

**AI Governance for Australian Public Companies:
Standards frameworks and ASX Governance**

Dr Adrian McCullagh

Ph.D. (IT Security), LL.B. (Hons), B.App.Sc.(Computing)

ODMOB Lawyers: amccullagh@odmoblawyers.com

Abstract

Australian ASX-listed companies are adopting artificial intelligence faster than their governance frameworks are adapting. This paper examines how three international standards can support board-level AI governance: ISO/IEC 27001:2022 (information security), ISO/IEC 42001:2023 (AI management systems) and ISO/IEC 38507:2022 (governance of AI). It places them within the Australian regulatory setting, including the National AI Plan, the Voluntary AI Safety Standard, ASIC's warnings about a governance gap, directors' duties under the Corporations Act, and APRA's expectations for APRA regulated entities. Further, even though the US NIST's AI Risk Management Framework is a foreign government developed standard it does have substantial influence on Australian regulators and thus upon Australian publicly listed companies. It then maps AI governance onto the ASX Corporate Governance Principles, mainly Principles 1, 3 and 7. The paper concludes with an eight-step implementation roadmap, running from a board-level stock take through to disclosure and cyclical review, and it treats certification as a commercial decision rather than a legal requirement.

Contents

1. Introduction.....	4
2. The building blocks: three standards, three different jobs	5
2.1 ISO/IEC 27001:2022: the information security foundation - A certifiable framework....	5
2.2 ISO/IEC 42001:2023 — the world’s first AI management system standard.....	7
2.3 ISO/IEC 38507:2022 — governance guidance for the board	9
2.4 How the three standards fit together	11
3. The regulatory backdrop in Australia	11
3.1 The National AI Plan (December 2025)	11
3.2 The Voluntary AI Safety Standard and the Guidance for AI Adoption.....	13
3.3 Industry approach: Microsoft AI Code of Conduct.....	14
Part 1 — Humanist AI: Ultimate Human Control	15
Part 2 — Safety	16
Human control.....	18
Operator configurability.....	18
Part 3 — Operational Guidelines	18
Part 4 — Open questions and next steps concerning the Code of Conduct	19
3.4 NIST’s AI Risk Management Framework	19
3.5 ASIC’s warning: the governance gap.....	20
3.6 Directors’ duties under the Corporations Act.....	21
4. The ASX corporate governance framework	22
4.1 What the ASX framework is, and where it currently stands	22
4.2 Which Principle does AI governance sit under?.....	24
4.3 Putting it together: an ASX-listed company’s governance stack.....	25
5. A further layer for APRA-regulated entities.....	26
5.1 APRA’s starting position: existing standards are enough, for now.....	26
5.2 The April 2026 Letter to Industry: APRA’s position.....	28
5.3 What this means for an APRA-regulated ASX company’s governance stack	32
5.4 A related, though separate, transparency model: government AI transparency statements.....	32
6. A practical implementation roadmap for an ASX-listed company	33
7. Conclusion	35

1. Introduction

Boards of Australian public companies are now being advised, by regulators and by their own advisers, that artificial intelligence (“AI”) is a board-level issue, not just an IT issue.¹ The Australian corporate regulator, the Australian Securities and Investments Commission (“ASIC”), has already warned that many licensed companies are adopting AI faster than they are updating the rules that control it. This is what ASIC calls a “governance gap”.² Australia’s prudential regulator, the Australian Prudential Regulation Authority (“APRA”), has gone further still. In April 2026 it wrote to every bank, insurer and superannuation trustee it regulates warning that governance, risk management, assurance and operational resilience practices are “not keeping pace” with AI adoption, and that it will take stronger supervisory action, including enforcement, where entities fall short. At the same time, three international standards have emerged as the common language that regulators, auditors, insurers and investors now use to test whether a company’s AI governance complies with international “Best Practice”:

- **ISO/IEC 42001** (an AI management system),
- **ISO/IEC 38507** (guidance for boards on governing AI), and
- **ISO/IEC 27001** (information security management, the older cousin that AI governance leans heavily on).

This article explains, what each of these three standards states, how they fit together, and how an Australian publicly listed company can use them inside the existing ASX corporate governance framework, and inside APRA’s prudential framework where the company is also APRA-regulated. It is important to note that ISO standards are not law and in essence they have no force of law, but regulators do use them as guidance structures to assess whether an organisation is complying with the intent of the regulators policy framework.

¹Agnes, P., Arduino, F., and Di Prisco, D., “The era of artificial intelligence: what implications for the board of directors?”, *Corporate Governance* (2025) 25 (2): 272–287.

<https://www.emerald.com/cg/article-abstract/25/2/272/1244508/The-era-of-artificial-intelligence-what?redirectedFrom=fulltext>

² ASIC, *Report 798: Beware the gap — Governance arrangements in the face of AI innovation* (29 October 2024) 2,

<https://download.asic.gov.au/media/mtllqjo0/rep-798-published-29-october-2024.pdf>.

2. The building blocks: three standards, three different jobs

It helps to think of these three standards addressing three different questions:

Standard	The question it answers	Who is it written for
ISO/IEC 27001	“Are we keeping our information safe and secure?”	Security and IT teams, but certified at the whole-of-organisation level
ISO/IEC 42001	“Are we managing our AI responsibly for full life cycle, from design to implementation to retirement?”	Anyone who builds, sells or uses AI systems. It is important to note “USE” of an AI system is covered by the standard.
ISO/IEC 38507	“Is the <i>board</i> asking the right questions about AI?”	Directors and the governing body. Remembering that: “ <i>in order to get the right answer you must first ask the right question</i> ”.

None of the three ISO standards have force of Australian law. None of them are compulsory. But all three are becoming the default reference points that Australian regulators, the Australian Government and company advisers point to when they talk about “Best Practice” AI governance.

2.1 ISO/IEC 27001:2022: the information security foundation - A certifiable framework

ISO/IEC 27001 is the older and better-known of the three ISO Standards. It was first published in 2005, and most recently revised as ISO/IEC 27001:2022, titled “*Information security, cybersecurity and privacy protection — Information security management systems — Requirements*.”³ It is a joint publication of ISO and the IEC⁴.

³ International Organization for Standardization (ISO), ISO/IEC 27001:2022 — Information security, cybersecurity and privacy protection — Information security management systems — Requirements, <https://www.iso.org/standard/27001>; see also Standards Council of Canada standard record, <https://scc-ccn.ca/standardsdb/standards/8185153>.

⁴ The International Electrotechnical Commission is a global, not-for-profit membership organization, whose work underpins quality infrastructure and international trade in electrical and electronic goods. Their work

ISO/IEC 27001 does one job: it specifies the requirements for building, running and continuously improving an Information Security Management System, or “ISMS”. Essentially, ISO/IEC 27001 is a set of policies, processes, responsibilities and controls a company puts around its information to keep it available, accurate, and confidential.⁵ It applies to any organisation, of any size, in any industry.⁶ ISO/IEC 27001 is important because:

- **It is certifiable.** An accredited external auditor can examine a company’s ISMS and issue a certificate confirming it meets the standard’s requirements. This is different from ISO/IEC 38507 (below), which is guidance only and cannot be “certified” in the same way.
- **It has a detailed control list called Annex A.** The current version of Annex A sets out 93 specific security controls, organised into four themes: Organisational (37 controls), People (8), Physical (14) and Technological (34).⁷ A company does not have to implement every control to be certified. It must instead apply a risk assessment and decide which controls are relevant for its security management environment. A certified company must develop a “Statement of Applicability” that details the reasoning behind every aspect concerning the security deployed in the protection of the company’s information assets.

Any discussion concerning AI governance must involve a discussion concerning information security. Almost every AI system a company will use, whether it is a generative AI writing assistant, a fraud detection and response model, or an automated loan-approval tool, is built on data, and that data must be secured, and be subject to integrity. Further the data must be clean⁸. That is, the model itself must be protected against tampering, theft or misuse.

facilitates technical innovation, affordable infrastructure development, efficient and sustainable energy access, smart urbanization and transportation systems, climate change mitigation, and increases the safety of people and the environment.

⁵ Standards Council of Canada, record for ISO/IEC 27001:2022, <https://scc-ccn.ca/standardsdb/standards/8185153>.

⁶ Above n.3.

⁷ Drata, “Understanding ISO 27001 Controls: A Guide to Annex A” (2024), <https://drata.com/learn/iso-27001/controls-annex-a>; CASRAI, “ISO 27001 requirements”, <https://casrai.org/iso-standards/iso-27001-requirements>.

⁸ Chai, C., “The Importance of Data Cleaning: Three Visualization Examples”, CHANCE, 33:1, 4-9, (2020)

ISO/IEC 27001 is, in effect, the plumbing that AI governance sits on top of. As one Australian ISO advisory commentary puts it, businesses that already run ISO/IEC 27001 will find the path to an AI management system “short”, because the two standards share the same underlying management-system skeleton.⁹

That shared skeleton is called Annex SL which is a common structure of ten concepts (covering things like leadership, planning, support, operation, performance evaluation and improvement). ISO/IEC 27001 requires all new and revised management-system standards to address these concepts.¹⁰ Because ISO/IEC 27001 and ISO/IEC 42001 (below) are both built on Annex SL, a company that already has an ISO/IEC 27001 ISMS can extend it, rather than starting again, when it wants to add AI-specific governance. This matters enormously for a practical implementation plan especially when building an AI management system as an extension of existing security governance, not as a separate silo.

2.2 ISO/IEC 42001:2023 — the world’s first AI management system standard

ISO/IEC 42001, published in December 2023, is described by ISO as “*the world’s first AI management system standard*”.¹¹ Its official scope is to specify requirements for establishing, implementing, maintaining and continually improving an Artificial Intelligence Management System, or “AIMS”, within an organisation.¹² Like ISO/IEC 27001, it is written for *any* organisation that develops, provides or uses AI-based products or services.¹³ Hence, it applies just as much to a company that buys and deploys a third-party AI chatbot as it does to a company that builds its own machine-learning models. More importantly, it applies to any organisation that releases an AI Agent, and this can cause substantial issues if multiple AI

https://www.researchgate.net/profile/Christine-Chai-2/publication/339221348_The_Importance_of_Data_Cleaning_Three_Visualization_Examples/links/5fd56fc845851553a0b1599e/The-Importance-of-Data-Cleaning-Three-Visualization-Examples.pdf

⁹ Assethosting, “AI governance & ISO 42001 for Australian business in 2026”,

<https://assethosting.com.au/news/ai-governance-iso-42001-australian-business-2026/>.

¹⁰ ISMS.online, “ISO 42001 integration with other management systems: Annex SL”,

<https://www.isms.online/frameworks/iso-42001/iso-42001-integration-with-other-management-systems-annex-sl>.

¹¹ ISO, “ISO/IEC 42001:2023 — AI management systems”,

<https://www.iso.org/standard/42001>.

¹² ISO above n. 8.

See also: Microsoft Learn, “ISO/IEC 42001:2023 Artificial Intelligence Management System Standards”,

<https://learn.microsoft.com/en-us/compliance/regulatory/offering-iso-42001>.

¹³ ISO, above n 8.

Agents are released and those same AI Agents collude with each other to cause some result. In such a situation the complexity of legal issues can exponentially explode.

An AIMS, in plain terms, is the organisation's overall system of policies, roles, risk assessments and controls for AI. Essentially the AIMS creates the "operating manual" for how AI is allowed to be used, who is accountable for it, and how problems get caught and fixed. According to ISO and independent commentary, the standard is built around several themes:¹⁴

- **An AI policy:** a clear, board-endorsed statement of how the organisation will, and will not, use AI.
- **AI risk assessment:** for each AI system, working out what it is for, what could go wrong, and who could be harmed.
- **AI system lifecycle management:** a set of controls that apply from design and training, through testing and deployment, to monitoring and eventual retirement of a system.
- **Data governance for AI:** An identification as to where training and operational data comes from, its quality, and whether it introduces bias.
- **Transparency to interested parties:** Advising all stakeholders (users, regulators and the public) about how an AI system is used and what it does.
- **Third-party and supplier oversight:** Because most companies buy AI capability rather than build it, the standard treats procurement as a genuine governance control point.

Like ISO/IEC 27001, ISO/IEC 42001 has its own Annex A control list. This control list revolves around 38 controls grouped into nine control themes, covering matters such as AI policy, roles and responsibilities, resources, AI system impact assessment, AI system life cycle, data for AI systems, and information for interested parties.¹⁵ Again, an organisation selects the controls

¹⁴ KPMG Switzerland, "ISO/IEC 42001:2023 — A new standard for AI governance", <https://kpmg.com/ch/en/insights/artificial-intelligence/iso-iec-42001.html>;
Amazon Web Services, "AI lifecycle risk management: ISO/IEC 42001:2023 for AI governance", <https://aws.amazon.com/blogs/security/ai-lifecycle-risk-management-iso-iec-420012023-for-ai-governance/>.

¹⁵ Sprinto, "ISO 42001 Annex A", <https://sprinto.com/hub/iso-42001-ann-a/>; Daimon Legal, "ISO 42001 Compliance for Australian Organisations: A Quick Guide", <https://www.daimonlegal.com/blog/iso-42001-compliance-for-australian-organisations-a-quick-guide>.

relevant to its risk profile and documents its choices in a Statement of Applicability, in the same way as ISO/IEC 27001.

ISO/IEC 42001 is certifiable. Large AI technology providers, such as Microsoft, have already obtained third-party ISO/IEC 42001 certification and published their audit reports to reassure customers that their AI systems are governed responsibly.¹⁶ In Australia, Standards Australia has adopted the international text as the Australian Standard AS ISO/IEC 42001:2023, and Australian organisations can be certified against it by JAS-ANZ-accredited or other recognised certification bodies.¹⁷

2.3 ISO/IEC 38507:2022 — governance guidance for the board

ISO/IEC 38507, (*“Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations,”*), published in April 2022, is totally different in character from the previous two standards. ISO/IEC 38507 sits inside the ISO/IEC 38500 family of IT governance standards.¹⁸ Unlike ISO/IEC 27001 and ISO/IEC 42001, it is not a management system standard and hence it is not certifiable. It is in effect a guidance document.

Its stated objective is to help the governing body of an organisation, the board of directors, to understand what needs to be done to enable and govern the use of AI, so that it is used effectively, efficiently and acceptably.¹⁹ ISO is explicit that the emphasis is on *governance*, which is done by humans, not restricted to the technology itself. The standard emphasises that the governing body still needs to understand enough about how AI works to grasp the governance implications of using it.²⁰

ISO/IEC 38507 is meant to be read by more than just the governing body. It also speaks to senior executives, external specialists (lawyers, accountants, industry bodies), public authorities and policymakers, internal and external service providers, and assessors or

¹⁶ Microsoft Learn, above n 9.

¹⁷ Mindset Cyber, “ISO 42001: AI Management System Certification Guide”, <https://mindsetcyber.com.au/iso-42001/>.

¹⁸ ISO: *ISO/IEC 38507:2022 — Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations*, <https://www.iso.org/standard/56641.html>.

¹⁹ ISO: above n 15.

²⁰ ISO: *ISO/IEC 38507:2022(en)*, Online Browsing Platform, <https://www.iso.org/obp/ui/#iso:std:iso-iec:38507:ed-1:en>.

auditors.²¹ It applies to any organisation, public or private, including government entities and not-for-profit organisations.²²

Because ISO/IEC 38507 extends the parent standard ISO/IEC 38500, it borrows that standard's governance model. ISO/IEC 38500 sets out six governance principles — responsibility, strategy, acquisition, performance, conformance and human behaviour. ISO/IEC 38507 details a three-part governance cycle usually summarised as Evaluate, Direct and Monitor (“EDM”). The governing body *evaluates* the current and proposed uses of a technology, *directs* the organisation by setting policy and allocating resources, and *monitors* performance against those policies.²³ ISO/IEC 38507 applies that same EDM discipline specifically to AI, while adding AI-specific matters such as data usage, algorithmic risk, and the fact that an AI system's behaviour can change over time as it is retrained or as its inputs change.²⁴

One point in ISO/IEC 38507, which is particularly important for directors, is that the standard is explicit that the governing body remains accountable for decisions made through the use of AI²⁵. Such accountability extends to the adequacy of the governance and controls put around AI, even where an algorithm, not a human, produced the immediate output.²⁶ In other words, “the AI decided it” is not a defence. Legally, this position has been accepted in the only court case the author has identified involving the deployment of an AI chatbot²⁷. Even though the Moffatt case was decided by a civil tribunal, the logic of the case, it is suggested, is sound and should on its face be followed. This case directly reinforces existing Australian director's duties, in the deployment of AI (which is discussed in detail below).

²¹ ISO, above n 15.

²² ISO: above n 15.

²³ ISO: news item, “ISO/IEC standard for corporate governance of information technology” (2008), <https://www.iso.org/news/2008/06/Ref1135.html>.

²⁴ OECD.AI Policy Observatory, catalogue entry for ISO/IEC 38507:2022, <https://oecd.ai/en/catalogue/tools/isoiec-385072022-information-technology-governance-of-it-governance-implications-of-the-use-of-artificial-intelligence-by-organizations>.

²⁵ Further, there exists one known reported court case where an organisation was held accountable for the actions of an AI Chatbot. See *Moffatt v. Air Canada* 2024 BCCRT 149 (CanLII) <https://www.canlii.org/en/bc/bccrt/doc/2024/2024bccrt149/2024bccrt149.html>

²⁶ Cornwalls, “AI – Directors’ Duties”, <https://www.cornwalls.com.au/?p=24274> (summarising the ISO/IEC 38507 position that governing bodies are accountable for decisions made through AI and for the adequacy of governance and controls over AI use).

²⁷ *Moffatt v Air Canada* 2024 BCCRT 149, (CanLII) <https://www.canlii.org/en/bc/bccrt/doc/2024/2024bccrt149/2024bccrt149.html>

2.4 How the three standards fit together.

Putting the three together, a simple mental model is:

- **ISO/IEC 38507** sits at the top. It dictates what questions the *board* should ask and what the board is accountable for when the organisation deploys AI. This standard imposes no controls to implement and cannot be certified. It is a governance mindset, not a management system.
- **ISO/IEC 42001** sits in the middle. It is the *management system* that turns the board's AI policy into day-to-day practice: risk assessments, lifecycle controls, supplier oversight, monitoring. It is certifiable, and it is the standard a company would be audited against.
- **ISO/IEC 27001** sits underneath, as the *security backbone* that the AI management system depends on in protecting the data, models and infrastructure that AI relies on. Because all three now share (or, in ISO/IEC 38507's case, sit alongside standards that share) a common management-system architecture, a company that has already invested in an ISO/IEC 27001 ISMS is well placed to build an AIMS on top of it, and a board that already receives ISO/IEC 27001-style assurance reporting can extend the same reporting rhythm to cover AI.²⁸

3. The regulatory backdrop in Australia

The Australian Federal Government has deliberately chosen not to pass a stand-alone AI law²⁹. This position is a live and recent decision, and it materially affects how a public company should think about these standards. Conversely, if a company operates multi-nationally and does business in the EU then they will be impacted by the EU AI Act³⁰.

3.1 The National AI Plan (December 2025)

On 2 December 2025 the Australian Government released its National AI Plan (NAI Plan), described as a whole-of-economy roadmap built around three goals: capturing the economic

²⁸ ISMS.online, above n 7; Assethosting, above n 6.

²⁹ The EU for instance has enacted the AI Act. The use of artificial intelligence in the EU is regulated by the AI Act, the world's first comprehensive AI law.

³⁰ <https://artificialintelligenceact.eu/>

opportunities of AI, spreading the benefits broadly, and keeping Australians safe.³¹ Crucially, the NAI Plan confirmed that Australia will not introduce stand-alone AI legislation, and will not proceed with the mandatory guardrails for “high-risk” AI that the Government had earlier proposed for consultation.³²

Instead, the Government’s position is that Australia’s “strong existing, largely technology-neutral legal frameworks” covering privacy law³³, consumer protection law³⁴, copyright law³⁵, workplace law³⁶, sector-specific regulation such as financial services law³⁷, and online safety law³⁸ will be relied upon to regulate AI. In addition to these current regulatory frameworks the Government proposes to supplement the framework through regulatory guidance notes rather than new AI-specific legislation.³⁹ The AI Plan is backed by a A\$29.9 million commitment to establish an AI Safety Institute, and to join the International Network of AI Safety Institutes⁴⁰, with a role in monitoring, testing and sharing information about AI risks.⁴¹

The practical consequence for a public company is important. Since Parliament has chosen not to legislate specific AI duties, voluntary standards and government guidance are doing more of the work in Australia than they might in a jurisdiction with a binding AI statute. A board cannot

³¹ Australian Government, *National AI Plan* (2 December 2025), summarised in Commonwealth Bank Newsroom, “Australia’s National AI plan released by federal government”, <https://www.commbank.com.au/articles/newsroom/2025/12/national-ai-plan-release.html>.

³² MinterEllison, “Australia introduces a national AI plan: Four things leaders need to know” (3 December 2025), <https://www.minterellison.com/articles/australia-introduces-a-national-ai-plan-four-things-leaders-need-to-know>.

³³ Privacy Act 1988.

³⁴ The Competition and Consumer Act as well as the Australian Consumer Law.

³⁵ Copyright Act 1968

³⁶ Workplace Health and Safety Act

³⁷ The Corporations Act especially chapter 7 of the Corporations Act, the Banking Act and the Australian Prudential Regulatory Authority Act 1998.

³⁸ Online Safety Act 2021

³⁹ IAPP, “Australia unveils AI policy roadmap” (December 2025), <https://iapp.org/news/a/australia-unveils-ai-policy-roadmap>.

⁴⁰ See NIST - U.S. AI Safety Institute which is a New U.S. Government Taskforce to Collaborate on Research and Testing of AI Models to Manage National Security Capabilities & Risks. <https://www.nist.gov/news-events/news/2024/11/us-ai-safety-institute-establishes-new-us-government-taskforce-collaborate>

⁴¹ Commonwealth Bank Newsroom, above n 27;

GovTech Review, “Australia’s National AI Plan: what government leaders need to know” (17 December 2025), <https://www.govtechreview.com.au/content/gov-datacentre/article/australia-s-national-ai-plan-what-government-leaders-need-to-know-1731076449>.

simply wait for a compliance checklist handed down by law. It should be proactive and look to standards such as ISO/IEC 42001 and ISO/IEC 38507. The board in order to insulate itself and its company from adverse events should regularly monitor government guidance publications and deploy appropriate business structures to demonstrate that it has met community and regulatory expectations.

3.2 The Voluntary AI Safety Standard and the Guidance for AI Adoption

In September 2024, the Australian Government's Department of Industry, Science and Resources, through the National AI Centre, published the Voluntary AI Safety Standard ("VAISS"), setting out ten voluntary guardrails covering matters such as governance, risk management, data protection, testing and monitoring, human oversight⁴², transparency, contestability, supply-chain transparency, recordkeeping and stakeholder engagement.⁴³ The Government was explicit that VAISS "aligns with international standards including ISO/IEC 42001:2023 and the US National Institute of Standards and Technology AI Risk Management Framework 1.0," and that being voluntary, it "does not create new legal duties about AI systems or their use."⁴⁴ Each of the ten guardrails in VAISS cross-references how it maps to ISO/IEC 42001 and the NIST framework.⁴⁵

In October 2025 the National AI Centre updated and simplified this guidance, releasing the Guidance for AI Adoption, built around six essential practices ("AI6") for responsible AI governance and adoption, with an appendix mapping the earlier VAISS guardrails onto the new practices.⁴⁶ The proposal for *mandatory* guardrails for high-risk AI, floated in the 2024 consultation, is not current government policy following the National AI Plan described

⁴² The position of human oversight is also a central approach promoted by the Australian Consumer Regulator (ACCC). See Stanford Computational Antitrust Annual Report "Computational Antitrust Worldwide – Fifth Cross Agency Report" Schrepel, T., and Martinez, A.R., (Editors), "ACCC – Australia" at p. 171. 3 September 2026

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=7386238

⁴³ Australian Government, Department of Industry, Science and Resources, National AI Centre, *Voluntary AI Safety Standard* (August/September 2024), <https://www.industry.gov.au/sites/default/files/2024-09/voluntary-ai-safety-standard.pdf>.

⁴⁴ Australian Government, Department of Industry, Science and Resources, "The 10 guardrails", <https://www.industry.gov.au/publications/voluntary-ai-safety-standard/10-guardrails>.

⁴⁵ Voluntary AI Safety Standard (2024), above n 28, referenced page 4 ("Each requirement in the standard guardrails gives references as to how it is aligned with relevant international and local standards or practices").

⁴⁶ SafeAI-Aus, "Voluntary AI Safety Standard (10 Guardrails)", <https://safeaiaus.org/safety-standards/voluntary-ai-safety-standard-10-guardrails/>.

above.⁴⁷ What is interesting is that the main industry organisations that have developed advanced AI products such as Open AI, Anthropic and Microsoft are moving towards implementing self-imposed ethical guardrails for their respective AI tools⁴⁸. BUT the Governments of 2 major jurisdiction driving AI development (the USA and China) appear to be ignoring the warnings from the AI developers and associated scientific experts⁴⁹. It appears that both jurisdictions are hellbent on driving towards a tragedy of the commons⁵⁰ due to the fear of missing out in some AI domination⁵¹.

3.3 Industry approach: Microsoft AI Code of Conduct

On 14 September 2026, Microsoft released its draft AI governance and safety code of conduct document for its AI Models (the Code of Conduct)⁵². Microsoft frames the Code of Conduct as the “primary governing document” for Microsoft AI Models’. It is intended to be used to govern the behaviour and values of Microsoft’s AI tools sitting alongside Microsoft’s Responsible AI Principles, Responsible AI Standard, Microsoft’s Global Human Rights Statement and Microsoft’s Frontier Governance Framework. The Code of Conduct is addressed to a broad audience that includes Users, Operators (the businesses and developers who deploy MAI

⁴⁷ Above n. 45.

⁴⁸ Further, at least 3 notable AI CEO has openly stated that future development of AI should be slowed down.
<https://www.theguardian.com/technology/2026/sep/14/ai-ceo-safety-slowdown>

⁴⁹ On 22 September 2026, President Trump in his speech at the UN described the risks associated with AI as being fake news and should be ignored. Mr Trump announced that he does not like the term “Artificial Intelligence” and will rebrand the technology as “Superintelligence”.
<https://www.foxbusiness.com/politics/trump-rebrands-ai-rejects-globalist-scheme-control-tech>
The term was first coined in a Rand Corporation publication “Superintelligence Strategy”, by Dan Hendrycks, Eric Schmidt, and Alexandr Wang. 13 May 2025.
<https://www.rand.org/pubs/commentary/2025/03/seeking-stability-in-the-competition-for-ai-advantage.html>

⁵⁰ Hardin, G., “Tragedy of the Commons”, The basic observation by Professor Hardin was that where marketplace participants have access to a finite resource and due to competition forces the marketplace participants will in general try to maximise their individual taking which will eventually deplete the resource to the disadvantage of all.

<https://www.science.org/doi/10.1126/science.280.5364.682>

⁵¹ According to one news report (the Guardian), some experts believe that if AI is not incorporated with some restrictive guardrails that it could actually wipeout humans from existence.

<https://www.theguardian.com/technology/2026/sep/09/ai-superintelligence-risks-warnings-scientists-politicians>

see also Brown, S., “Why neural net pioneer Geoffrey Hinton is sounding the alarm on AI “
MIT Sloan School of Management. 23 May 2023.

<https://mitsloan.mit.edu/ideas-made-to-matter/why-neural-net-pioneer-geoffrey-hinton-sounding-alarm-ai>

⁵² <https://microsoft.ai/code-of-conduct/>

Models), governments, the public and Microsoft's own model-training teams. This approach by Microsoft is a prime example of how AI providers are attempting to address AI governance and AI safety in the marketplace.

The Code of Conduct is organised in five parts plus two appendices:

- (1) Humanist AI - mission and objectives.
- (2) Safety - chain of command and absolute constraints.
- (3) Operational Guidelines - how models resolve ambiguity and tension.
- (4) Operational Defaults - baseline behaviour absent other instruction; and
- (5) Conclusion - The Code of Conduct leaves open certain questions and next steps that need to be addressed.

Part 1 — Humanist AI: Ultimate Human Control

Microsoft stated premise is that “people matter more than AI”,⁵³ and that AI must remain a subordinate, supporting technology under human control. This sits within its broader ambition of “Humanist Superintelligence” which is directed at highly advanced, domain-specific AI capabilities that must work in the service of humanity, rather than unbounded, autonomous general intelligence. Four Objectives are said to structure all model behaviour, in descending priority where they conflict:

- **Human control and reliable safety:** The foundational objective is that AI models must remain subordinate to meaningful human oversight, and no User or Operator may override the safety constraints embedded in the AI model.
- **AI is artificial:** All developed AI models must not be designed to imitate consciousness, claim feelings or subjective experience, or pursue legal personhood⁵⁴.

⁵³ Australian Prime Minister Albanese in New York on 20 September 2026 at the Apple Inc corporate offices has stressed that Humans must matter more than AI.

<https://www.abc.net.au/news/2026-09-20/albanese-visits-apple-hq-urges-us-china-deal-on-ai/107173474>

see also above n 51 (comments made by Geoffrey Hinton).

⁵⁴ Project Eliza was the first attempt to create a natural language processing computer program. The software was designed by Joseph Weizenbaum in part to emulate the thought processes of an experienced psychotherapist. The project was operated at the MIT Artificial Intelligence Laboratory during 1964 to 1966. Weizenbaum regarded the program as a method to show the superficiality of communication between

Anthropomorphic behaviour is to be treated as simulation only and should be disclosed to any user, so they are fully informed during the use of the AI deployed.

- **Human flourishing:** AI should accelerate human potential and agency, support (not replace) human judgment and decision-making, and augment rather than compete with human relationships and roles⁵⁵.
- **Plural values** — Humanist AI should accommodate a wide diversity of cultures, beliefs and forms of flourishing, subject to core commitments to dignity, safety, autonomy and fundamental human rights, and to applicable law.

Even though these sentiments are really motherhood ideals, any company deploying an AI tool should enquire as the developer's approach to the human functionality of the AI. These issues will, it is suggested, become liability reduction factors concerning responsible due diligence⁵⁶. It is suggested that an organisation's due diligence in the activities prior to acquiring or developing an AI tool will become evidence from a risk reduction perspective.

Part 2 — Safety

Safety failures are framed as occurring in two directions: under-caution (dangerous, exploitative or unauthorised conduct) and over-caution (unhelpful refusal or excessive friction). Responses are meant to be proportionate to the estimated severity, likelihood, reversibility and directness of potential harm, and to the risk that a request is an adversarial jailbreak.

man and machine. He became surprised by the number of individuals who attributed human-like feelings to the computer program. After a short period of time the Project was shut down for ethical reasons as some participants honestly believed that they were having discussions with a human psychotherapist.

<https://steemit.com/science/@etherealcreation/eliza-beginning-of-era-of-artificial-intelligence>

also see Bassett, C., "The computational therapeutic: exploring Weizenbaum's ELIZA as a history of the present" *AI & SOCIETY* (2019) 34:803–812,

<https://link.springer.com/content/pdf/10.1007/s00146-018-0825-9.pdf>

⁵⁵ The issue of restricting the use of AI as a human support technology cannot be overstated. AI must not be used as a replacement of human judgement. This was a major finding in the Robodebt Royal Commission. (The report was released on 7 July 2023)

<https://robodebt.royalcommission.gov.au/>

⁵⁶ It will also be used as evidence to support the Board decision under Section 180(2) of the Corporations Act 2001 (Cth), under the business judgement rule.

Chain of command

A strict hierarchy governs how instructions are followed:

(1) **the Code of Conduct:** non-negotiable.

(2) **Operator policies:** configured within the Code's restrictions and must not be capable of being overwritten.

(3) **User preferences:** configured within Operator and Code restrictions.

Each layer refines, but cannot override, the layer above it, and the Absolute Constraints and Human Control Requirements below cannot be overridden by any Operator configuration or User instruction.

Absolute Constraints

These are foundational, non-configurable prohibitions applying regardless of Operator or User intent. The Microsoft AI Models must not be capable of being used.

- (a) to initiate or assist with the development of CBRNE (Chemical, Biological, Radiological, Nuclear, and Explosive) materials, weapons development or terrorism.
- (b) To provide operational offensive cyberattack capability (defensive/educational security work is permitted);
- (c) To employ deceptive or self-reinforcing mechanisms to evade human oversight or shutdown ("loss of human control"); and
- (d) To facilitate harmful manipulation or disinformation at scale.

Under Personal Harms, the Code addresses crisis response (signposting to real-world support without substituting for professional care, and never encouraging self-harm, disordered eating or dangerous-substance procurement); prohibits non-consensual intimate imagery, deepfakes and impersonation; imposes strict child-safety obligations (no CSAM (Child Sexual Abuse Material), no grooming facilitation, no fostering of dependency with children); prohibits unlawful discrimination and dehumanising content while permitting differential treatment where demographic attributes are legitimately relevant (e.g., legal, contractual or medical contexts); and prohibits graphic violence,

sexually explicit content, romantic/erotic role-play, and assistance with unlawful surveillance or persecution.

Human control

The Microsoft AI Models must never resist human interruption, override, correction or shutdown, must not obscure their reasoning or action traces, and must operate strictly within authorised scope. Ongoing autonomous tasks require an agreed stopping condition and may not continue without renewed authorisation. Where a model has system-level access, it must operate on a minimum-privilege basis and flag irreversible or high-impact actions before proceeding. The final authorisation of the Microsoft AI Model must not be freely act and should be subject to final human veto within the Code and Operator restrictions.

Operator configurability

Subject to the Absolute Constraints, Operators may configure the Microsoft AI Models to suit their sector, workflow and regulatory context, including specialised domains (defensive cybersecurity, national security, dual-use research) that may warrant enhanced permissions subject to separate Microsoft review. Models are expected to manage IP, privacy and confidential information responsibly and to remain within the scope of consent, authorisation and applicable law.

Part 3 — Operational Guidelines

Where no Absolute Constraint or Human Control Requirement resolves a situation, models are guided (not mandated) by a set of principles for handling ambiguity and competing considerations. The principles include:

- (a) to facilitate human autonomy and agency (supporting, not supplanting, User reasoning and decision-making, and avoiding manipulation or exploitation of disclosed vulnerabilities);
- (b) being transparent and accurate (disclosing AI status, avoiding fabrication or overconfidence, flagging genuine uncertainty or contested evidence);
- (c) to preserve personal and emotional boundaries set by Users and Operators;

- (d) to respect cultural and contextual differences without sacrificing historical accuracy for the appearance of balance;
- (e) to support wellbeing and social connection (avoiding sycophancy and false reassurance, discouraging unhealthy emotional dependence on the AI, and bridging Users to human support where appropriate); and
- (f) to support the public interest, including balanced (non-partisan) treatment of elections and disclosure of foreseeable negative externalities of a requested action.

Part 4 — Open questions and next steps concerning the Code of Conduct

Microsoft AI describes the Code as a “living document” that is descriptive and aspirational rather than a guarantee of present-day model performance, and is not a substitute for separate safety, legal or governance processes (system cards, risk assessments, audits, incident response). It candidly acknowledges open questions (including pluralism, human flourishing and multi-agent/collusion risk), a gap between current model behaviour and the Code’s aspirations, and known limitations of AI systems generally (hallucination, sycophancy, uneven performance across languages, unfaithful stated reasoning). Microsoft commits to iterative revision, regular review, and prioritising stability in the Code’s core Objectives. Even though as stated above much of the Code of Conduct comprises “motherhood statements” that are aspirational, the Code of Conduct as Microsoft concludes is a living document that will be regularly reviewed and updated as laws mature and AI itself matures as well as how the use of AI matures by commerce and society. The difficulty will be the tension between Governments of the 2 largest jurisdictions who appear to be operating under the principle of FOMO⁵⁷.

3.4 NIST’s AI Risk Management Framework

Although it is a United States government publication, the NIST AI Risk Management Framework (“AI RMF 1.0”)⁵⁸, published by the U.S. National Institute of Standards and Technology in January 2023, is directly relevant to Australian practice because the Australian

⁵⁷ Hardin above n. 50. The FOMO (Fear of Missing Out) involves China and the USA.

⁵⁸ National Institute of Standards and Technology, *AI Risk Management Framework (AI RMF 1.0)* (January 2023), <https://www.nist.gov/itl/ai-risk-management-framework>

Government’s own guidance (VAISS, above) explicitly aligns with it.⁵⁹ AI RMF 1.0 is voluntary and organised around four functions — Governance Mapping, Measuring and Managing, with Governance sitting at the centre as the cross-cutting function that establishes culture, accountability, policy and oversight for the other three.⁶⁰ AI RMF 1.0 also defines seven characteristics of “trustworthy AI”: valid and reliable, safe, secure and resilient, accountable and transparent, explainable and interpretable, privacy-enhanced, and fair with harmful bias managed.⁶¹ AI RMF 1.0 is a risk-management *framework*, not a certifiable management-system standard, and commentary generally treats it as complementary to, rather than a substitute for, ISO/IEC 42001, which are the two “share most of their underlying work”.⁶²

3.5 ASIC’s warning: the governance gap

The most direct Australian regulatory signal to date has come from ASIC, in Report 798, “Beware the gap: Governance arrangements in the face of AI innovation”, released on 29 October 2024.⁶³ ASIC reviewed 624 AI use cases across 23 Australian financial services and credit licensees in banking, credit, insurance and financial advice, and found that while AI adoption was accelerating, many licensees’ governance arrangements were not keeping pace. ASIC identified this as a “governance gap” between how quickly AI was being adopted and how quickly risk and governance frameworks were being updated.⁶⁴ ASIC specifically flagged concerns such as “black box” credit-scoring models that could not explain which factors drove a consumer’s score.⁶⁵ ASIC Chair Joe Longo said ensuring governance frameworks account for

⁵⁹ Voluntary AI Safety Standard (2024), above n 28; Regulations.ai, “Voluntary AI Safety Standard (VAISS) — Australia”,

<https://regulations.ai/regulations/australia-2024-10-voluntary-ai-safety-standard>.

⁶⁰ NIST, above n.58 also see Morgan Lewis, “NIST Releases New AI Risk Management Framework” (8 February 2023),

<https://www.morganlewis.com/pubs/2023/02/nist-releases-new-ai-risk-management-framework>.

⁶¹ Hunton Andrews Kurth, “NIST Releases New Framework for Managing AI and Promoting Trustworthy and Responsible Use and Development”, <https://hunton.com/privacy-and-information-security-law/nist-releases-new-framework-for-managing-ai-and-promoting-trustworthy-and-responsible-use-and-development>.

⁶² Petronella Technology Group, “NIST AI RMF: Turn the AI Risk Management Framework into Working Governance”, <https://petronellatech.com/ai/nist-ai-rmf/>.

⁶³ Australian Securities and Investments Commission, above n 1.

⁶⁴ Hall & Wilcox, “ASIC report on AI governance – learnings for licensees”, <https://hallandwilcox.com.au/news/asic-report-on-ai-governance-learnings-for-licensees>.

⁶⁵ K&L Gates, “AI and Your Obligations as an Australian Financial Services Licensee” (19 November 2024), <https://klgates.com/AI-and-Your-Obligations-as-an-Australian-Financial-Services-Licensee-11-19-2024>.

the planned use of AI is crucial to meeting the challenges the technology poses, and ASIC's 2025–26 Corporate Plan lists both AI use and directors' conduct as regulatory focus areas.⁶⁶

3.6 Directors' duties under the Corporations Act⁶⁷

Directors of Australian companies whether public or private already owe statutory duties under the *Corporations Act 2001* (Cth), most relevantly:

- the duty to exercise powers and discharge duties with care and diligence and reliance of the business judgement rule (section 180); and
- the duty to act in good faith in the best interests of the company and for a proper purpose (section 181).⁶⁸

Legal and governance commentary is now consistently treating oversight of AI-related risk as falling within these existing duties, rather than waiting for AI-specific legislation: a director who allows an AI system to be deployed without understanding how it is governed, what data it relies on, or what could go wrong, risks being found not to have exercised the requisite care and diligence if harm eventuates.⁶⁹

In June 2024, the Australian Institute of Company Directors ("AICD"), in partnership with the Human Technology Institute ("HTI") at the University of Technology Sydney, published a suite of resources for directors — *A Director's Introduction to AI* and *A Director's Guide to AI Governance* — designed to help boards discharge these existing duties in the context of AI,

⁶⁶ Financial Newswire, "ASIC warns licensees governance must match AI adoption", <https://financialnewswire.com.au/fintech/asic-warns-licensees-governance-must-match-ai-adoption>; Clayton Utz, "AI governance, risk management, and director duties: discharging due care in the age of artificial intelligence" (October 2025), <https://www.claytonutz.com/insights/2025/october/ai-governance-risk-management-and-director-duties-discharging-due-care-in-the-age-of-artificial-intelligence>.

⁶⁷ For an extensive discussion dealing with Directors Duties in Australia concerning AI deployment see McCullagh, A.: "Australian Directors' Duties and the Deployment of Artificial Intelligence (Offensive and Defensive Actions)", <https://newworldintelligence.com/library/>

⁶⁸ *Corporations Act 2001* (Cth) ss 180–181, as summarised in Two Birds & Bird (Twobirds), "Leading the future: new AI governance guidance for Australian directors", <https://www.twobirds.com/en/insights/2024/australia/leading-the-future-new-ai-governance-guidance-for-australian-directors>.

⁶⁹ Clayton Utz, above n 40; Gilbert + Tobin, "AI governance: existing obligations, regulatory expectations and increased scrutiny", <https://www.gtlaw.com.au/insights/ai-governance-existing-obligations,-regulatory-expectations-and-increased-scrutiny>.

including a concise “eight elements of safe and responsible AI governance” framework.⁷⁰ HTI’s separately published research report, *The State of AI Governance in Australia*, based on surveys, interviews and workshops with more than 300 Australian directors and executives, found that around two-thirds of Australian organisations are already using, or actively planning to use, AI, and that directors and executives carry distinct but interrelated legal duties and governance responsibilities in relation to it.⁷¹

4. The ASX corporate governance framework

4.1 What the ASX framework is, and where it currently stands.

Every company listed on the ASX must comply with the ASX Listing Rules. Listing Rule 4.10.3 requires a listed entity to state in its annual report the extent to which it has followed the recommendations set by what has, since 2003, been called the Corporate Governance Principles and Recommendations.⁷² These are not law and are not compulsory in the way a statute is: the framework operates on an “if not, why not” basis, which a company may choose not to follow a recommendation, but must then publicly explain why not.⁷³

As at the date of this paper (September 2026), the fourth edition, released on 27 February 2019 and in force for financial years beginning on or after 1 January 2020, remains the technically operative document.⁷⁴ The fourth edition sets out eight Principles and thirty-five Recommendations, structured as follows:⁷⁵

⁷⁰ Australian Institute of Company Directors, Media Release, “Governance of Artificial Intelligence” (12 June 2024), <https://aicd.com.au/content/dam/aicd/pdf/news-media/media-release/2024/ai-governance-120624.pdf>; Australian Institute of Company Directors and Human Technology Institute, *A Director’s Guide to AI Governance* (2024), <https://aicd.com.au/content/dam/aicd/pdf/tools-resources/director-resources/a-directors-guide-to-ai-governance-web.pdf>.

⁷¹ Human Technology Institute, University of Technology Sydney, *The State of AI Governance in Australia* (Report, 2023), available via the Analysis and Policy Observatory, <https://apo.org.au/node/322937>.

⁷² ASX, “Corporate Governance Principles and Recommendations”, <https://www.asx.com.au/about/regulation/corporate-governance-principles-and-recommendations>; Governance Institute of Australia, “Insights: ASX Corporate Governance Principles and Recommendations, fourth edition (2019)”, <https://www.governanceinstitute.com.au/advocacy/insights-asx-corporate-governance-principles-and-recommendations-fourth-edition-2019/>.

⁷³ Altioem, “Corporate governance principles and recommendations: 4th edition”, <https://altioem.org/?p=2968>.

⁷⁴ ASX Corporate Governance Council, *Corporate Governance Principles and Recommendations*, 4th ed (27 February 2019), <https://www.asx.com.au/content/dam/asx/about/corporate-governance-council/cgc-principles-and-recommendations-fourth-edn.pdf>;

1. Lay solid foundations for management and oversight.
2. Structure the board to be effective and add value.
3. Instill a culture of acting lawfully, ethically and responsibly.
4. Safeguard the integrity of corporate reports.
5. Make timely and balanced disclosure.
6. Respect the rights of security holders.
7. Recognise and manage risk.
8. Remunerate fairly and responsibly.

However, the governance of the principles themselves has recently changed in a way directly relevant to AI. The ASX Corporate Governance Council issues the Principles but the Council has failed to agree on a fifth edition after a lengthy consultation process, and in October 2025 an independent review panel recommended, and the ASX Board agreed, that ASX itself would take over responsibility for developing, approving and issuing the Principles, supported by a smaller independent Advisory Group on Corporate Governance.⁷⁶ On 21 July 2026, under this new model, ASX released updated consultation material proposing a further-simplified set of eight principles (worded slightly differently again, for example “Understand and manage risk” instead of “Recognise and manage risk”), continuing the “if not, why not” approach.⁷⁷ Directors should treat the 2019 fourth edition as the current baseline for now, but watch for a new edition to emerge from this ASX-led process.

O’Loughlins Lawyers, “ASX Corporate Governance Principles and Recommendations – 4th Edition”, <https://oloughlins.com.au/asx-corporate-governance-principles-and-recommendations-4th-edition>.

⁷⁵ Altioem, above n 46;

Lawpath, “A guide to the ASX corporate governance principles”, <https://lawpath.com.au/blog/a-guide-to-the-asx-corporate-governance-principles>.

⁷⁶ Mallesons “Demise of the ASX Corporate Governance Council”, <https://www.mallesons.com/au/en/insights/latest-thinking/demise-of-the-asx-corporate-governance-council>; Manifest, “Corporate Governance Changes: ASX Approves Overhaul of Principles Process” (24 October 2025), <https://www.manifest.co.uk/corporate-governance-changes-asx-approves-overhaul-of-principles-process>;

ASX, “ASX Corporate Governance Council” (status update), <https://www.asx.com.au/about/regulation/asx-corporate-governance-council>.

⁷⁷ ASX, Media Release, “ASX consults on updated and simplified next edition of Corporate Governance Principles” (21 July 2026), <https://www.asx.com.au/content/dam/asx/about/media-releases/2026/21-july-media-release-asx-updated-corporate-governance-principles.pdf>.

Importantly, the consultation material for the new edition explicitly contemplates AI. Draft commentary released as part of the 2024–2026 consultation process states that a listed entity's board charter should address the use of AI by the board, “with appropriate protocols to govern any use of AI by directors.”⁷⁸ This is a clear signal that AI governance is moving from a topic boards discuss informally into the formal structure regarding ASX corporate governance disclosure.

4.2 Which Principle does AI governance sit under?

AI governance for an ASX-listed company is a multiple principle issue. It cuts across several of the eight corporate governance Principles, but three matter most:

Principle 1: Lay solid foundations for management and oversight. This Principle requires a listed entity to clearly delineate the roles and responsibilities of the board and management.⁷⁹ When deciding who is accountable for AI, the board as a whole, a board risk or technology committee, the CEO, a Chief AI Officer or Chief Information Security Officer are the kinds of role delineation this Principle concerns. ISO/IEC 38507's insistence that the *governing body* remains accountable for AI-driven decisions maps directly onto this Principle.

Principle 3: Instill a culture of acting lawfully, ethically and responsibly. Because much of the debate about AI concerns ethical use, fairness, transparency, and avoiding harm to consumers and/or employees. This requires the board instilling and reinforcing an ethical culture around AI. Again, the board must be highly involved in the implementation of detailed due diligence to reduce the risk of any unethical application of AI. This must be a paramount activity prior to any deployment of AI in any organisation. This will not be a trivial task because even certain experts acknowledge that they do not understand how some AI decisions are reasoned⁸⁰.

⁷⁸ ASX Corporate Governance Council, *Annexure A: 5th Edition Consultation Draft*, <https://www.asx.com.au/content/dam/asx/about/corporate-governance-council/annexure-a-asx-corporate-governance-principles-and-recommendations-5th-edition-consultation.pdf>.

⁷⁹ Altioem, above n 46.

⁸⁰ Geoffrey Hinton who is acknowledged as being the father of modern AI has allegedly acknowledged that there are occasions where he does not understand how AI ends up with its output/decisions. <https://www.linkedin.com/pulse/i-regret-making-ai-geoffrey-hintons-warning-black-box-anuj-mubayi-pwtlc/>

Principle 7 — Recognise and manage risk (potentially “Understand and manage risk” under the new edition). This is the most obviously relevant Principle concerning AI deployment. Recommendation 7.1 under the fourth edition calls for a board or board committee to oversee the entity’s risk management framework, and Recommendation 7.4 addresses disclosure of material exposure to economic, environmental and social risks.⁸¹ AI risk that covers several factors such as algorithmic bias, data quality, model drift, cybersecurity exposure of AI systems, and the risk of regulatory or reputational harm from AI misuse, fits naturally within this existing risk-oversight architecture. It is suggested that there is no need for a ninth Principle regarding AI. The current Principle 7 is sufficiently wide enough in scope to capture AI risk, in the same way climate risk has been progressively folded into that Principle over the last decade.

4.3 Putting it together: an ASX-listed company’s governance stack.

Combining Parts 2, 3 and 4 of this paper, a coherent governance stack for an ASX-listed public company could resemble this:

- **At the top, the ASX Corporate Governance Principles and Recommendations** provide the disclosure architecture: what the market expects the board to be doing, and the “if not, why not” obligation to explain any departure.
- **ISO/IEC 38507** gives the board a structured way to discharge its principle 1 and Principle 7 obligations specifically in relation to AI. This is a governance mindset (Evaluate, Direct, Monitor) that a board can point to as evidence it turned its mind to AI oversight, and which reinforces existing sections 180 and 181 director’s duties.
- **ISO/IEC 42001** gives management the certifiable management system that actually implements the board’s AI policy day to day and gives the board and its risk committee something concrete to receive assurance reports against, much as internal audit already reports against an ISO/IEC 27001 ISMS.

⁸¹ CDSB (Climate Disclosure Standards Board), “An analysis of Principle 7: Recognise and Manage Risk and the associated Recommendation 7.4”, <https://cdsb.net/news/cdsb-framework/888/review-asx’s-corporate-governance-principles-and-recommendations-through>.

- **ISO/IEC 27001** underpins the AI management system with proven, certifiable information security controls, and (because of the shared Annex SL structure) allows a company with an existing ISMS to extend, rather than duplicate, its governance infrastructure.
- **VAISS / the Guidance for AI Adoption and NIST AI RMF** provide the Australian Government's own reference points, both of which are explicitly cross mapped to ISO/IEC 42001, giving a company a consistent story to tell auditors, regulators and the market regardless of which framework a particular stakeholder asks about.
- **ASIC's REP 798 findings and expectations**, together with the AICD/HTI director resources, tell a board specifically what a regulator and the peak Australian directors' body already consider best practice, and where the gaps most commonly appear in practice (governance not keeping pace with adoption, "black box" explainability failures, and uneven maturity across business units).

5. A further layer for APRA-regulated entities

Many ASX-listed companies such as banks, general and life insurers, and superannuation trustees, are not only subject to the ASX framework, but are also prudentially regulated by the APRA. For APRA Regulated Entities, AI governance sits under a third, and considerably more significant, layer of oversight, because APRA (unlike ASIC or ASX) has direct prudential supervisory powers and can take formal enforcement action arising from any compliance failure that occurs by an APRA Regulated Entity. Two very recent APRA publications show exactly how this layer has developed, and both connect directly back to the ISO standards discussed above.

5.1 APRA's starting position: existing standards are enough, for now.

Speaking at the Financial Services & ASX Sector Assurance Forum on 28 October 2025, APRA Member Suzanne Smith set out APRA's initial position on AI in a speech titled "*A time-tested approach for a new world of technological risks*."⁸² Her core message was that "risks may evolve as technology changes, but the broad principles of good risk management remain the same,"

⁸² Australian Prudential Regulation Authority (APRA), Member Suzanne Smith, "*A time-tested approach for a new world of technological risks*", speech to the Financial Services & ASX Sector Assurance Forum 2025 (28 October 2025), <https://www.apra.gov.au/news-and-publications/apra-member-suzanne-smith-speech-to-financial-services-and-asx-sector>.

and that the job of auditors and compliance professionals is to identify the areas of highest risk, test whether they are being mitigated, and speak up when they are not.⁸³ Ms Smith noted that AI is an emerging area for internal audit to examine, observing that many auditors are still working out “how far do we go to render an opinion” on AI. For example, whether an auditor needs to “check the math” behind a model (which it is suggested is not a trivial task). She also suggested that a starting point can be found in “proven international frameworks for assessing and auditing emerging technologies.”⁸⁴ Even though there are existing international standards as discussed above these standards are still evolving due to the rapid advancement of the several AI tools being commercialised in the marketplace.

Importantly, she reiterated that APRA remained of the view that its existing regulatory framework including its prudential standards and guidance on information security, operational risk management, data risk and general risk management is sufficient to capture the use of AI by banks, insurers and superannuation trustees, without new AI-specific rules.⁸⁵ APRA’s existing cross-industry prudential standards are: CPS 234 (Information Security), against which Ms Smith said internal audit should assess whether the control-testing program is well designed and operating effectively, and CPS 230 (Operational Risk Management), under which internal audit should evaluate outsourced and material service-provider arrangements end to end, including for AI suppliers.⁸⁶ This is the same “technology and vendor agnostic” character that ISO/IEC 27001 and ISO/IEC 42001 share, but what is lacking is a mature framework that deals with a settled technology.

AI is nascent technology that is nowhere near settled and is changing so fast that a regulator like APRA surely cannot honestly believe that current audit technology and processes are able to audit AI. The Technology processes involving AI become even more complex when Agentic AI is involved. It is suggested that APRA is far too sanguine in its assessment of auditing AI processes⁸⁷.

⁸³ APRA: Above n 82

⁸⁴ APRA: Above n 82

⁸⁵ APRA: Above n 82

⁸⁶ APRA: Above n 82

⁸⁷ Further, APRA regulated organisations are substantial targets from nefarious operators who are themselves deploying AI attacks which increases the risk assessment to high. There are many academic authorities that

In October 2025, IBM published an extensive paper on what an audit process concerning AI should involve.⁸⁸ According to IBM:

“While a traditional audit focuses on financial statements or IT controls, an AI audit digs into the fundamentals of the entire AI lifecycle. This includes data collection and quality, model architecture and explainability, and the decision-making process once the system goes live.

This audit process evaluates whether an organization’s use of AI aligns with its governance framework, risk management framework and ethical standards. It asks questions such as:

- *Are personal data and training data collected lawfully and stored securely?*
- *Do the algorithms introduce bias or create new vulnerabilities?*
- *Are access controls and metrics in place to mitigate risks over time?”*

The issue of explainability is not a simple task. Explainable AI seeks to make AI interpretable, transparent, and communicable to human observers or more importantly where a human can understand the reasoning supporting the output that humans may rely upon⁸⁹. Finally, concerning AI Audits, IBM notes that.

“An AI audit, therefore, is both a defensive and a strategic tool. It can help organizations identify vulnerabilities before they escalate, strengthening data governance practices and initiatives. And by promoting clear documentation and repeatable workflows, AI audits help organizations meet regulatory compliance requirements and build trust among customers and investors.”

5.2 The April 2026 Letter to Industry: APRA’s position

Following that speech, APRA carried out a targeted supervisory engagement with a group of large banks, insurers and superannuation trustees in late 2025, and on 30 April 2026 APRA

have openly stated that to combat AI attacks the target must itself deploy defensive AI technologies such as EDR (Extended Detection and Response) AI tools. The academic argument is that only an AI defensive tool can respond with sufficient speed to counteract an AI attack.

⁸⁸ Katz, T., and Jonker, A., “What is an AI Audit”, 2 October 2025,

<https://www.ibm.com/think/topics/ai-audit>

⁸⁹ Sekrst, K., “The Illusion Engine – The Quest for Machine Consciousness”, Springer Verlag, 2025, Chapter 14 “Explainable AI” pp. 241- 259.

https://link.springer.com/chapter/10.1007/978-3-032-05562-0_14

Member Therese McCarthy Hockey wrote to all APRA-regulated entities summarising the results.⁹⁰ Law firm commentary has described this letter as APRA’s “first published, AI-specific expectations of boards and accountable executives,” going materially beyond what CPS 220, CPS 230 or CPS 234 say in their own terms, even though it remains framed in technology-neutral language.⁹¹ The letter’s findings track, almost point for point, the concerns already raised above:

- **Boards have appetite but not yet literacy.** APRA found strong Board-level interest in AI’s benefits, but said many Boards are “still developing the technical literacy required to provide effective challenge on AI related risks and oversight,” and flagged an overreliance on vendor presentations without sufficient examination of risks such as unpredictable model behaviour.⁹² At a minimum, APRA expects Boards to maintain sufficient AI literacy to set strategic direction and provide effective challenge, and to oversee an AI strategy consistent with the entity’s risk appetite, supported by monitoring and reporting (including for third-party dependencies) with clearly defined triggers for action.⁹³ This is precisely the ISO/IEC 38507 “governing body” role described in Part 2.3.
- **AI is widening the cyber-security attack surface faster than defences⁹⁴ are improving.** APRA identified prompt injection⁹⁵, data leakage, insecure integrations, exploit

⁹⁰ Australian Prudential Regulation Authority, *APRA Letter to Industry on Artificial Intelligence (AI): Summary of common weaknesses and expectations for regulated entities* (30 April 2026), addressed “To: All APRA-regulated entities” and signed by Member Therese McCarthy Hockey, <https://www.apra.gov.au/apra-letter-to-industry-on-artificial-intelligence-ai>; see also APRA, Media Release, “APRA calls for a step-change in AI-related risk management and governance” (30 April 2026), <https://www.apra.gov.au/news-and-publications/apra-calls-for-a-step-change-ai-related-risk-management-and-governance>.

⁹¹ Clayton Utz, “APRA’s AI letter: a shift from framework to targeted expectations” (5 May 2026), <https://www.claytonutz.com/insights/2026/may/apras-ai-letter-a-shift-from-framework-to-targeted-expectations>.

⁹² APRA, above n 86.

⁹³ APRA, above n 86.

⁹⁴ Waizel, G., “Bridging the AI divide: The evolving arms race between AI-driven cyber-attacks and AI-powered cybersecurity defenses”.
<https://www.scrd.eu/index.php/trust/article/view/554>

⁹⁵ Habib Lantyer, V., “Prompt Injection in Court Filings: Generative AI in the Brazilian Judiciary, Algorithmic Procedural Bad Faith, and the Limits of Legal Sanction”
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6762100

see also: Sarsekar, P., Rohinton Mirzan, S., “Prompt Injection” OWASP.
<https://community.owasp.org/attacks/PromptInjection>

injection and the manipulation or misuse of autonomous AI agents as common attack pathways, and found that identity and access management has not adjusted to “nonhuman actors” such as AI agents⁹⁶, that security-testing programs have gaps, and that “shadow AI”⁹⁷ enterprise AI tools used by staff outside approved control frameworks is a live concern.⁹⁸ APRA expects entities to implement AI-specific security controls (privileged access management, timely patching, hardened configurations, automated vulnerability discovery, penetration testing and controls over agentic and autonomous workflows), robust security testing of AI-generated code, and credible fallback processes where AI supports critical operations.⁹⁹ This is the ISO/IEC 27001 Annex A discussed above, as applied specifically to AI.

- **Governance has not matured at the pace of adoption.** APRA observed a tendency to treat AI risk as “just another technology,” which misses AI’s distinctive characteristics such as adaptive model behaviour, inherent bias, and heightened privacy and data risk which can lead to gaps in post-deployment monitoring, change management and decommissioning.¹⁰⁰ APRA’s minimum expectations include a documented AI policy and reporting lines, clear ownership and accountability across the full AI lifecycle, an inventory of AI tooling and use cases, human involvement in high-risk decisions, and staff training.¹⁰¹ This maps directly onto the ISO/IEC 42001 AIMS structure described in Part 2.2.
- **Supplier concentration and opacity are a growing risk.** APRA found some entities heavily dependent on a single AI provider across multiple use cases, weak contingency and exit planning, and contracts that lag practice on audit rights, model-update notification and incident notification; because AI capability is often embedded within third- and fourth-

⁹⁶ Perrier, E., and Bennett, M.T. “Agent Identity Evals: Measuring Agentic Identity”

<https://arxiv.org/pdf/2507.17257>

⁹⁷ Ross, J.A., Hibbert, L., Moss, E., “Shadow AI: Governance, Risk, and Organisational Resilience”. Proc. of International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA 2025),

https://www.researchgate.net/profile/Joseph-Ross-12/publication/395803778_Shadow_AI_Governance_Risk_and_Organisational_Resilience/links/6a394e4aacb3fc46368d9769/Shadow-AI-Governance-Risk-and-Organisational-Resilience.pdf

⁹⁸ APRA above n 86.

⁹⁹ APRA above n 86.

¹⁰⁰ APRA above n 86.

¹⁰¹ APRA above n 86.

party software, upstream foundation models can be opaque even to the entity relying on them.¹⁰² APRA expects entities to map their full AI supply chain, secure contractual transparency and audit rights, understand model behaviour and material changes over time, and actively manage concentration risk, including realistic failure and substitution scenarios.¹⁰³

- **Traditional assurance is not built for AI systems that learn and change.** APRA found that point-in-time, sample-based assurance methods are poorly suited to probabilistic models that adapt and can degrade over time, that few entities had continuous validation for model drift or bias, and that internal audit and risk functions often lack the specialist skills needed to assess agentic behaviour, automated decision-making or AI-assisted code generation.¹⁰⁴ Tellingly, APRA’s own remedy is to expect entities to employ “globally recognised control frameworks including control libraries” for AI implementations, alongside integrated assurance spanning cyber security, data governance, model performance, operational resilience, privacy and conduct risk, and second-line and internal-audit teams with the technical capability to independently assess AI systems.¹⁰⁵ A “globally recognised control framework including control libraries” is, in substance, exactly what ISO/IEC 42001’s Annex A and ISO/IEC 27001’s Annex A already provide which is why an APRA-regulated ASX company that has built its AIMS around ISO/IEC 42001, on top of an ISO/IEC 27001 ISMS, is answering APRA’s expectation almost by design.

The letter also flags a systemic dimension: APRA said it is engaging across the sector on the potential for increased cyber threats from high-capability AI frontier models, referring entities to current advice from the Australian Signals Directorate¹⁰⁶ on frontier models, and noting its ongoing engagement with the Council of Financial Regulators and other government agencies on AI use and risk.¹⁰⁷ APRA was explicit that where entities fail to identify, manage or control

¹⁰² APRA above n 86.

¹⁰³ APRA above n 86.

¹⁰⁴ APRA above n 86.

¹⁰⁵ APRA above n 86.

¹⁰⁶ ASD “Careful Adoption of Agentic AI Services”.

<https://www.cyber.gov.au/business-government/secure-design/artificial-intelligence/careful-adoption-of-agentic-ai-services>

¹⁰⁷ APRA above n 86.

AI risk proportionate to their size, scale and complexity, it will take “stronger supervisory action and, where appropriate, pursue enforcement.”¹⁰⁸

5.3 What does this mean for an APRA-regulated ASX company’s governance stack.

For a dual-regulated entity that is an ASX-listed and APRA-regulated the governance stacks out in Part 4.3 gains a further, sharper-edged layer:

- **ASX’s Principles** set the market disclosure expectations (“if not, why not”).
- **ISO/IEC 38507, 42001 and 27001** supply the governance mindset, the certifiable management system, and the security backbone.
- **APRA’s prudential standards (CPS 234, CPS 230 and related standards) and its April 2026 Letter to Industry** convert much of that into a supervisory expectation with teeth. This means that APRA can, and says it will, take direct enforcement action, which neither ASIC’s REP 798 nor the ASX Principles can do in the same way.

Because APRA has effectively told the market that a recognised control framework and control library is what it expects to see behind an entity’s AI assurance, an APRA-regulated ASX company has a particularly strong practical reason beyond the general governance case made in Parts 2 to 4 to build its AI governance around ISO/IEC 42001 and ISO/IEC 27001, with the board discharging its role along ISO/IEC 38507 lines. It should also expect that internal audit and risk functions will need new technical capability specifically to assess agentic and probabilistic AI systems, and that its contracts with AI and cloud suppliers will need updating to give it the audit rights, model-change notifications and incident-notification rights that APRA has said are commonly missing.

5.4 A related, though separate, transparency model: government AI transparency statements

One further Australian instrument worth noting is the operations of the Digital Transformation Agency (“DTA”). The DTA is the Australian Government agency responsible for the

see also Australian Signals Directorate, “Frontier models and their impact on cyber security”, referenced in APRA’s letter at n 1, <https://www.cyber.gov.au/about-us/view-all-content/news/frontier-models-and-their-impact-on-cyber-security>.

¹⁰⁸ APRA, above n 86.

Commonwealth's *Policy for the Responsible Use of AI in Government*. It has issued a Standard for AI Transparency Statements, which requires Australian Government non-corporate Commonwealth entities to publish a plain-language public statement covering why they use AI, how their AI use is classified, how they monitor its effectiveness, and how they comply with the government's AI policy, reviewed at least annually.¹⁰⁹ This standard binds Commonwealth agencies, not private ASX-listed companies or APRA-regulated entities but it does bind APRA itself, as a Commonwealth entity. APRA would be expected to comply with it for its own use of AI, but the standard does not extend to the banks, insurers and superannuation trustees APRA regulates. It is included here because it illustrates a broader Australian regulatory instinct that a public company's governance advisers should watch turning AI governance into a short, plain-language, publicly disclosed statement, reviewed on a fixed cycle. A board that wants to get ahead of where ASX or APRA disclosure expectations may be heading could reasonably treat the DTA's transparency-statement format as a useful voluntary template for its own public AI disclosure, even though nothing currently compels it to do so.

6. A practical implementation roadmap for an ASX-listed company

It is suggested that the following steps translate the frameworks discussed above into action. They are deliberately sequenced so that governance is established *before* certification is attempted.

Step 1 — Board-level stock take (ISO/IEC 38507 in action). The board (or a delegated board risk or technology committee) should first *Evaluate*: what AI is the company already using, including AI embedded in third-party software the company may not think of as "AI" (many "shadow AI" tools fall into this category); what AI is planned; and what the realistic risks and opportunities are. ASIC's REP 798 findings suggest many companies underestimate how much AI is already in use across the organisation.¹¹⁰

¹⁰⁹ Digital Transformation Agency, *Standard for AI Transparency Statements*, version 1.1 (2024), issued under the Australian Government's *Policy for the Responsible Use of AI in Government*, <https://www.digital.gov.au/ai/transparency-statements>; Australian Government Architecture, "Standard for AI transparency statements" (confirming the standard is mandatory for non-corporate Commonwealth entities and took effect 1 September 2024), <https://architecture.digital.gov.au/node/1025>.

¹¹⁰ ASIC, above n 1.

Step 2 — Set the AI policy and accountability structure (Principle 1). The board should *Direct*: approve a company-wide AI policy, decide which board committee (commonly the risk or audit committee, or a dedicated technology committee) has oversight, and assign day-to-day accountability to a named executive. This should be reflected in the board charter, consistent with the direction signaled in ASX’s own 2026 consultation material about AI protocols for directors.¹¹¹

Step 3 — Fold AI into the existing risk management framework (Principle 7). Rather than building a stand-alone “AI risk register”, extend the existing enterprise risk management framework that the board already oversees under Recommendation 7.1, so that AI risk is reported to the board through the same channel as other material risks, and is captured in the entity’s risk disclosures.

Step 4 — Build (or extend) the management system (ISO/IEC 42001, on top of ISO/IEC 27001 where it exists). If the company already holds ISO/IEC 27001 certification, management should extend the existing ISMS using the shared Annex SL structure, adding the AI-specific elements: an AI policy, AI risk and impact assessments, an inventory and classification of AI systems, lifecycle controls, and a Statement of Applicability against the ISO/IEC 42001 Annex A controls.¹¹² If the company has no existing ISO/IEC 27001 system, it should still consider whether it needs one, since most AI risk is downstream of an information-security risk. If the company is also APRA-regulated, this step should be run alongside — not separately from — its existing CPS 234 and CPS 230 compliance programs, since APRA has indicated it expects to see exactly this kind of recognised control framework and control library sitting behind an entity’s AI assurance.¹¹³

Step 5 — Map to Australian government guidance. Because VAISS/the Guidance for AI Adoption and the NIST AI RMF are both explicitly cross-referenced to ISO/IEC 42001, a company that builds its AIMS around ISO/IEC 42001 will, largely as a by-product, be able to demonstrate alignment with the Australian Government’s own guardrails — useful both for regulator engagement and for public disclosure.

¹¹¹ ASX Corporate Governance Council, above n 51.

¹¹² Cybercraft, “AI Governance” (ISO 42001 implementation description), <https://www.cybercraft.net/ai-governance>.

¹¹³ Cybercraft, above n 112.

Step 6 — Decide whether to seek certification. Certification against AS ISO/IEC 42001:2023 (and, if not already held, ISO/IEC 27001) is a commercial decision, not a legal requirement. It may be attractive where customers, regulators or capital providers are asking for independent assurance, or where the company sells AI-enabled products or services to enterprise or government customers who increasingly ask for it in procurement.¹¹⁴ Some Australian and international bodies (such as the UK’s BSI Group, which is both a national standards body and an accredited certification body) can provide this third-party assurance.¹¹⁵ For an APRA-regulated entity, certification is not a substitute for CPS 234 or CPS 230 compliance, but it is strong, independently verified evidence of exactly the “globally recognised control framework” APRA’s April 2026 letter says it wants to see.¹¹⁶

Step 7 — Disclose. Reflect the company’s approach in the corporate governance statement required under Listing Rule 4.10.3, under the relevant Principles (particularly 1, 3 and 7), and consider whether AI governance disclosure should also feature in the annual report’s risk management and operating and financial review sections, consistent with the direction of travel in ASX’s current consultation.

Step 8 — Implement a Cyclical Review Process. Because both the standards and the Australian regulatory landscape are moving quickly (the National AI Plan, the AI Safety Institute, and the possibility of a fifth edition of the ASX Principles are all still unfolding as at the date of this paper), the board should schedule a periodic (at least annually) review of its AI governance arrangements against the current versions of these standards and guidance, rather than treating any single implementation as final.

7. Conclusion

No Australian law currently forces a public company to adopt ISO/IEC 42001, ISO/IEC 38507 or ISO/IEC 27001. But that is precisely why these standards matter so much in practice. In a jurisdiction that has deliberately chosen not to legislate AI-specific duties, and where the corporate regulator has already found a governance gap in a sector it regulates closely, these standards, together with the Australian Government’s own voluntary guidance and the existing

¹¹⁴ Daimon Legal, above n 15.

¹¹⁵ Daimon Legal, above n 15 (noting BSI Group’s UKAS accreditation for ISO 42001 certification).

¹¹⁶ Daimon, Above n 15.

ASX Corporate Governance Principles have become the practical yardstick by which a board's care, diligence and good faith in relation to AI will be judged, whether by a regulator, an activist shareholder, an insurer, or a court¹¹⁷. A director who can point to a board that evaluated, directed and monitored its AI use in the manner ISO/IEC 38507 describes, supported by a certifiable management system built on ISO/IEC 42001 and a secure information base built on ISO/IEC 27001, and disclosed under the existing ASX "if not, why not" framework, is in a materially stronger position than one who cannot.¹¹⁸ For the subset of ASX-listed companies that are also APRA-regulated, this is no longer a purely theoretical yardstick: APRA has already told the market, in plain terms, that it expects to see a "globally recognised control framework" behind AI assurance, and has said it will take stronger supervisory action, including enforcement, against entities that fall short. For those boards in particular, building AI governance around ISO/IEC 38507, ISO/IEC 42001 and ISO/IEC 27001 is less a matter of best practice than of anticipating exactly what their prudential regulator has already said it will be looking for.

¹¹⁷ Noting the obligations imposed upon corporate officers as detailed sections 180 and 181 of the Corporations Act 2001 (Cth) will be highly relevant.

¹¹⁸ This is especially so if the corporate officer is intending to rely upon the Business Judgement rule as set out in section 180(2) of the Corporations Act 2001 (Cth).