KPMG | AI Reimagined

Trusted
Responsible AI

Data, Generative AI & Automated Decisioning

February 2024

kpmg.com.au

"If this technology goes wrong, it can go quite wrong...we want to be vocal about that"

Sam Altman,
CEO, OpenAI (ChatGPT)

How is AI defined differently around the world?

Perhaps the broadest definition is the one posed in the AI Act drafted by the European Union. In Australia, the CSIRO has defined AI and the Department of Industry, Science and Resources has outlined Australia's AI Ethics Principles

Definition of AI

"'Artificial intelligence (AI) may be defined as a collection of interrelated technologies used to solve problems autonomously and perform tasks to achieve defined objectives, in some cases without explicit guidance from a human being." *CSIRO, Department of Industry, Science and Resources*

International definitions

Definition of AI

"'artificial intelligence system' (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with" *Draft EU AI Act*
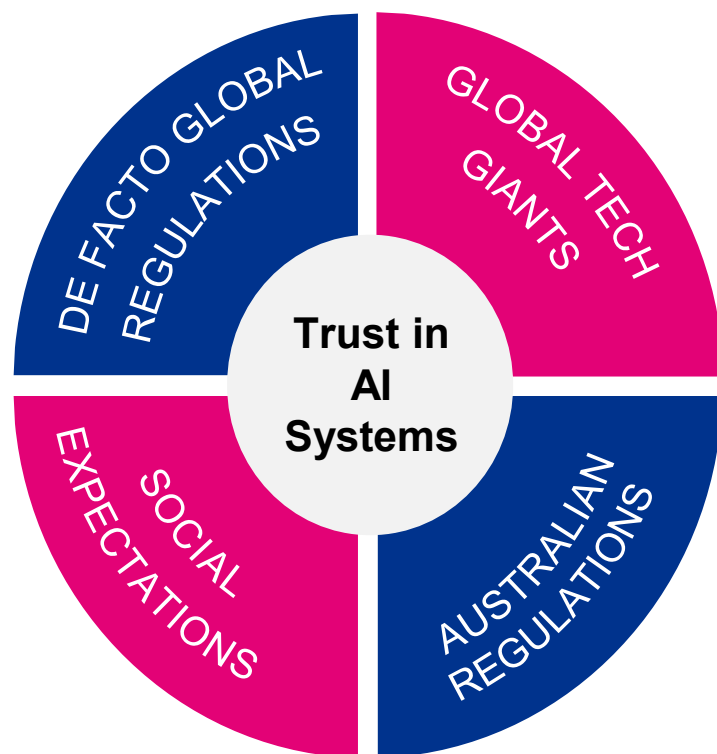
Definition of AI

"'An AI system is a machine-based system that is capable of influencing the environment by producing an output (predictions, recommendations or decisions) for a given set of objectives." *OECD*
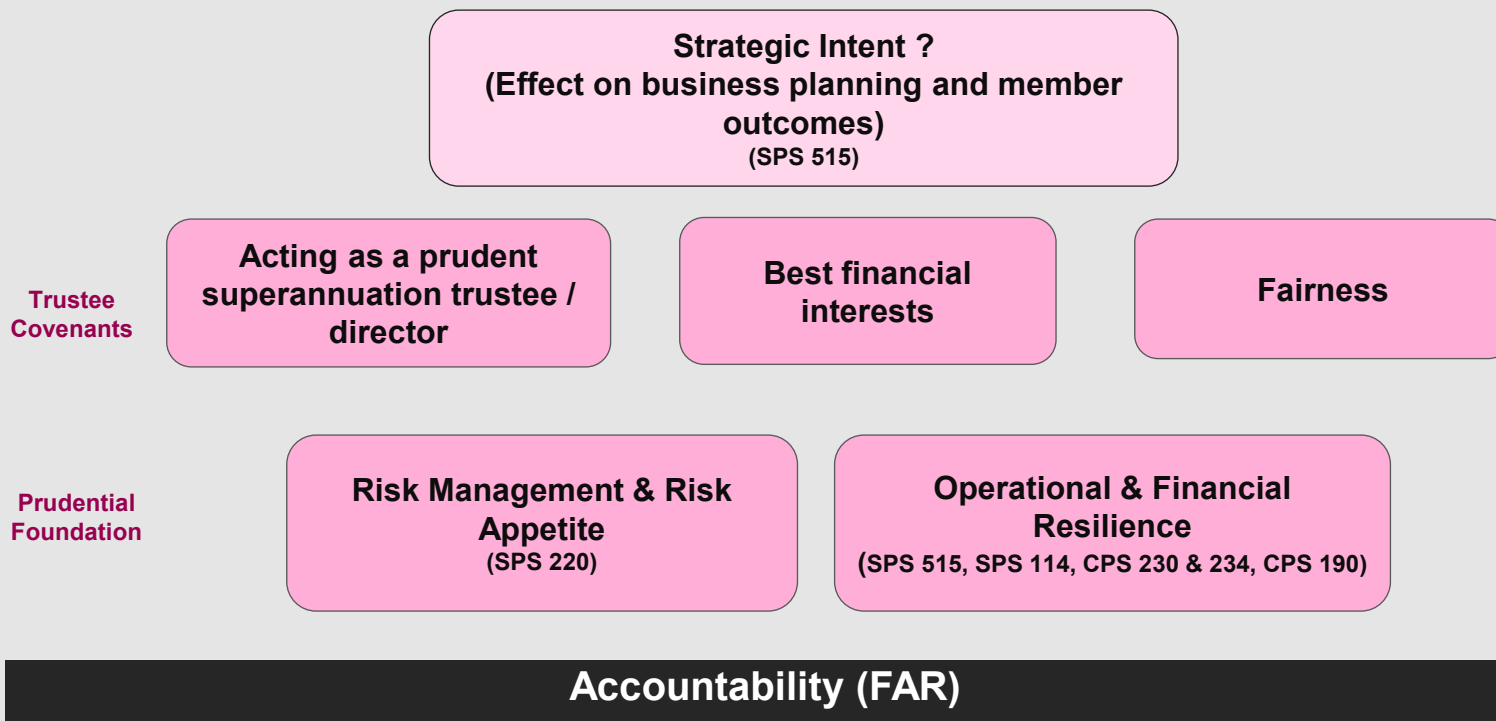
Is there "Trust" in AI Systems? What are the Guiding Principles for Trustees?

Even since 2022 the landscape for the use of AI systems changed significantly as a response to the unprecedented change in technology, the demand for data, and the combination of these two factors informing the use of AI Systems. This has led to calls to "tread carefully".

A multiplicity of actors including global regulators, global tech giants, Australian regulators, consumers and society more broadly, call for greater regulation of AI.

Trust in AI Systems

- DE FACTO GLOBAL REGULATIONS
- GLOBAL TECH GIANTS
- SOCIAL EXPECTATIONS
- AUSTRALIAN REGULATIONS

Against this emerging context what are the guiding principles for trustees?

Strategic Intent ?
(Effect on business planning and member outcomes)
(SPS 515)

Trustee Covenants

- Acting as a prudent superannuation trustee / director
- Best financial interests
- Fairness

Prudential Foundation

- Risk Management & Risk Appetite **(SPS 220)**
- Operational & Financial Resilience **(SPS 515, SPS 114, CPS 230 & 234, CPS 190)**

Accountability (FAR)

While there are many opportunities associated with AI, it also increases the potential for risk

There are a multitude of risks associated with the use of AI systems and algorithms. Many of the risks are unintended and emerge as a consequence of a series of decisions across the lifecycle of the AI systems' development and delivery. Effective mitigation strategies include, but are not limited to, ensuring:

- Adequate **human oversight** through each step of your AI journey;
- Adequate information out how AI **arrived at its output**; and
- Robust governance arrangements, including **policies and procedures** are in place.

Bias & discrimination	Misinformation & misleading information	Cyber Security	Poor Member Experience	Issues with efficacy, reliability & predictability	Legal & Regulatory Risk
Lack of transparency & explainability	Ethical concerns & unintentional consequences	Illegal surveillance	Job displacement & loss	Privacy considerations and potential breaches	Lack of accountability

> " *AI systems should aid in human decision-making, not replace it* "

Responsible AI: How is the regulation of AI systems evolving?

Since 2022, the regulatory landscape for the use of AI systems changed significantly as a response to the rapid development of technology, the demand for data, and the combination of these two factors informing the use of AI Systems. A multiplicity of actors including global regulators, global tech giants, Australian regulators, consumers and society more broadly, call for greater regulation of AI.

De facto global AI regulation

- As the largest economy in the world, the European Union is a global leader in AI regulation. Last year, it implemented the **Digital Services Act (DSA)** and **Digital Market Act (DMA)**, to protect human rights, mitigate disinformation and provide a fair market for consumers.

- This year, the EU also passed an **AI Act**. Expected to take effect in 2025, it represents the most significant change to the way AI systems are used, including banning some forms of AI.

- More recently, the **Biden Administration** signed an Executive Order enforcing rules for the safe use of AI. In November 2023, **Bletchley Declaration on AI Safety** was signed by Australia, the EU and 27 counties, including China, the US and UK.

Social expectations for AI regulation

- In Australia, and 16 other countries, 71% of participants surveyed in a recent global study believed that AI regulation was required.

- The study conducted by the University of Queensland, titled **Trust in AI**, found that 61% of responders were wary about trusting AI systems, and 3 out of 4 people reported that Assurance mechanisms would enhance trust in AI systems.

The changing regulatory landscape for the use of AI Systems:
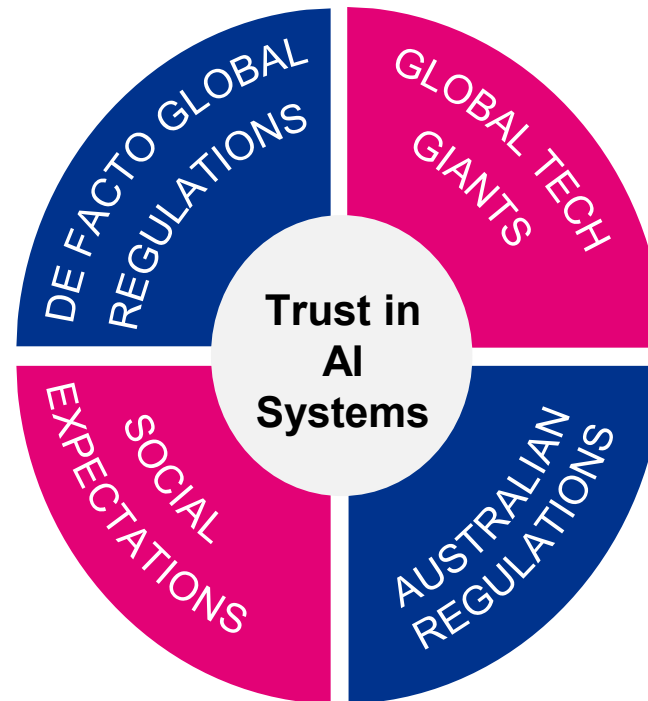
Global tech giants call for AI regulation

- In September 2023, CEOs from global tech giants called for greater AI regulation.

- At Capital Hill, Mark Zuckerberg, Elon Musk, Sundar Pichai, and Bill Gates, amongst others, met with US Senators with a call to regulate AI to protect people from the worst effects of AI.

- Elon Musk described the meeting as a '**service to humanity**', and their call for AI regulation as necessary for the '**future of civilisation**'.

Australian regulators call for AI regulation

- Existing tech agnostic laws apply to the use of technology, including AI systems. This covers the training of data, the models used, the decisions made, the individuals affected and the outputs.

- Last year the ACCC successfully took enforcement action against an online accommodation booking company for using a misleading algorithm (resulting in a $45m fine).

- ASIC has brought 2 enforcement actions against an insurer for using misleading algorithms relating to premium discounts. APRA/ASIC are also looking into bias and discrimination.

- The Australian Industry and Science Minister is currently working on the development of 'modern laws for modern technology' to address the rapid development of AI systems, including considering new laws related to AI, similar to the EU and US.

- The Attorney General is planning to increase rules for surveillance. Australia recently signed the 2023 Bletchley Declaration on AI Safety. On 15 November, Australia commenced *AI Month* for a **national debate regarding a national AI strategy, regulatory framework, and an AI Commission which mirrors similar initiatives in the US and EU.**

Where are you on the AI path?

Not all in the banking, financial services or superannuation sectors are keen adaptors of AI. Warrant Buffett famously stated that he was 'worried' about AI, and that 'we won't be able to uninvent it'. Conversely, there have been many early AI adopters, including in Australia, who have developed AI systems and, with it, started to establish guardrails, including an ethical AI policy and risk management processes, to ensure that AI remains safe, trusted and responsible.

	SKEPTICAL ABOUT AI	JUST GETTING STARTED	PoC FACTORY	CORE AI CAPABILITIES	CRACKED THE AI CODE
Activity in AI	• Have not started using AI yet • Don't see a value in AI • Concerned about AI that is not controlled	• Might have a small number of Proof of Concepts (PoCs). • Identified use cases/needs • Unsure of next steps	• Large number of PoCs • Not achieved ROI • Limited number of use cases in production	• Invested strategically in core business capabilities • Not investing in innovation, but a close follower	• Innovating at scale and deploying into production • Investing in AI start-ups and complementary solutions
Key questions	• Why should I use AI? • How can it be used effectively in my business? • Once 'unleased', how can it be controlled?	• Where should we invest first? • How do we differentiate ourselves? • Is this the right thing for our customers and/or stakeholders?	• Why are PoCs not moving to production? • Do we have the right strategy?	• How do we scale from here? • How do we re-use existing investments? • Are we treating customers fairly?	• How do we ensure our partners adhere to Responsible AI practices? • How do we manage risk in a democratised environment?

AI level

What should Trustees think about to manage AI risk?

In our experience Trustees should consider adopting a structured approach of frameworks and policies...

	RAI Policy	RAI Assurance Framework	RAI Assessment & Management Plan (for each use case)	Establish RAI review panel charter	RAI Training & Awareness
Outcomes					
Description	The RAI Policy compliments the AI strategy by defining the principles guiding the use of AI, and the tolerances for using AI, including the nature of harm	The RAI Assurance Framework outlines the minimum standards and controls required to ensure that AI use adheres to safeguards at an operational level	For each AI use case, the RAI Assessment & Management Plan will document: the AI owner, the business rules, overall objectives, the unintended consequences to be monitored, and the mitigation steps	Establishing a RAI review panel and charter is a key component of ensuring an independent review process for each AI system, especially for medium to high-risk category AI systems	Consisting of two sessions - 1) at the board and executive committee-level: to coach them on their responsibilities for AI systems; and 2) at the AI system owner and user level to build a culture that is sufficiently adept at managing Responsible AI
Benefits	Protects the brand and reputation of the Trustee; Protect members from the worst effects of AI; Ensures the Trustee adheres to legal and regulatory standards	Ensures that AI systems maintain high standards; Introduces controls and processes for developing and maintaining at a use case level	Defines the risk rating for each AI System; Clearly defines how the owner will mitigate the unintended consequences; Provides a feedback loop to the Board	Ensures an independence and oversight process; Provides the ability to challenge the ethical and technical decisions made by the owner of the AI	Ensures the culture is ready for managing AI effectively; Provides further clarity for how actors are engaged and accountable for AI

NOTE: this is not an exhaustive list

How do you prepare your workforce for the future?

DAVOS World Economic Forum 2024

Karim Haji, KPMG Global Head of FS stated:

"**It's been a real eye opener for me to explore the role of AI in bolstering, not replacing, our workforce**. I have spent a lot of time in the last year, focused on our future skills need as a practice and global industry. But this week **I've really been stretched to think about how to build teams differently – based on skills, not roles, job titles or seniority**. We can start mapping the supply and demand of skills. **Some of the most crucial skills over the next decade will be more to do with EQ than IQ**…"

> # *"AI systems should aid in human decision-making, not replace it "*

What should form the basis for human-centred AI design practice?

Be honest job displacement

Start talking openly about how your organization plans to adopt AI, including the questions relating to if it will have the effect of displacing your colleagues

Change the culture

Start changing the culture to one of greater transparency and accountability and where people feel safe to openly discuss and debate ethics and responsible for AI respectfully

Upskill of your colleagues

Support the training and upskilling of your colleagues. The paradox is that your colleagues are going to know more about your business, and at the same time understand what the AI system

Remove the hype

In your business strategy and AI design practice, moderate what AI can do to practical, realistic, limited objectives. This has the benefit of largely serving a more practical purpose, but not overselling the need to displace people

Introduce the notion of harm

AI decisioning does have a number of unintended consequences such as bias, discrimination, dis/mis-information, privacy breaches, lack of transparency/explainability, amongst others. Acknowledge the potential for harm into your AI policy and standards, with the onus on how organisation intends to mitigate that. It sends a message to your colleagues that you give them permission and authority to protect members from harm (which in turn serves to protect the brand)

KPMG AI **Re**imagined

Contact
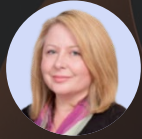
Lisa Butler Beatty
Partner, Consulting
National Lead, Governance & Trust
T: +61 477 752 941
E: lisabbeatty@kpmg.com.au
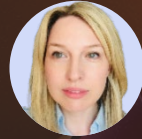
Melinda Rankin
Associate Director, Responsible AI Lead
T: +61 2 9335 7045
E: mrankin1@kpmg.com.au

Danielle Malone
Partner, Consulting AI Lead
T: +61 2 9335 7850
E: daniellem@kpmg.com.au

KPMG.com.au

KPMG | THE UNIVERSITY OF QUEENSLAND AUSTRALIA

UTS | Human Technology Institute

We work with advanced research partners specializing in Responsible AI