THE TRUST AI FRAMEWORK™

The Cognitive Debt Briefing.

A board-ready briefing on the hidden risk accumulating inside your AI strategy — and the personal liability exposure it creates under CPS 230 and FAR.

Prepared by
Trust AI™

Author
Anam Munshey

Audience
Boards, CROs, Heads of Risk, Accountable Persons (FAR)

Classification
Board Briefing — For Discussion

EXECUTIVE SUMMARY

The Gap Between AI Speed and Governance Readiness

Australian financial institutions are adopting artificial intelligence at an unprecedented pace. GenAI tools are being deployed across member services, claims processing, investment analysis, and operational workflows — often faster than governance structures can evaluate, contain, or oversee them.

This gap — between the speed of AI adoption and the organisation's capacity to govern it — is what we call Cognitive Debt.

Unlike technical debt, which lives in code, Cognitive Debt lives in decision-making. It accumulates when leaders approve AI strategies they cannot audit, when risk teams inherit model portfolios they cannot map, and when boards receive assurance reports that no longer reflect the organisation's actual AI exposure.

Left unaddressed, Cognitive Debt compounds — quietly creating regulatory exposure, decision fragility, and operational blind spots that no single audit can resolve.

> *"Cognitive Debt is to AI governance what technical debt is to software engineering — except the interest payments are denominated in regulatory penalties, reputational damage, and personal liability under FAR."*
>
> *— Trust AI Framework™*

This briefing is designed to be shared with boards, risk committees, and accountable persons. It sets out the nature of Cognitive Debt, the regulatory context that makes it urgent, and three questions every board should be asking about their AI governance posture.

It also introduces the Trust AI Framework™ — a structured methodology for resolving Cognitive Debt across governance, innovation, and human judgment.

This briefing addresses three questions every board should be asking:

1. Can we map every AI system touching member outcomes — and name the accountable person for each?
2. If APRA asked us to demonstrate AI operational resilience under CPS 230 tomorrow, could we?
3. Are our people making better decisions because of AI — or just faster ones?

THE REGULATORY CONTEXT

Why This Is Urgent Now

Three intersecting regulatory forces have transformed AI governance from a best-practice aspiration into a personal liability exposure:

APRA CPS 230 — Operational Risk Management

Effective 1 July 2025, CPS 230 requires APRA-regulated entities to identify, assess, and manage material operational risks — including those arising from AI and third-party technology providers. Institutions must demonstrate that critical operations can withstand disruption, and that material service providers (including AI vendors) are subject to appropriate oversight.

AI systems that support or execute activities involving member data, investment decisions, or claims processing will generally be treated as part of critical operations, and so fall within CPS 230's operational risk and resilience expectations.

Financial Accountability Regime (FAR)

FAR establishes personal accountability obligations for senior executives and directors of APRA-regulated entities — building on the foundations of the earlier Banking Executive Accountability Regime (BEAR) and extending across all prudentially regulated sectors. Accountable persons must take "reasonable steps" to ensure their areas of responsibility comply with financial services laws. There is no carve-out for AI-related decisions. If an AI system within your accountability remit contributes to harm, regulatory non-compliance, or member detriment, regulators will look to whether you took "reasonable steps" in overseeing its governance — and liability exposure becomes personal if those steps are found wanting.

Best Financial Interests Duty (BFIM)

For superannuation funds, the BFIM duty requires trustees to act in the best financial interests of members when making decisions about the fund. AI systems that influence member outcomes — from robo-advice to claims triage to investment allocation — must be governed with the same fiduciary rigour applied to any other material decision. This extends to decisions about acquiring, configuring, and monitoring AI systems — not just the outcomes they produce. The duty bites at the point of decision-making and expenditure, connecting AI procurement and configuration choices directly to BFIM exposure. Ungoverned AI is, by definition, a potential breach of this duty.

> The convergence: CPS 230 demands operational resilience for AI systems. FAR makes governance failures personally attributable. BFIM ensures that member-facing AI must meet fiduciary standards. Together, they transform Cognitive Debt from a strategic concern into a concrete compliance exposure, with named individuals expected to evidence how they are reducing it over time.

THE ANATOMY OF COGNITIVE DEBT

Where It Accumulates

Cognitive Debt does not announce itself. It accumulates in the spaces between teams, between approval cycles, and between what the board believes and what is actually happening. These are the five most common accumulation patterns we observe in APRA-regulated entities:

1. Shadow AI Proliferation

Teams adopt AI tools — ChatGPT, Copilot, third-party analytics — without central visibility. Each instance creates an ungoverned decision surface. The risk team cannot map what it cannot see.

> Target state: A single enterprise AI registry with risk ratings and usage approvals.

2. Accountability Gaps

AI systems span multiple business units, but no single accountable person (under FAR) owns the end-to-end governance. When something goes wrong, accountability is diffused — which under FAR means everyone is potentially liable.

> Target state: A RACI map aligning every AI system to a named FAR accountable person.

3. Governance Theatre

Policies exist on paper — an AI ethics statement, a responsible use charter — but lack operational controls, testing regimes, or enforcement mechanisms. The board receives assurance that governance is in place; the reality is that governance is performative.

> Target state: Policies tied to specific controls, testing schedules, and assurance metrics.

4. Third-Party Opacity

AI vendors provide outputs but not transparency. Model behaviour, training data provenance, and drift characteristics are opaque. Under CPS 230, material service provider risk must be actively managed — not passively accepted.

> Target state: Vendor transparency clauses, model documentation, and ongoing drift monitoring.

5. Decision Fragility

Executives make strategic decisions informed by AI-generated analysis they cannot independently verify. The human-in-the-loop becomes a human-rubber-stamping-the-loop. Cognitive Clarity — the capacity to critically evaluate AI outputs — erodes silently.

> Target state: Decision integrity frameworks where leaders can articulate the basis, assumptions, and limitations of AI-informed decisions.

THE BOARD AGENDA

Three Questions Every Board Should Be Asking

Q1. Can we map every AI system touching member outcomes — and name the accountable person for each?

If the answer is no, you have an accountability gap that FAR will not overlook. The first step in resolving Cognitive Debt is achieving complete visibility: a registry of AI systems, their risk classification, their data inputs, and the named individual accountable for their governance under FAR.

Red flag: If no one can produce a current AI inventory within 48 hours, assume significant Cognitive Debt.

Q2. If APRA asked us to demonstrate AI operational resilience under CPS 230 tomorrow, could we?

CPS 230 requires documented evidence of operational resilience for critical operations. If AI systems are embedded in claims, investment, or member services workflows, they are critical operations. The question is not whether AI is in scope — it is whether your resilience documentation reflects the AI that is actually deployed.

Red flag: If AI is embedded in critical processes but not named in CPS 230 resilience plans, assume a documentation gap.

Q3. Are our people making better decisions because of AI — or just faster ones?

Speed without clarity is not a competitive advantage; it is a compounding risk. If your executives cannot articulate how an AI-generated recommendation was produced, what data informed it, and where it might be wrong — then AI is not augmenting human judgment. It is replacing it. And when it fails, the accountability still sits with the human who signed off.

Red flag: If executives routinely approve AI-generated outputs without being able to explain their basis, decision integrity has been compromised.

THE TRUST AI FRAMEWORK™

A Codified Approach to Resolving Cognitive Debt

The Trust AI Framework is a structured methodology for governing enterprise AI — built from 17+ years of financial services transformation experience across global firms including EY and IBM. It addresses Cognitive Debt through three integrated pillars:

PILLAR I

AI Governance

The Controls.

Mapping your AI stack to CPS 230, FAR, and BFIM requirements. Accountability mapping, regulatory gap analysis, and operational resilience frameworks.

> CPS 230 mapping, FAR "reasonable steps" evidence packs, BFIM artefacts.

PILLAR II

Responsible Innovation

The Velocity.

Building safety nets for GenAI adoption. Ethical use policies, safety wrappers for member-facing AI, and innovation guardrails.

> Operational risk controls for new AI use cases and change-management expectations.

PILLAR III

Human Judgment

The Clarity.

Protecting executive decision quality. Cognitive Clarity tools, decision integrity frameworks, and AI literacy programs.

> FAR accountable persons' duty to understand the risks they sign off on.

The Engagement Methodology

01 Assess — Map the terrain.

11-pillar AI Governance Readiness Scorecard. (2–3 weeks)

Deliverable: AI system registry with risk classifications.

02 Activate — Build the controls.

Governance structures and accountability maps. (4–6 weeks)

Deliverable: RACI map for FAR accountable persons and CPS 230 evidence pack.

03 Embed — Change the culture.

Training, decision frameworks, internal capability. (6–8 weeks)

Deliverable: Decision integrity framework and AI literacy program.

04 Scale — Compound the advantage.

Cross-BU extension and independent verification. (Ongoing)

Deliverable: Cross-business-unit governance rollout and assurance reporting.

YOUR NEXT STEP

Your AI governance gap is someone's personal liability.

The Trust AI Framework begins with a 30-minute Framework Alignment Session — a diagnostic conversation to assess where your Cognitive Debt sits, which regulatory obligations are most exposed, and what resolution looks like.

Request a Framework Alignment Session

Scan the QR code or visit:

form.jotform.com/253269232889066

hello@trustai.au

trustai.au

Stay Human in the AI Revolution.

This briefing is for general information and governance discussion only and does not constitute legal advice.