



POINTS OF VIEW

You cannot govern AI by pretending people aren't using it

A Tasmanian parole case shows the cost of discovering too late that AI has entered important work without anyone being able to explain where, how or with what effect.

A document relied on by the Tasmanian Parole Board cited nonexistent cases. The board conceded that its secretary had used AI to prepare a submission supporting a condition that would have prevented a parolee from speaking publicly about her claim of wrongful conviction.

In September, the Supreme Court of Tasmania set aside the decisions imposing the condition after the State accepted that the parolee had been denied procedural fairness and that the condition was uncertain and legally unreasonable. It remains unclear which tool was used or whether its use was authorised. Later that month, the Tasmanian Government appointed KordaMentha to conduct an independent review of the governance, systems and controls supporting parole decision-making, including the circumstances surrounding the use of unverified AI-generated material.

That review does not mean earlier decisions were wrong. It means the review must now establish where AI entered the work and whether it mattered. Whatever time the original use saved, the organisation is paying for it through legal challenge, public scrutiny and retrospective review.

The model is not the service. The Tasmanian case exposes the next problem: an organisation may not know which model is being used, where it has entered the service or what work depends on it.

Shadow AI is demand made invisible

Tasmania is not an isolated example of AI entering important work without adequate controls.

In 2025, Deloitte refunded part of a \$440,000 fee after fabricated references and a misquoted Federal Court judgment were found in a report prepared for the Commonwealth Department of Employment and Workplace Relations. The corrected report disclosed that Deloitte had used an AI tool licensed by the department and hosted in its own cloud environment.

The Australian legal profession has seen similar failures. A lawyer acting in an immigration appeal filed submissions citing 17 nonexistent cases after using a public AI tool for research. In Victoria, a solicitor tendered authorities produced through AI-assisted legal research software without checking them and later said he had not fully understood how the product worked.

These cases differ. Deloitte used an approved tool in the client's environment. One lawyer used a public tool, while another used specialist legal software. In Tasmania, it is not known whether the use was authorised. What links them is visibility. AI entered important work without being visible enough for the right checks to operate before someone relied on the result. Approval did not necessarily mean the use was understood. Access to specialist software did not mean the user knew how it worked. A requirement for human review did not mean the sources were checked.

This is where shadow AI enters the discussion.



Shadow AI usually refers to AI being used for work without the organisation's approval, visibility or effective control. The same blind spot can also arise through suppliers, AI features embedded in existing software and approved products being used beyond their authorised purpose.

None of these cases has been established as shadow AI. The distinction matters because approval and visibility are not the same. An approved tool can still be used badly. Shadow AI goes further by removing the organisation's opportunity to apply controls at all.

The risk is simple: you cannot govern what you cannot see. If the organisation does not know a tool is being used, it cannot assess the information going in, check how the output is used, preserve the right records or know what needs to be revisited when something goes wrong.

The label is new. The behaviour is not. People find their own way to get work done when approved processes are too slow, available tools do not meet the need or nobody has decided what is allowed.

Most staff are not trying to breach policy. They are clearing a backlog, finding an old decision, summarising a long document, preparing a first draft or answering a question before the next meeting. AI removes friction from those tasks. Prohibition does not remove the demand.

People turn to outside services when the approved option is slower, harder to use or less capable than what is readily available. It may take too long to access, be unable to search the internal files they need, produce poorer meeting notes or offer little help with contracts, policies and technical standards. When the work is due, people choose the tool that helps them get it done.

Most of these uses will not produce a visible failure. That can make the arrangement seem safer than it is. The organisation receives the efficiency benefit while carrying hidden information, quality, legal and records risks. The absence of an incident is mistaken for evidence that controls are working.

The information risk exists even when the answer is accurate. Samsung employees reportedly pasted proprietary code and internal meeting material into a public generative AI service to make technical and administrative work easier. Nothing needed to be fabricated. Sensitive information had been sent outside the organisation's controlled environment.

No single study reliably measures unauthorised AI use across Australian workplaces. UNSW Canberra heard examples of public servants using free AI tools on personal devices, including where use was not authorised or was prohibited. Microsoft and LinkedIn found widespread international use of personally sourced AI at work. UpGuard found that many workers looked for another way in when a tool was blocked.

These sources use different methods, and some come from technology vendors. They do not establish a precise Australian figure. They point to a consistent behavioural problem: access, demand and formal approval are moving at different speeds.

Policy is necessary. Abstinence is not an operating model.

Frameworks need visibility

Governments and mature organisations are increasingly putting AI governance arrangements in place. Queensland requires agencies evaluating AI solutions to complete a Foundational Artificial Intelligence Risk Assessment or agreed equivalent. Commonwealth policy requires in-scope AI uses to have accountable owners, internal registers and impact assessments. Large private organisations have technology, privacy, security, quality and enterprise risk arrangements of their own.

These arrangements matter. They help organisations consider the purpose of a proposed use, the information involved, possible consequences, required controls and who remains accountable.

They also depend on the use becoming visible.

That visibility gap is not theoretical. In 2025, the Queensland Audit Office found that the department responsible for the state's AI governance framework had limited visibility of AI use and emerging ethical risks across government. It also found that the Department of Transport and Main Roads lacked department-wide oversight and had not completed dedicated ethical risk assessments for the two systems examined.

A framework cannot assess a personal account nobody has disclosed. It cannot assess an AI feature introduced through a routine software update if nobody notices it. It cannot govern a supplier's use if the contract does not require disclosure. It cannot detect that a product approved for drafting routine correspondence is now helping prepare decisions affecting people's rights.

The issue is not simply whether an AI framework exists. It is whether the organisation can see how AI is actually being used, keep that view current and make controls operate where the work is done.

This is where formal governance can create false confidence. A policy has been published. An approved product has been selected. Access to other products has been blocked. Staff have completed training. The organisation appears to be in control.

Meanwhile, the demand that led people to use outside tools remains.

If every uncertain use requires a committee paper, people will make the decision themselves. If early disclosure automatically triggers punishment, mistakes will remain hidden. If suppliers are not required to explain how they produced their work, the organisation may carry risks created outside its own systems.

Control begins with understanding the work, not selecting the tool.

Give people an approved option worth using

Many AI strategies fail here. The organisation buys a general productivity assistant, blocks everything else and assumes the need has been met. It clearly has not if the product cannot access the information people actually need.

A useful service starts with the work. It must be quick enough to use between meetings and grounded in current, authoritative information. A polished interface over poor information will produce faster confusion. Good information trapped behind a slow or clumsy service will send people elsewhere.

Many organisations, including law firms, now use permission-aware AI services that can search correspondence, decisions, reports and policies a person is authorised to access. If I can find what I need through a good interface, see the sources and stay inside the organisation's controls, I have less reason to take the question outside.

Existing permissions must be tested before anything is indexed. AI search can turn stale or excessive access into an active disclosure problem by making sensitive material much easier to find.

The service must also protect information, preserve appropriate records and make its sources and uncertainty clear. Even then, approval is not assurance. Deloitte's tool sat inside the client's environment, yet unreliable references still reached the final report.

The aim is not to make AI safe in the abstract. It is to provide a useful, controlled service without removing professional judgement, source checking or responsibility for the result.

Visibility has to extend beyond employees

An organisation can control its own systems and remain blind to how paid work is produced.

Suppliers should disclose material AI use, preserve the evidence behind their work, control subcontractors and embedded tools, report incidents and explain material changes to the products or models supporting the service. That does not mean a client should reproduce every piece of work it buys. Primary responsibility for quality remains with the supplier. Client assurance should test the method, evidence and controls in proportion to the consequence of getting it wrong.

The same principle applies when work moves between internal teams. Where AI contributes materially to advice, a client deliverable or a consequential decision, four things should remain visible:

1. What the tool was used for.
2. Which sources informed the output.
3. Who checked the facts, authorities and reasoning.
4. Who accepted responsibility for the final work.

A rewritten meeting invitation does not need an assurance file. Legal advice, a safety assessment or a decision restricting someone's freedom needs verified sources, a clear record of reasoning and someone able to explain the result.

A named reviewer is not assurance if that person lacks the expertise, source access, time or authority to challenge the work. Managers cannot reward faster output while pretending verification has no workload cost.

Accountability must also come with authority: the person responsible needs access to the evidence and the ability to restrict or stop the use when the controls are not operating.

People also need a safe way to report mistakes and uncertain use. Early disclosure should trigger help and containment. The organisation can then determine what work may be affected, whether information left its control, who needs to be told and what must be corrected.

If the automatic response is punishment, the use will not stop. It will become harder to see.

A practical first 90 days

If an organisation is serious about understanding its AI use and doing something practical about the risks, the first 90 days need to do more than produce another policy.

Ninety days is enough to establish ownership, expose the most consequential uses, contain immediate concerns, set initial boundaries and test one useful service against real work. It is not enough to find every use or resolve every technology, information, procurement, supplier and workforce issue.

What follows is not a universal implementation plan. It is a practical structure for getting started. Discovery, containment, decision-making and testing will overlap, and the scope will need to reflect the size, responsibilities and consequences carried by the organisation.

Start with the uses that could cause real harm. Use existing governance arrangements, but make accountability and authority clear. Review known AI use in decisions, advice and other high-consequence work immediately. Where there is a credible concern, preserve the evidence, contain the risk and establish what work may need to be checked.

Find the demand and the workarounds. Speak with staff about the tasks driving AI use and where approved processes or products are falling short. Do not rely on a survey alone. Examine the products already in use, AI

features arriving through software updates, reported incidents and supplier practices. Give people a safe way to disclose uncertain use and mistakes.

Set practical boundaries as the evidence emerges. Decide which uses can involve public information, which must stay inside an approved environment and which require stronger assessment or should remain prohibited. Make it clear who can approve a use, who accepts the remaining risk and who can restrict or stop it when controls fail.

Test one contained service against real work. Choose a problem where demand is strong, the information can be bounded and the consequences can be managed. Establish how the task performs now. Define the sources, permissions, checking steps, records requirements, supplier obligations and incident response before configuring the technology. Use the people who do the work and examples drawn from the work itself.

Measure the completed task, not just the speed of the first draft. Look at accuracy, correction rates, source quality, information handling, user behaviour, incidents and near misses. Find out why people still leave the approved environment and use that evidence to improve the service and the rules.

At the end of 90 days, the organisation should not claim that shadow AI has been solved. It should be able to explain where its most consequential uses sit, what has been contained, which boundaries apply, who can make decisions and whether the approved service is useful enough to change behaviour.

The Tasmanian parole failure is extreme, but the organisational conditions behind it are not unusual. Powerful tools are readily available. Staff and suppliers are under pressure to get work done. Formal approval moves more slowly than behaviour.

The answer is not to pretend AI can be kept outside the organisation. It is also not to buy an approved product and declare the problem solved.

Understand the demand. Make material use visible. Provide a protected and genuinely useful way to work. Keep responsibility for the result clear. Put the strongest checks where the consequences sit.

The technology will keep changing. The obligation does not. **The service still has to hold.**

Sources

- ABC News, “AI hallucinations in parole board submission scuttles media ban”, 15 September 2026.
- ABC News, “Revelations lead to review of AI use in parole decisions”, 19 September 2026.
- Tasmanian Government, “Independent review of the parole board”.
- Digital Transformation Agency, Policy for the responsible use of AI in government, version 2.0.
- Digital Transformation Agency, Artificial intelligence impact assessment tool, 1 December 2025.
- Queensland Government, Foundational Artificial Intelligence Risk Assessment framework.
- Queensland Audit Office, Managing the ethical risks of artificial intelligence.
- UNSW Canberra Public Service Research Group, The future of Generative AI in policy work, 2025.
- Microsoft and LinkedIn, 2024 Work Trend Index Annual Report.
- ACS Information Age, “Deloitte to refund government over AI errors”, 7 October 2025.
- Victorian Legal Services Board + Commissioner, Statement on the Mr Dayal matter, 2 September 2025.
- The Guardian Australia, report on an Australian lawyer’s use of generative AI in an immigration appeal, 1 February 2025.
- Bloomberg News, report on Samsung’s restrictions following the disclosure of sensitive information to a public generative AI service.
- UpGuard, The State of Shadow AI Report 2025 and The Shadow Supply Chain Report.