accountability becomes formal rather than effective.

That thin-record problem is the article's decisive stress test. The issue is not confined to fully automated decision-making. A statutory decision-maker may remain formally in place, yet rely in substance on an algorithmic score, vendor-generated recommendation, or generative draft. In such a case, a file containing only the final conclusion and a brief statement of reasons may conceal how the outcome was shaped, what was checked, whether the tool was used inside its validated domain, or whether known limitations were ignored [10–14]. The practical consequence is an erosion of contestability. The doctrinal consequence is a weakened capacity to show legality, reasonableness, and procedural fairness.

The article's original contribution is twofold. First, it advances procurement as the principal operational lever for AI assurance in government. Procurement is not merely a purchasing mechanism. In public administration it functions as a quasi-regulatory device capable of imposing ex ante evidence requirements, audit rights, change-notice duties, and pass-through obligations across the supply chain [15–17]. Secondly, the article proposes a minimum reviewable trace. This is a bounded set of artefacts required for a specific AI-influenced decision to remain contestable and administratively

defensible. The trace is deliberately narrower than full technical transparency and narrower than any demand for source-code disclosure. It asks for what review bodies and oversight institutions actually need: enough information to reconstruct how a tool was configured, what it produced, how it was used, what checks occurred, and how an affected person can challenge the result [3,9,18,19].

The paper also introduces selective hardening as the mechanism for avoiding both under-regulation and compliance theatre. Not every AI use in government warrants identical controls. Internal productivity tools that do not affect individuals directly should not attract the same evidentiary burdens as systems influencing entitlements, sanctions, safety, or service access. The article therefore proposes an Assurance Requirement Level, framed not as a mathematical risk score but as a qualitative policy heuristic that combines impact, likelihood of harm, and opacity. Its purpose is to guide when guidance should become mandatory, when independent evaluation is needed, and when a minimum reviewable trace must be preserved [2,20,21].

Two concrete illustrations anchor the analysis. The first is a welfare eligibility or debt-assessment context, chosen because the Robodebt episode remains Australia's clearest example of the legal and human cost of a defective administrative record [22,23]. The second is an automated clinical triage scenario in a public emergency setting, used to show how the same trace logic operates where queue order and prioritisation materially affect outcomes even if no final legal entitlement is determined. Together they show that the assurance stack is not a slogan. It is a practical evidence architecture.

The remainder of the article proceeds in four parts. Section 2 explains the doctrinal and functional method and the criteria used to assess the adequacy of the proposed governance design. Section 3 develops the assurance stack itself, the procurement lever, the minimum reviewable trace, and the selective hardening logic. Section 4 addresses implementation risks, including bargaining asymmetry, confidentiality claims, and institutional capacity. Section 5 concludes that Australia's comparative advantage may lie not in immediate comprehensive codification, but in building auditable evidence pathways through existing legal and administrative structures.

2. Method and evaluative frame

This article adopts a combined doctrinal and functional method. The doctrinal dimension identifies the legal and administrative obligations already capable of supporting accountability for public-sector AI in Australia. These include legality, procedural fairness, rationality and reasonableness, record-making, reason-giving, procurement discipline, privacy and data governance, and reviewability through both internal and external forums [3,6,24–26]. The functional dimension asks how those duties can be made operational when AI systems influence public action. In practical terms, the article treats legal obligations not only as abstract norms but also as demands for artefacts, workflows, and evidence that can be tested by oversight institutions [21,27,28].

The analysis proceeds in three steps. First, it maps the current Australian governance environment for public-sector AI, drawing together ethics principles, finance and digital-government frameworks, procurement rules, audit practice, administrative law, privacy, and related accountability channels [1,2,15,29]. Secondly, it organises these materials into an assurance stack whose layers are assessed by reference to the artefacts they should produce, the accountability forums able to scrutinise them, and the characteristic failure modes that arise if those artefacts are absent. Thirdly, it stress-tests the stack against two recurrent problems: supply-chain opacity and the thin administrative record.

These stress tests are structured hypotheticals rather than empirical case studies. That choice is deliberate. The immediate purpose is not to estimate incident frequency, but to examine whether Australia's existing accountability mechanisms can still function once AI reshapes how facts, risks, recommendations, and explanations are produced. The first stress test asks whether an agency relying on an external model or vendor system can substantiate key claims about performance, safety, bias, and limitations without direct visibility into upstream artefacts. The second asks whether a person affected by a decision, together with a reviewer, auditor, or ombudsman, could reconstruct how the system influenced the outcome and whether reliance on it was lawful and fair [10,30,31].

The article assesses governance adequacy against four criteria. The first is contestability: whether an affected person can understand and challenge the material influence of AI on a decision. The second is auditability: whether an internal or external assessor can verify that the agency's own requirements were met. The third is allocability: whether responsibility can be traced across agencies, vendors, integrators, and decision-makers. The fourth is proportionality: whether the evidentiary burden is calibrated to risk and does not overwhelm low-risk uses [20,21,32]. These criteria reflect a broader normative commitment. AI governance in government should strengthen the defensibility of public administration, not weaken it.

3. Australia's public-sector AI assurance stack

The assurance stack is best understood as a translation layer between abstract duties and practical evidence. Each layer of the stack does different work, but the layers become meaningful only when they generate artefacts that can travel between decision-makers, contract managers, reviewers, auditors, and courts. The aim is not to impose one uniform governance mechanism. It is to ensure that ethics, policy, law, procurement, operational monitoring, and review institutions are linked through documents and verification pathways rather than left as disconnected commitments.

At a high level, the stack has six layers. The normative baseline comprises ethics principles, use-case registration, and purpose framing. Public-law duties require reasons-ready records, lawful reliance, and reviewable administrative files. Data governance adds provenance, retention, and lawful basis controls. Evaluation and risk management require validation, bias testing, monitoring, and change discipline. Supply-chain control is delivered mainly through procurement and contracting. Operational oversight involves telemetry, incident response, and escalation. Figure 1 demonstrates the architecture.

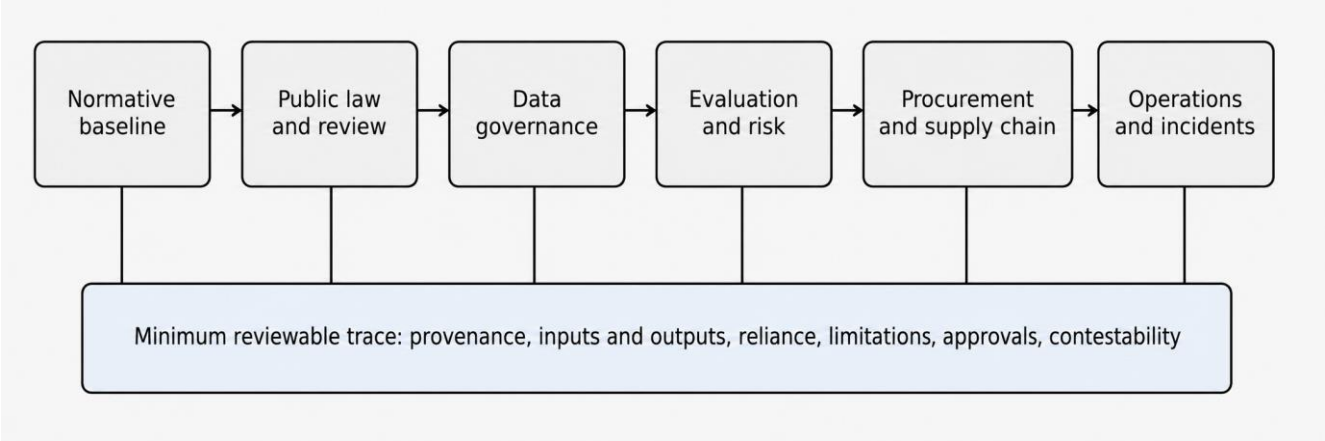

Figure 1. Public-sector AI assurance stack and the minimum reviewable trace.

3.1. The stack architecture and its artefacts

The normative baseline layer is the least coercive but still valuable. Government ethics principles, transparency statements, and policy frameworks shape the language of responsible use and can require agencies to declare the purpose, scope, and human-rights implications of an AI use case [5,29,33]. Their weakness is obvious: principles alone do not create consequences. Their value lies in establishing an administrative expectation that agencies should know what they are deploying, for what purpose, under whose authority, and with what stated limitations.

The public-law layer is more demanding. Administrative law remains the core site in which official action must be justified as lawful, rational, and procedurally fair. In practice, those obligations are mediated through the administrative record—decision files, reasons, source materials, review documents, and identifiable chains of responsibility [3,8,18,19]. The relevant failure mode is therefore the thin administrative record. Where AI shapes a decision but leaves no reviewable footprint, public-law accountability becomes fragile.

The data-governance layer concerns provenance, retention, lawful basis, access, and data-quality discipline. It matters because many contestability problems are actually data problems. A person may be unable to challenge an outcome because the underlying data source is unknown, inaccurate, stale, or difficult to retrieve. Privacy law and related guidance do not solve all AI harms, but they force attention to data lineage and handling practices [6,25,26].

The evaluation layer translates risk into testing, monitoring, and documented limitation. Here the most useful materials are not abstract ethical claims but evaluation summaries, robustness testing, error-rate analysis, red-teaming outputs, and change-control records [21,28,30]. These documents are particularly important where the deployed tool is vendor-supplied, foundation-model based, or updated over time.

The supply-chain control layer is where procurement becomes central. Public agencies often do not build the relevant systems themselves. They buy, integrate, configure, and rely upon them. That means the availability of upstream artefacts cannot be assumed. Procurement is therefore the practical mechanism through which evidence requirements can be transformed from desirable requests into deliverables, warranties, audit rights, and notification obligations [15–17].

The final operational layer concerns monitoring and incident response. Silent failure, delayed disclosure, and unmanaged model drift can defeat an otherwise well-designed assurance regime. For that reason, the assurance stack must include telemetry, incident thresholds, escalation pathways, and post-incident review. A system that was validated at procurement may nonetheless become unsafe or unreliable after updates, changing inputs, or a change in operating context [21,28]. Table 1 below summarises the dynamics of the layers.

Table 1. Public-sector AI assurance stack.

Stack layer	Primary instrument	Core artefact	Main accountability forum	Characteristic failure mode	Hardening option
Normative baseline	Ethics principles and AI policy	Use-case register; purpose statement; role and rights framing	Internal governance; ministerial and parliamentary scrutiny	Principles without consequence	Mandatory registration for higher-risk systems
Public law duties	Administrative law; review law; reasons practice	Reasons-ready file; reliance statement; review record	Internal review; tribunals; courts; ombudsman	Thin administrative record	Minimum reviewable trace for material AI influence
Data governance	Privacy and records frameworks	Data provenance map; lawful basis mapping; retention schedule	Privacy oversight; audit; FOI processes	Unverifiable data lineage	Standard lineage and retention controls
Evaluation and risk	Assurance frameworks; standards	Evaluation summary; error analysis; bias testing; monitoring plan	Audit; procurement assurance; executive oversight	Unvalidated performance or drift	Independent evaluation above threshold
Supply-chain control	Procurement and contracting	Assurance clauses; audit rights; change notices; pass-through obligations	Procurement challenge; contract remedies; audit	Vendor opacity and lock-in	Evidence package as contractual deliverable
Operations and incidents	Monitoring and incident response	Telemetry log; incident register; post-incident review	Auditor-General; regulators; executives	Silent failure and delayed disclosure	Mandatory material-incident notification

The stack is not simply a list. Its architecture matters because the layers interact and the dynamic amongst each layer. A transparency statement cannot compensate for the absence of upstream validation evidence. A strong procurement clause does not cure an empty decision file. A detailed case trace does not justify an unlawful or poorly evaluated system adoption. The system works only when adoption-stage assurance and case-level traceability are linked.

3.2. Procurement as the central operational lever

Any argument that procurement can function as governance must be anchored in black-letter law and practical procurement discipline, not presented as a purely conceptual claim. In the Australian context, the strongest doctrinal anchors are the accountable authority obligations in the PGPA Act, the Commonwealth Procurement Rules, and the Government Procurement (Judicial Review) Act 2018 [15,16,24].

Section 15 of the PGPA Act requires accountable authorities to govern in a way that promotes the proper use and management of public resources. “Proper use” is defined in terms of efficient, effective, economical, and ethical use. That duty does not mention AI specifically, but it strongly supports an argument that agencies should not procure or deploy opaque, inadequately evidenced systems in ways that expose them to foreseeable legal and operational failure [24]. AI assurance is therefore not an optional ethical add-on. In higher-risk settings it is part of sound resource governance.

The Commonwealth Procurement Rules provide a second anchor. Their enduring "value for money" orientation is often treated as primarily financial, but that would be too narrow. A system whose claims cannot be substantiated, whose changes cannot be tracked, or whose outputs cannot be defended may be poor value even if the purchase price is low. For public-sector AI, value for money should therefore include evidentiary sufficiency, auditability, portability, and the capacity to support lawful administration over time [15]. On that view, an apparently cheaper system that offers no meaningful access to evaluation evidence or change notices may be inferior to a more transparent alternative.

The Government Procurement (Judicial Review) Act 2018 adds a third element by underscoring that procurement obligations are not merely internal management preferences. They sit within a legal environment in which procurement rules can be contested and compliance challenged [16]. The significance for AI assurance is less that suppliers will litigate over every evidence clause, and more that procurement discipline already operates in a consequence-bearing legal setting. This strengthens the claim that procurement can function as a governance channel rather than as a discretionary aspiration.

Seen this way, procurement is the most practical mechanism for pushing evidence duties upstream. It can require an agency's suppliers, integrators, and subcontractors to deliver a defined assurance package—system description, versioning information, intended use boundaries, evaluation summaries, known limitations, security posture, monitoring plan, change-notice triggers, and incident-reporting obligations. It can also establish verification rights scaled to risk. For lower-risk systems, attestation and documentation review may be enough. For higher-risk systems, agencies should require access to evaluation processes, independent assurance outputs, and material change logs [21,28,30].

This is also the point at which consumer-law style substantiation norms become relevant. If vendors represent a system as "explainable", "fair", "safe", or "bias-reduced", agencies should not treat those claims as marketing gloss. They should require substantiation and incorporate the relevant representations into contractual assurance obligations. ACCC guidance on false or misleading claims reinforces the basic proposition that claims capable of influencing purchasing decisions should be supportable [17]. In AI procurement, that principle supports explicit evidence-delivery clauses.

Procurement is limited at the time, especially when agencies deal with large multinational providers. Bargaining asymmetry is real. Some providers may resist meaningful audit access or offer only highly abstracted documentation. That does not negate the governance role of procurement. It means that inability to obtain evidence should itself be treated as a governance signal. If a supplier cannot support lawful and reviewable reliance in a high-stakes public context, the appropriate response may be to narrow the use case, add compensating controls, or avoid deployment altogether.

3.3. The minimum reviewable trace

The article's central operational proposal is the minimum reviewable trace. The concept responds directly to the thin administrative record problem. It asks what information must exist in the file so that a decision materially influenced by AI remains contestable and administratively defensible. The trace does not aim to make every technical aspect of a model transparent. It is a bounded evidentiary discipline focused on what review institutions actually need.

The minimum reviewable trace has six core components.

First, provenance and configuration. The file should identify the system used, its version or endpoint, the relevant model or service, and the key configuration settings applicable at the time of use.

This is necessary because a later reviewer cannot assess a tool's role without knowing what system was actually in play.

Secondly, inputs and outputs. The decision file should preserve the material inputs supplied to the system for the case and the relevant output generated, whether a score, classification, ranking, recommendation, or draft. Privacy and security considerations may justify controlled rather than public access, but the information must exist.

Thirdly, reliance and verification. The file should record how the output was used by the human decision-maker, what checks were performed, whether it was treated as advisory or determinative, and whether any override occurred.

Fourthly, validated use and limitations. The file should note the evaluation basis relevant to the decision class and any limitations material to lawful reliance, translated into administrative rather than purely technical language.

Fifthly, human accountability. The record should identify the responsible officer, the approval pathway that authorised the tool's use for the decision class, and any required escalation.

Sixthly, contestability pathway. The file should preserve enough information to enable an affected person to request reasons, seek review, or raise a challenge, including contact points and explanation materials where appropriate [3,8,9].

Two distinctions are important. The first is between the upstream evidence package and the case-level trace. The upstream package justifies procurement, configuration, and continued operation of the system. It is aimed mainly at procurement, risk, audit, executive governance, and contract management. The case-level trace justifies reliance in a particular decision. Upstream assurance cannot compensate for an empty case file, and a carefully documented case trace cannot cure an unlawfully adopted or badly validated system. The second distinction is between explainability and defensibility. Public-law review generally does not require philosophical access to a model's inner workings. It requires enough information to assess whether official reliance on the system was lawful, reasonable, and fair. In this sense the trace is a doctrine-sensitive response to epistemic opacity. It does not demand full technical explainability. It demands administrative defensibility [13,14,18,19].

3.4. Concrete illustrations

The first illustration is a welfare eligibility or debt-assessment tool. Suppose an agency uses a machine-assisted matching and risk-flagging system to identify potential overpayments and prioritise matters for investigation. The final notice is still issued by a human officer, but queue position, recommended follow-up intensity, and draft reasoning are all shaped by the tool. Without a minimum reviewable trace, the file may contain only the final notice and a short statement of reasons. What is missing is precisely what later matters most: which model version was used, what data fields were relied on, how the risk flag was generated, whether the tool had been validated for that cohort, whether the officer checked anomalies, and whether the system's draft language overstated confidence.

This is where the Robodebt lesson becomes concrete. The Royal Commission exposed not only substantive unlawfulness, but the governance cost of weak record quality, delayed correction, and contested reasoning [22]. A minimum reviewable trace would not by itself have guaranteed legality, but it would have changed the administrative posture in at least four ways. First, it would have required preservation of the actual matching logic, data assumptions, and versioned configuration relevant to the

decision class. Secondly, it would have required a reliance statement showing what a human officer did with the output rather than allowing a general fiction of human endorsement. Thirdly, it would have linked the decision file to the upstream evaluation basis and known limitations. Fourthly, it would have made earlier scrutiny by internal reviewers, tribunals, auditors, and advisers more feasible. In other words, the trace is not presented here as a magic shield. It is presented as a structural antidote to the opacity that allowed legal error to persist.

A second illustration is an automated emergency-care triage support tool in a public hospital or ambulance setting. The tool does not determine a legal entitlement in the classic administrative-law sense. It does, however, influence queue position, escalation, and attention. If it ranks certain patients as lower priority because the training data under-represents their presentation patterns, the resulting harm may be immediate. Here the minimum reviewable trace would preserve the inputs used, the triage recommendation produced, the human clinician's reliance and override record, the validated scope and limitations of the system, and the escalation pathway used where the recommendation conflicted with clinical judgement. This example matters because it shows that the trace is not restricted to final rights determinations. It also applies to high-stakes prioritisation systems whose effects are operational yet serious.

These examples illustrate a broader point. The minimum reviewable trace is most useful where AI influence is easy to deny after the event. Its purpose is not to freeze innovation. Its purpose is to prevent agencies from occupying a governance grey zone in which AI meaningfully shapes action but responsibility remains diffuse and evidence thin.

3.5. Selective hardening and the Assurance Requirement Level

Selective hardening is the article's answer to the risk of overreach. Not every AI use requires identical discipline. Internal drafting assistance, low-impact search tools, and routine productivity applications do not warrant the same controls as systems affecting entitlement, sanction, liberty, safety, or access to essential services. The challenge is therefore to identify a practical trigger for moving from general guidance to mandatory evidence duties.

The proposed Assurance Requirement Level should be understood as a qualitative policy heuristic rather than a quantitative risk model. It combines three broad considerations: likely impact, likelihood of harm, and opacity. Opacity matters because assurance effort rises sharply where system behaviour, upstream changes, or vendor documentation gaps make claims difficult to verify. A simple matrix is more appropriate here than a mathematically presented formula. Agencies can assess each dimension as low, medium, or high, then map the combined result to a governance tier. The point is to guide assurance effort, not to create a false impression of precision [2,20,21].

For practical purposes, the following triggers should move a use case into a hardened tier.

First, rights-affecting decisions. If a system materially influences an outcome affecting entitlements, obligations, sanctions, or serious adverse consequences, a minimum reviewable trace should be mandatory.

Second, high-stakes triage or prioritisation. If a system changes queue order, investigation intensity, service access, or monitoring level, the agency should preserve the ranking objective, key evaluation basis, and an explanation pathway.

Third, reliance on third-party or foundation models. External dependency raises opacity and reduces direct control. That should trigger stronger upstream evidence requirements, change notices, and contingency planning.

Fourth, adaptive or frequently updated systems. Where model behaviour may shift over time, change logs, monitoring, and defined re-evaluation intervals become essential.

Fifth, generative drafting of notices, reasons, or recommendations. Where generative tools shape the language through which official action is justified, prompt records, configuration settings, and human verification steps should be preserved.

For use cases above the relevant threshold, the operational baseline should include at least the following—contractual audit rights; pass-through obligations for subcontractors and upstream providers; change-notice requirements; incident reporting thresholds; portability or exit protections; and a minimum reviewable trace in the case file where the tool materially influenced the outcome. These requirements are not presented as universal legislation. They are presented as the minimum hardening needed to make existing accountability channels work in practice.

3.6. Institutions that must read the artefacts

A recurring weakness in AI governance writing is the assumption that creating documents is enough. It is not. Evidence artefacts only matter if institutions know how to read them, test them, and attach consequences to them. That is why institutional capacity is not peripheral to the assurance stack. It is part of the stack.

Audit offices need enough capability to examine evaluation evidence, change control, and assurance claims rather than treating AI projects as ordinary digital procurements [7]. Tribunals and courts do not need deep technical mastery, but they do need a practical understanding of how AI-generated artefacts fit into reasons, records, relevance, and lawful reliance. Ombudsmen and integrity bodies may need guidance on when absence of a trace should itself be treated as a governance concern. Contract managers need to know that assurance clauses are not boilerplate. They are operational tools that must be used, enforced, and escalated.

This institutional point also reinforces why the minimum reviewable trace is bounded. Review institutions require artefacts they can reasonably work with. A governance model that depends on universal source-code inspection is unrealistic. A model that requires no more than a vague assurance statement is useless. The trace seeks a middle path grounded in the practical needs of public-law scrutiny.

4. Residual risks and limits

The assurance stack is not a complete answer to every AI governance problem. Its most obvious limitation is bargaining asymmetry. Public agencies may be able to obtain extensive documentation from smaller vendors while struggling to secure meaningful access or tailored terms from dominant cloud and foundation-model providers. In some cases, evidence gaps will remain even after careful procurement. Those gaps should not be normalised. They should be treated as signals that the use case may need to be narrowed, independently tested at the output layer, or abandoned.

A second limit is confidentiality and security. Some systems operate in environments where broad disclosure of prompts, features, or outputs would be unsafe or unlawful. The assurance stack accommodates

this by focusing on controlled preservation and controlled access rather than universal publication. What matters is that the evidence exists and can be accessed by legitimate accountability forums through appropriate channels.

A third limit is ritualistic compliance. An agency may produce registers, templates, and evaluation summaries without meaningful scrutiny. That is why selective hardening must be joined to verification routines. High-risk deployments should be sampled by internal audit, revisited after significant changes, and connected to actual consequence pathways in contract management and executive oversight. Documentation without scrutiny is paperwork, not accountability.

Finally, the article does not claim that Australia should never adopt more comprehensive legislation. The narrower claim is that, on current settings, a significant amount of improvement is possible through evidence-focused redesign of existing legal and administrative mechanisms. The assurance stack should therefore be understood as an immediate governance programme rather than as a final constitutional settlement.

5. Conclusion

Australia does not face a binary choice between a comprehensive AI statute and unstructured reliance on voluntary principles. A third pathway is available. Existing law, policy, procurement, audit, and review mechanisms can be organised into an assurance stack that delivers auditable accountability, provided those mechanisms are made to produce evidence rather than aspiration.

The core prescription is the minimum reviewable trace. It converts rule-of-law demands into a bounded evidentiary discipline for AI-influenced public decisions. It is narrower than full technical transparency, but materially stronger than generic promises of explainability. The second prescription is procurement-as-governance. In public administration, procurement is the practical means by which agencies can demand substantiation, preserve upstream visibility, require change notices, and allocate assurance responsibilities across supply chains.

The article has also argued for selective hardening. Higher-risk uses, especially those affecting rights, safety, entitlements, sanctions, and prioritisation, should attract stronger evidence duties and institutional scrutiny. Lower-risk uses should not be burdened with the same controls. The Assurance Requirement Level is therefore offered as a flexible policy heuristic, not as a pseudo-scientific scoring model.

The broader conclusion is that public-sector AI governance is, at bottom, governance of evidence. Whether official action can be justified, challenged, corrected, and defended depends on what can be shown. An assurance stack that treats evidence as its primary output offers a legally grounded and operationally realistic route to trustworthy AI adoption in Australia.

Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the author used generative AI tools (Copilot and ChatGPT) only to improve language and readability. The author takes full responsibility for the content of the manuscript.

Conflicts of interest

The author declares no conflicts of interest.

References

[1] Australian government department of finance. National framework for the assurance of artificial intelligence in government. 2024. Available: https://www.finance.gov.au/government/public-data/data-and-digital-ministers-meeting/national-framework-assurance-artificial-intelligence-government (accessed on 15 May 2025).

[2] Australian government digital transformation agency. policy for the responsible use of AI in government. 2025. Available: https://architecture.digital.gov.au/policy/responsible-use-of-ai-in-government?utm_source (accessed on 3 January 2026).

[3] Australian Government. Administrative decisions (Judicial Review) act 1977. 1977. Available: https://www.legislation.gov.au/C2004A01697/latest/versions (accessed on 20 December 2025).

[4] European Union. Regulation (EU) 2024/1689 (Artificial Intelligence Act). 2024. Available: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng (accessed on 15 May 2025).

[5] Australian Government. Standard for AI transparency statements. 2025. Available: https://www.digital.gov.au/ai/ai-in-government-policy/standard-ai-transparency-statements?utm_source (accessed on 20 December 2025).

[6] Australian Government. Freedom of information act 1982. 1982. Available: https://www.legislation.gov.au/C2004A02562/latest/downloads (accessed on 2 April 2025).

[7] Leocádio D, Malheiro L, Reis JCGD. Auditors in the digital age: a systematic literature review. 2024. Available: https://www.emerald.com/dts/article/4/1/5/1244156/Auditors-in-the-digital-age-a-systematic (accessed on 25 March 2025).

[8] Australian Government Attorney-General's Department. Administrative review council best practice guide: statements of reasons. 2025. Available: https://www.ag.gov.au/legal-system/publications/administrative-review-council-best-practice-guide-statements-reasons (accessed on 11 August 2025).

[9] Perry M. Statements of reasons: issues of legality and best practice. *Brief* 2021, 48(2):32–35.

[10] Citron DK. Technological due process. *Wash. UL Rev.* 2008, 85(6):1249–1313.

[11] Appel SM, Coglianese C. Algorithmic governance and administrative law. In *The Cambridge Handbook of the Law of Algorithms*. Cambridge: Cambridge University Press, 2020.

[12] Coglianese C, Lehr D. Transparency and algorithmic governance. *Admin. Law Rev.* 2019, 71(1):1–64.

[13] Wachter S, Mittelstadt B, Floridi L. Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. *Int. Data Priv. Law* 2017, 7(2):76–99.

[14] Veale M, Edwards L. Clarity, surprises and further questions in the article 22 right not to be subject to automated decision-making. *Comput. Law Secur. Rev.* 2018, 34(2):398–404.

[15] Australian Government Department of Finance. Commonwealth procurement rules. 2026. Available: https://www.finance.gov.au/government/procurement/commonwealth-procurement-rules?utm_source (accessed on 20 November 2025).

[16] Australian Government. Government procurement (Judicial Review) act 2018. 2018. Available: https://www.legislation.gov.au/C2018A00129/latest/text (accessed on 20 November 2025).

[17] Australian Competition and Consumer Commission. False or misleading claims. 2015. Available: https://www.accc.gov.au/consumers/advertising-and-promotions/false-or-misleading-claims (accessed on 2 February 2026).

[18] Minister for Immigration and Citizenship v Li (2013) 249 CLR 332. 2013. Available: https://www.hcourt.gov.au/cases-and-judgments/judgments/judgments-1998-current/minister-immigration-and-citizenship-v-li (accessed on 2 February 2026).

[19] Minister for Immigration and Border Protection v SZVFW (2018) 264 CLR 541. 2018. Available: https://www.hcourt.gov.au/cases-and-judgments/judgments/judgments-1998-current/minister-immigration-and-border-protection-v-szvfw (accessed on 2 February 2026).

[20] OECD. OECD framework for the classification of AI systems. 2022. Available: https://doi.org/10.1787/cb6d9eca-en (accessed on 22 February 2023).

[21] NIST. Artificial intelligence risk management framework (AI RMF 1.0). 2023. Available: https://doi.org/10.6028/NIST.AI.100-1 (accessed on 22 February 2025).

[22] Royal Commission. Report of the Royal Commission into the Robodebt Scheme. 2023. Available: https://robodebt.royalcommission.gov.au/publications/report (accessed on 22 February 2025).

[23] Parliament of Australia. Social security Act 1991. 1991. Available: https://www.aph.gov.au/parliamentary_business/committees/senate/community_affairs/completed_inquiries/2008-10/disability_trusts/ss_act_1991 (accessed on 18 June 2025).

[24] Australian Government. Public governance, performance and accountability act 2013. 2013. Available: https://www.legislation.gov.au/C2013A00123/latest/downloads (accessed on 20 December 2025).

[25] Office of the Australian Information Commissioner. Australian privacy principles guidelines. 2026. Available: https://www.oaic.gov.au/__data/assets/pdf_file/0032/263777/12-May-Consolidated-APP-Guidelines-Updated-Chapter-3.pdf (accessed on 18 February 2026).

[26] Australian Law Reform Commission. For your information: Australian privacy law and practice (ALRC report 108). 2008. Available: https://www.alrc.gov.au/publication/for-your-information-australian-privacy-law-and-practice-alrc-report-108/ (accessed on 18 Jan 2025).

[27] Kroll JA, Huey J, Barocas S, Felten EW, Reidenberg JR, *et al*. Accountable algorithms. *Univ. Pa. Law Rev.* 2017, 165(3):633–705.

[28] International Organization for Standardization. ISO/IEC 42001:2023 Information technology—artificial intelligence—management system. 2023. Available: https://www.iso.org/standard/42001 (accessed on 18 Jan 2025).

[29] Australian Government Department of Industry, Science, Energy and Resources. Australia's AI ethics principles. 2019. Available: https://www.industry.gov.au/publications/australias-ai-ethics-principles (accessed on 20 December 2025).

[30] NIST. Artificial intelligence risk management framework: generative artificial intelligence profile. 2024. Available: https://doi.org/10.6028/NIST.AI.600-1 (accessed on 23 Jan 2025).

[31] UK Government. The roadmap to an effective AI assurance ecosystem. 2021. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1039146/The_roadmap_to_an_effective_AI_assurance_ecosystem.pdf (accessed on 24 Jan 2025).

[32] OECD. AI principles. 2019. Available: https://www.oecd.org/en/topics/ai-principles.html (accessed on 24 Jan 2025).

[33] Australian Government Department of Finance. Cornerstones of Assurance. 2025. Available: https://www.finance.gov.au/government/public-data/data-and-digital-ministers-meeting/national-framework-assurance-artificial-intelligence-government/cornerstones-assurance (accessed on 18 December 2025).