



LETTERS

APRA Letter to Industry on Artificial Intelligence (AI)

Summary of common weaknesses and expectations for regulated entities - April 2026

To: All APRA-regulated entities

Artificial Intelligence (AI) is being rapidly adopted across APRA-regulated industries as entities seek to realise benefits to their businesses and customers. AI presents great opportunity for productivity and efficiency, and failing to embrace AI may put businesses at a strategic disadvantage. AI also has the potential to create new risks and escalate existing challenges. To understand and assess the current state of AI adoption and associated prudential risks, APRA conducted a targeted engagement on a group of selected large banks, insurers and superannuation trustees in late 2025. The purpose of this letter is to outline these observations and APRA's expectations in managing AI related risk. Lessons drawn from APRA's observations of these larger entities, will assist other entities who may be earlier in their AI adoption journey.

APRA found that, while AI is being actively adopted by all the entities we engaged with, there are differing levels of maturity across functions such as governance, risk management and operational resilience. In addition, assurance practices are not keeping pace with the scale, speed and complexity of AI. With respect to Boards, APRA observed strong interest and pursuit for AI's potential benefits and strategic imperatives, particularly in relation to productivity, efficiency and customer experience. However, APRA observed many Boards are still developing the technical literacy required to provide effective challenge on AI related risks and oversight. APRA also noted an overreliance on vendor presentations and summaries without sufficient examination of key AI risks such as unpredictable model behaviour and the impact on critical operations.

APRA expects Boards, at a minimum, to:

- maintain sufficient understanding and literacy with respect to AI in order to set strategic direction and provide effective challenge and oversight
- oversee an AI strategy which is consistent with the entity's risk appetite and tolerance settings, supported by effective monitoring and reporting (including for third party dependencies), with clearly defined triggers aligned to resilience objectives to enable timely action when not operating as expected.

APRA's observations and expectations for accountable executives are provided in the attachment to this letter. These are provided to assist CRO, CTO and CISOs in understanding APRA's expectations and to support prompt action to address gaps given the fast-moving environment.

APRA is also engaging across the sector on the potential for increased cyber threats from high capability AI frontier models such as Anthropic Mythos. APRA has heard clear recognition from regulated entities of the need for a step change in cyber practices and a continuing uplift in capabilities to protect IT assets in an evolving threat environment. This uplift could also include the use of AI in identifying and resolving vulnerabilities. APRA has been engaged with the Council of Financial Regulators (CFR) and government agencies on AI use and risks, and entities should note current ASD advice on frontier AI models.¹

APRA's principle-based prudential framework is technology and vendor agnostic. APRA requires regulated entities to ensure appropriate risk management of AI, including setting risk appetite, managing AI related exposures, and ensuring appropriate oversight and accountability. APRA emphasises the need for entities to prioritise and strengthen their security defences, including the timely implementation of patching, closure of vulnerabilities and focused attention to cyber hygiene. APRA will apply its supervisory focus to entities' AI adoption and manage the resulting risks. Where entities fail to adequately identify, manage or control AI risks in a manner proportionate to their size, scale and complexity, we will take stronger supervisory action and, where appropriate, pursue enforcement.

APRA is currently finalising its forward plan with regards to supervision of AI risks, taking a proportional approach to entity prudential reviews, thematic activities and AI supplier engagement. APRA will continue and monitor the use of AI to assess potential prudential risks and consider whether further APRA policy action may be needed.

While this letter provides guidance based on current observations, APRA strongly encourages entities to engage early with APRA's Non-Financial Risk Team via your supervisors on any unexpected or heightened AI-related risk concerns, including where existing risk management approaches may be challenged.

Yours sincerely,

Therese McCarthy Hockey

Member

AI Supervisory Engagement Debrief: Observations for Executive Management

The observations below reflect the issues identified by APRA from a deep-dive exercise on a sample of the largest banks, insurers and superannuation trustees that are published for the benefit of all regulated entities. The observations highlight areas where governance, risk management, assurance and operational practices are failing to keep pace with the scale, speed, and complexity of AI adoption.

Observations

AI threats are increasing, but information security practices are struggling to keep pace

APRA observes that AI adoption is materially changing the cyber threat landscape for regulated entities. The use of AI increases the pathways that cyber attackers can use and lead to more frequent cyber attacks. Common attack pathways observed include prompt injection, data leakage, insecure integrations, exploit injection and the manipulation or misuse of autonomous AI agents. AI can shorten the attack cycle and increase speed, coordination and impact. At the same time, entities are using AI to improve threat hunting and vulnerability identification, with the challenge being remediating at the speed with which vulnerabilities are identified.

Concerns were noted across several areas. Identity and access management capabilities have not yet adjusted to nonhuman actors such as AI agents. The volume and speed of AI assisted software development is placing strain on the effectiveness of change and release management controls. APRA observed gaps in the scope and coverage of security testing programmes for both AI implementation and responding to the AI augmented threat environment. The implementation timelines for information security remediation activities, such as patching and configuration management, are not consistently aligned to the accelerated threat environment. These issues are compounded by the variability across organisations technology deployments and the increasing volume of discovered vulnerabilities and threats requiring priority remediation, without a significant backlog.

The use of enterprise AI tools by staff outside approved control frameworks is also a concern. Whilst strategies to encourage staff experimentation and progress cultural change are commended, the calibration of these activities to risk appetite appears weak. In many cases, preventative controls were lacking, with entities relying primarily on policy direction or detective, after-the-fact measures, rather than enforceable technical restrictions or robust preventative controls.

APRA expects entities to actively manage information security vulnerabilities and threats. This would include:

- assessing the implications of AI reliance for operational resilience and business continuity. Where AI supports critical operations, credible fallback processes are required;
- security controls and capabilities that effectively address AI-specific threats and attack paths. This would include strong privileged access management, timely patching, hardened configurations, automated vulnerability discovery, penetration testing, and controls over agentic and autonomous workflows;
- robust security testing across AI-generated code, software components and libraries; and
- ongoing consideration of third-party and concentration implications in relation to common platforms, services, and providers.

AI adoption is moving fast, but governance maturity is lagging

APRA observed AI adoption is accelerating across all regulated industries. Entities are moving beyond experimentation from internal productivity use cases to customer facing applications. Many entities are

already trialling or introducing AI capabilities in areas such as software engineering, claims triage, loan application processing, fraud and scam disruption, customer interaction and insight generation.

However, governance has not matured at the same pace. While most entities recognise that existing prudential standards apply to AI risk, few have operationalised governance in practice. APRA observed a tendency to treat AI risk as 'just another technology'. This misses key differences such as the distinct characteristics of predictive systems, adaptive behaviour in models, ethical considerations such as inherent bias, and privacy and data risks. This has resulted in gaps in the management of AI across its lifecycle. Gaps include weak controls over post deployment monitoring, weak model behaviour monitoring, change management, and decommissioning of AI capabilities.

APRA expects entities to establish consistent governance arrangements that include, at a minimum:

- frameworks (policy, standard, guidance) and reporting lines to promote safe, responsible and sustainable adoption of AI;
- ownership and accountability across the AI lifecycle, from design and development through to deployment, monitoring and decommissioning;
- an inventory of AI tooling and AI use cases;
- human involvement for high-risk decisions and accountability; and
- training and education of staff on AI use, misuse, limitations and secure practices.

Supplier risk management is in place, but supplier concentration and opacity present challenges

APRA observed some entities heavily dependent on a single provider for multiple AI use cases. Few entities had demonstrated robust contingency planning or tested exit and substitution strategies for critical AI providers. Contractual arrangements often lagged practice, with limited evidence of specific provisions addressing audit rights, model updates and deviations, incident notification or changes to data handling.

AI capabilities are increasingly embedded within software, platforms or developer tools. This can mean upstream dependencies such as foundation models, training data sources and fourth party service providers are opaque which limits entities' ability to independently assess model performance, bias, resilience and security. Taken together these variables challenge an entity's ability to completely and effectively assess and manage risk.

APRA expects entities to manage supplier risks, this would include, at a minimum:

- mapping and maintain visibility over the full AI supply chain, including material, third-party and fourth-party dependencies;
- contractual and governance arrangements which provide sufficient transparency, auditability and assurance over AI services.

- entities should have the ability to understand model behaviour, material changes, performance issues and outcomes, and risk management practices across the service lifecycle; and
- active management of concentration risk. This would include plausible and systemic failure scenarios, the credibility and feasibility of substitution, portability or exit arrangements for critical AI providers.

Traditional change management and assurance is in place but is not sufficient for dynamic AI solutions

AI risks can cut across multiple domains at regulated entities. This includes operational risk, cyber and information security, data governance, model risk, change control and release management, legal and regulatory compliance, privacy and conduct risk, procurement and third party dependency. Existing change and assurance management approaches are often fragmented and may not effectively provide sufficient assurance.

APRA also observed reliance on point in time and sample based assurance methods, despite these methods being ill suited to probabilistic models that learn, adapt and degrade over time. Few entities had continuous validation or monitoring in place to detect issues such as model drift, bias, failure modes, or control breakdowns in a timely manner.

APRA observed internal audit and risk functions are challenged. Many lack the specialist skills and tools required to engage in AI assessment or audit. This is particularly true where agentic behaviour, automated decision making or AI assisted code generation were involved. As a result, assurance activities often lagged AI deployment.

APRA expects entities will adopt effective assurance mechanisms and approaches. This would include at a minimum:

- employing globally recognised control frameworks including control libraries, and change control for AI implementations;
- applying integrated assurance across cyber security, data governance, model performance risk, operational resilience, privacy, and conduct risks;
- second line risk management and internal audit functions possess technical capability and tooling to independently assess AI systems including probabilistic models and agentic workflows; and
- conducting comprehensive risk and information security assessments prior to deployment and throughout the lifecycle. Monitoring should be continuous and proportionate to the criticality of the use case, including consideration of model purpose, limitations, explainability and potential customer impacts.

Footnotes

¹<https://www.cyber.gov.au/about-us/view-all-content/news/frontier-models-and-their-impact-on-cyber-security>

2026