Shadow AI in the Enterprise

Governing the AI tools your organisation did not approve

White Paper | September 2026 | Version 2.0

Cetus AI Pty Ltd | Brookwater, Queensland

www.cetusai.com.au | hello@cetusai.com.au

CUSTOMER RESOURCE

Contents

Executive Summary

Shadow AI is the fastest-growing compliance gap in Australian enterprise. As artificial intelligence transforms the workplace, employees are rapidly adopting AI tools to improve productivity. However, when this adoption occurs without formal approval, governance, or oversight, it introduces profound risks. This unsanctioned use—known as shadow AI—exposes organisations to direct Privacy Act violations, breaches of the APS AI Policy, and the loss of intellectual property.

Cetus AI provides the core enterprise AI governance platform required to bring shadow AI into the light. By leveraging the **Teams**, **Control**, **Evidence Fabric**, and **Value Control** components, organisations can discover unsanctioned tools, ingest approved telemetry, and apply AI-specific policy decisions without discarding the enterprise stack they already trust. This paper explores the shadow AI problem, its regulatory implications in Australia, and a practical architectural framework for governing the AI tools your organisation did not approve.

The Shadow AI Problem

The rapid proliferation of artificial intelligence has fundamentally altered the enterprise technology landscape. Shadow AI occurs when employees use generative AI tools, models, services, or agents without the knowledge, sanction, or oversight of the technology and risk teams. Unlike traditional shadow IT, which often involved unauthorised software-as-a-service applications for relatively benign tasks, shadow AI introduces a unique set of risks due to the nature of how AI models process, retain, and utilise information.

This phenomenon is driven by a powerful and understandable pull factor: employees want to work faster, more creatively, and more efficiently. AI tools offer unprecedented capabilities for drafting documents, analysing complex datasets, summarising meetings, and generating code. When the enterprise fails to provide approved, secure AI tools, or when the approval process introduces excessive friction and delay, staff inevitably find their own solutions [2]. They bypass official channels, prioritising immediate productivity gains over abstract security or compliance concerns.

The consequences of unmanaged AI adoption are significant and multifaceted. The most immediate risk is data exposure. Employees may unintentionally upload sensitive client data, financial records, strategic business plans, or intellectual property into public large language models (LLMs) [3]. Once data is fed into a public model, the organisation loses control over it. There is often no mechanism to retrieve or delete the information, and depending on the vendor's terms of service, the data may be used to train future iterations of the model [2]. This can result in confidential information being exposed to the public domain or inadvertently surfaced in responses provided to other users of the AI service.

Beyond data leakage, shadow AI creates a critical governance deficit. The organisation lacks visibility into who is using which models, for what purpose, and with what data. This opacity prevents the technology team from assessing the security posture of the tools being used, the risk team from evaluating the compliance implications, and the finance team from tracking the hidden costs associated with individual or departmental subscriptions. Furthermore, unmanaged AI tools can generate inaccurate, biased, or hallucinated outputs, which employees may rely upon for critical business decisions without understanding the limitations of the technology [2]. The resulting lack of accountability and auditability represents a profound vulnerability for any mature enterprise.

Regulatory Exposure in Australia

The use of shadow AI is not merely an internal policy violation; it directly and materially impacts an organisation's ability to comply with stringent Australian regulatory frameworks. As government and industry regulators increase their focus on data protection and responsible AI usage, the inability to govern AI tools presents a severe compliance risk.

The Privacy Act 1988

The Privacy Act 1988 places clear, legally binding obligations on Australian businesses and government agencies that collect, hold, or process personal information. Shadow AI has emerged as the largest and most complex source of Privacy Act exposure for organisations that have not established clear, operational AI policies and controls [4].

When staff input client data, employee records, or other personal information into an unapproved AI tool, they are effectively transmitting that information to a third-party vendor's infrastructure outside the organisation's approved security boundary. This action triggers immediate compliance concerns across several Australian Privacy Principles (APPs):

- **APP 1 (Open and transparent management of personal information):** Organisations are required to maintain a clearly expressed and up-to-date privacy policy that accurately describes how personal information is managed. If unapproved AI tools are processing personal information—even as a background service or temporary processing step—the organisation's published privacy policy likely no longer reflects its actual data handling practices. This misalignment constitutes a direct compliance failure under APP 1 [4].
- **APP 6 (Use or disclosure of personal information):** Personal information collected for a primary purpose generally cannot be used for a different, secondary purpose without the individual's consent. Using client data as input for a public generative AI model, or allowing that data to be used by the vendor for model training, represents a secondary use that falls far outside the reasonable expectations of the individual at the time of collection [4].
- **APP 8 (Cross-border disclosure of personal information):** The majority of public AI platforms, including prominent LLMs and generative tools, process and store data on infrastructure located outside Australia, typically in the United States or Europe. APP 8 requires organisations to take reasonable steps to ensure that overseas recipients handle personal information in accordance with Australian privacy laws. Using shadow AI tools bypasses these required assessments and contractual protections, exposing the organisation to significant cross-border disclosure risks [4].
- **APP 11 (Security of personal information):** APP 11 mandates that organisations take reasonable steps to protect personal information from misuse, interference, loss, unauthorised access, modification, or disclosure. Unsanctioned AI tools inherently operate outside the organisation's established security perimeter, access controls, and monitoring systems. The inability to secure data once it enters a shadow AI platform fundamentally undermines the protections required by APP 11 [4].

APS AI Policy

For government entities, departments, and their technology suppliers, the regulatory landscape is even more explicit and demanding. The Digital Transformation Agency's policy for the responsible use of AI in government mandates strict, comprehensive governance requirements to ensure the safe and ethical adoption of AI across the public sector.

The policy requires agencies to establish a strategic approach to AI adoption, operationalise responsible use principles, maintain detailed internal use case registers, and undertake rigorous, risk-based AI use case impact assessments. Furthermore, agencies must designate clear accountability for every AI use case and publish transparency statements regarding their AI deployments.

Shadow AI fundamentally and systematically undermines every one of these mandatory requirements. An agency cannot possibly assess the impact, risk, or ethical implications of an AI use case if the technology team is entirely unaware that the tool is being used. It cannot record the tool in an internal register, nor can it disclose its use in a public transparency statement. Most critically, the policy's requirement for designated, accountable officials for AI use cases is rendered impossible when the tools are adopted informally by individual employees or isolated teams. Shadow AI in a government context is not just a security risk; it is a direct failure of mandated public sector accountability.

A Practical Governance Framework

Banning AI is neither practical nor effective; employees will simply find harder-to-detect workarounds. The solution is to transition from policy intent to operational control using a structured framework.

Governance Layer	Customer Responsibility	Cetus AI Platform Support
Discovery	Identify the extent of shadow AI usage through network analysis, DLP tools, and employee surveys.	**Teams** component supports enterprise visibility and governance-intake, allowing organisations to register discovered tools.
Policy Definition	Establish clear, unambiguous expectations about which AI tools are allowed and for what tasks [2].	**Control** component evaluates metadata against the defined policy, returning approve, deny, or conditional decisions.
Runtime Enforcement	Implement controls to prevent sensitive data from reaching unapproved endpoints.	**Control** component provides policy, identity, runtime guardrails, and exception handling.
Evidence Correlation	Maintain reviewable records of AI usage, policy enforcement, and data flows.	**Evidence Fabric** connects approved records across the customer environment to create a coherent governance record.
Economic Attribution	Track the cost and sustainability impact of AI usage across the enterprise.	**Value Control** module distinguishes forecast, observed, attributed, finance-validated, and risk-adjusted value.

Application to the Cetus AI Platform

Cetus AI provides the core enterprise AI governance platform designed specifically to address the complex challenge of shadow AI. The platform is built and available for controlled enterprise deployment using Release 2026.09.4 capabilities. Rather than forcing a mandatory middleware proposition that disrupts existing workflows, Cetus AI is architected to integrate seamlessly with the systems an organisation already trusts, adding the missing AI-specific policy, control, and evidence layers.

The journey from shadow AI to governed AI begins with visibility. The platform's **Teams** component establishes the necessary enterprise visibility and governance-intake process. It allows organisations to systematically discover, catalog, and assess the risk of AI tools operating within their environment, transforming unknown shadow applications into documented, reviewable assets.

Once visibility is established, governance requires operational enforcement. When a user, application, or autonomous agent attempts to interact with an AI service, the **Control** component provides the critical policy, identity, runtime guardrail, exception, and human-oversight capabilities. Through the Policy Decision Point, the platform evaluates the request against defined enterprise rules. It can enforce dynamic runtime guardrails—such as automatic PII redaction to protect sensitive data, model substitution to enforce the use of approved endpoints, or mandatory human-review gates for high-risk operations—ensuring that even sanctioned tools operate within strict safety boundaries.

Crucially, governance is incomplete without proof. The **Evidence Fabric** is the shared evidence capability that connects approved records across the customer environment. It correlates selected telemetry and events from across the enterprise stack—including existing identity platforms, API gateways, and cloud monitors—into a coherent, tamper-evident audit chain. This provides the review-ready evidence required by regulators, auditors, and boards, proving exactly who had authority, which policy applied, what data was involved, and how the AI model responded.

Furthermore, the integrated **Value Control** module ensures that AI consumption is economically attributed and governed. It distinguishes between forecast, observed, attributed, finance-validated, and risk-adjusted value, allowing organisations to track the true cost and return on investment of their AI initiatives, moving beyond estimations to accountable financial records.

Limitations and Customer Responsibilities

Cetus AI provides robust technical infrastructure, but it does not replace the customer's governance responsibilities. The platform's capabilities can support a customer process within configured coverage; they do not make the customer compliant, prove causation, replace an auditor, or guarantee completeness.

Customer-specific integrations and deployment patterns require validation against the customer's architecture, sources, identity, security, performance, failure behaviour, and operating model. Evidence Fabric correlates approved source records and makes gaps visible, but correlation does not establish causation. Customers remain responsible for data classification, defining acceptable use policies, providing employee training, and making final risk acceptance decisions.

Recommended Next Steps

To address the shadow AI risk, organisations should take immediate action:

1. **Assess Current Exposure:** Conduct a baseline assessment to identify unsanctioned AI tools currently operating within the environment.
2. **Define Acceptable Use:** Update privacy policies and acceptable use guidelines to explicitly address generative AI and third-party models.
3. **Establish Visibility:** Deploy the Cetus AI **Teams** component to formalise governance-intake and maintain an automated-decision register.
4. **Implement Guardrails:** Configure the **Control** component to enforce PII redaction and policy decisions on outbound AI requests.
5. **Review Architecture:** Book an Enterprise Architecture Review with Cetus AI to validate deployment patterns and integration with existing identity and security platforms.

For more information or to discuss your AI governance requirements, contact **hello@cetusai.com.au**.

References

[1] Cloud Security Alliance, "Shadow AI Apps: The Enterprise Attack Surface That Outpaces Monitoring," May 30, 2026. Available: https://labs.cloudsecurityalliance.org/research/csa-research-note-shadow-ai-apps-enterprise-20260530-csa-sty/ [2] D. Vedeneev and H. Ma, "Shadow AI: Why your biggest AI threat may come from within," CyberCX, Mar. 27, 2026. Available: https://cybercx.com.au/blog/shadow-ai-why-your-biggest-ai-threat-may-come-from-within/ [3] Leapfrog Market, "Hidden Shadow AI Business Risk: What Australian Businesses Need to Know in 2026," Aug. 22, 2026. Available: https://leapfrogmarket.com.au/shadow-ai-business-risk/ [4] J. Richards, "AI Implementation and the Privacy Act in Australia: A Compliance Guide," XCD IT, May 24, 2026. Available: https://xcdit.com.au/ai-implementation-privacy-act-australia/