AI Governance — Australia

Regulators have shifted from principle-based guidance to active supervision. AI risk is now a board-level accountability.

AS OF	SOURCES	TOPIC
12 May 2026	6 syntheses	ai-governance-au

The thesis

AI-related risk is a board accountability. Proportionate, evidenced action is expected now. Enforcement is on the table.

THE REGULATOR STACK

Four signals, one direction of travel.

SIGNAL	DATE	WHAT IT SAYS
APRA letter to industry on AI	30 Apr 2026	AI governance lags adoption; treat as a specific risk domain inside the ERMF
ASIC open letter	8 May 2026	Cyber resilience first principles, *Mythos*-class capability named explicitly
ASIC REP 798 "Beware the Gap"	earlier	Confirmed governance gap across financial services
OAIC	ongoing	Existing privacy laws apply to AI — **no grace period**

Consolidated message: AI risk is board-level. Proportionate action expected now.

50%

of CyberCX's 2025 AI penetration tests found a severe vulnerability — **double** the 26% web-app rate.

The strongest single AU/NZ data point behind APRA's "controls lag adoption" claim. [[2026-05-12-cybercx-2026-hack-report]]

WHERE THE EMPIRICAL PICTURE CONFIRMS THE REGULATOR PICTURE

Operational evidence aligns with the "act now" framing.

- FIRST **Financial Services overtook Healthcare** as the most-impacted sector (18% vs 12%) for the first time.
- UNIVERSAL **MFA bypass via session hijacking** in *every* CyberCX BEC case where TOTP/push MFA was enforced — policy-level MFA produces zero defensive value.
- FIRST **Offensive GenAI observed in DFIR** — 2025 case of a threat actor using GenAI for bespoke scripts and payloads.
- 2× **AI pen-tests find severe vulnerabilities at twice the web-app rate** — AI is reaching production at lower security maturity.

THE BOARD-ACCOUNTABILITY FRAME

What APRA expects of directors and accountable persons.

- **AI literacy** sufficient for strategic direction and effective challenge of management.
- AI oversight as **part of discharging existing director duties** — duty of care and diligence, best interests.
- **Reliance on vendor narrative without independent challenge** is itself flagged as a gap.
- **FAR-accountable persons** must understand AI use within their remit, the specific risks, and keep them within Board-approved appetite.
- **Super trustees**: discretions requiring personal/human involvement cannot be fettered by AI.

TWO HORIZON RISKS REGULATORS HAVE NOT YET NAMED

Embedded AI and cryptographic agility.

EMBEDDED / SILENT UPDATES

Chrome ships **Gemini Nano (~4 GB)** to endpoints by default; opt-out is enterprise-policy gated, not opt-in.

Procurement-style AI controls bind only deliberate adoption. They do not see AI shipped via routine software updates.

CPS 230 contractual uplift now needs explicit AI clauses (default-state, opt-out, data-handling).

CRYPTOGRAPHIC AGILITY / PQC

AU banks publicly preparing **PQC migration for card payments by 2030**, per AFR.

"Harvest now, decrypt later" — exposure starts before CRQC arrival.

APRA's 30 April letter is silent on cryptographic agility; whether CPS 234 / CPS 230 reach this remains an open governance question.

THE DIAGNOSTIC TO RUN

Six questions for boards and accountable persons.

01 **AI risk literacy** at board level — sufficient to challenge management?

02 Alignment to the **Risk Appetite Statement** — explicit, not implied?

03 **Independence** from vendor narrative on capability and risk?

04 **Super-trustee discretion boundaries** — where is human involvement non-fetterable?

05 **FAR-accountable-person evidence** — who owns AI risk within whose remit?

06 **Three-lines technical capability** for continuous AI assurance — including third- and fourth-party dependencies under CPS 230.

What to watch

The signals that move governance from supervision to enforcement.

OPEN THREADS

Six signals worth tracking.

- How APRA exercises "stronger supervisory action" in practice — **first enforcement signals**.
- Whether OAIC issues AI-specific guidance vs. continuing to clarify existing laws.
- Industry response to **CPS 230 + AI**: contract uplift, concentration-risk mapping.
- Whether APRA supervision extends to **embedded AI** (AI an entity runs by default without ever choosing).
- Whether APRA / ASIC issue an explicit **cryptographic-agility / PQC-migration** expectation analogous to the 30 April AI letter.
- Whether AU regulators name **MCP-specific authentication risk** as a control expectation, given CyberCX's bidirectional-data-flow observation.

SOURCES

Citations

§ [[2026-05-08-apra-ai-governance]] — MinterEllison synthesis of APRA's 30 April + ASIC's 8 May letters

§ [[2026-04-21-firefox-mythos-zero-days]] — Mozilla on Claude *Mythos* and the offence-defence rebalance

§ [[2026-05-10-vizza-chrome-silent-llm]] — Tony Vizza on Chrome shipping Gemini Nano by default

§ [[2026-05-12-cybercx-2026-threat-report]] — CyberCX 2026 annual threat report

§ [[2026-05-11-afr-quantum-banks]] — AFR on AU banks' paired quantum + AI cyber-threat exposure

§ [[2026-05-12-cybercx-2026-hack-report]] — CyberCX STA 2026 Hack Report

See also

§ [[ai-security-defense]] — dedicated dossier on the offence/defence rebalance