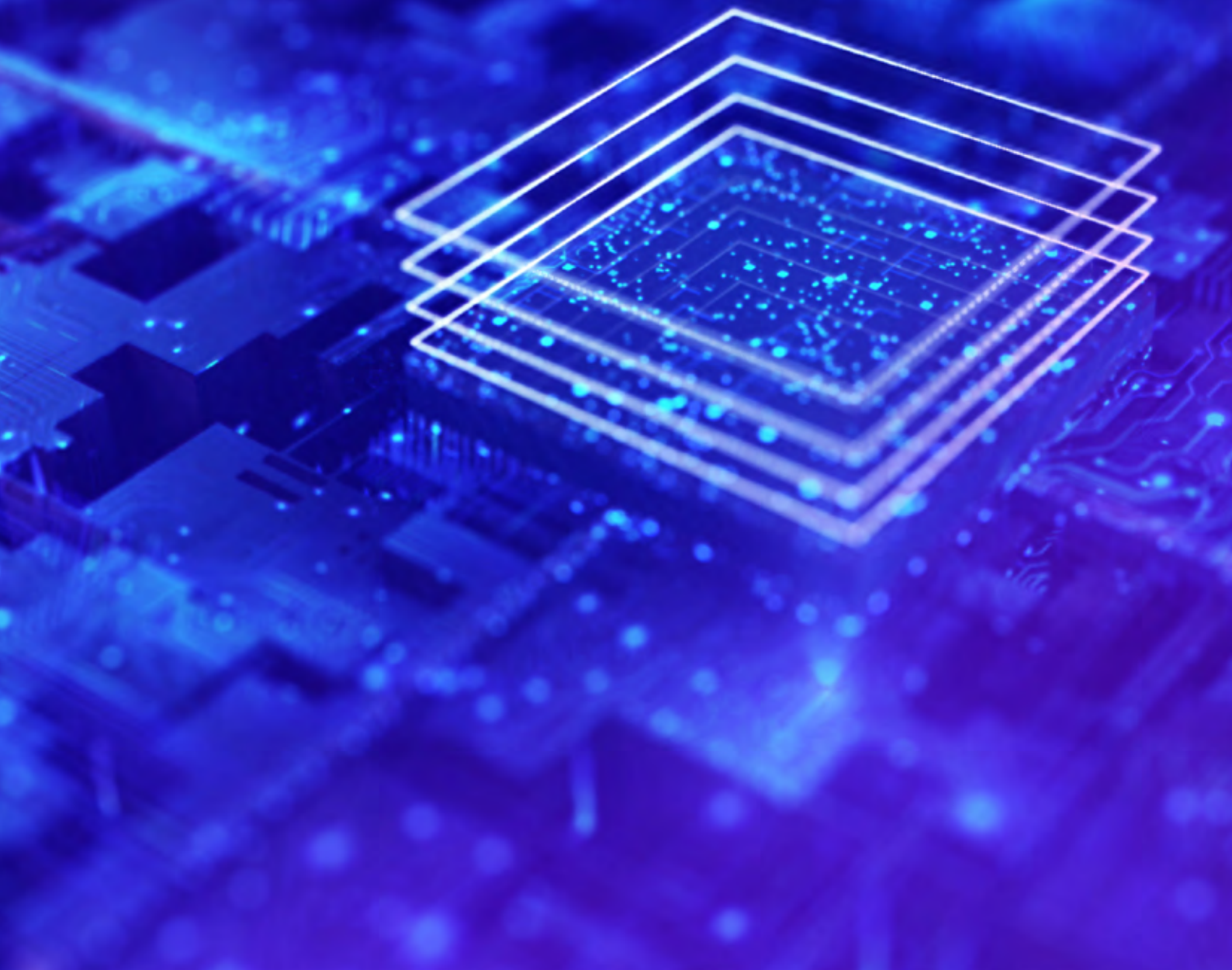




# Navigating AI

**Analysis and guidance on the use and adoption of AI**



## Citation

This report draws on findings and insights from the following report:

Gillespie, N., Lockey, S., Curtis, C., Pool, J., & Akbari, A. (2023). Trust in *Artificial Intelligence: A Global Study*. The University of Queensland and KPMG Australia.  
doi:10.14264/00d3c94

## KPMG advisors

Dean Grandy, Dhruw Joshi, Samantha Kwong, Paul Fitzgerald, Amanda Heather

## Acknowledgements

We are grateful for the insightful input, expertise and feedback on this research provided by James Mabbott, Partner, KPMG

AIIA members: Simon Bush (AIIA CEO), Andrew Hammond (KJR Australia), Guy Simons (UiPath), Lee Hickin (Microsoft), Ray Greenwood (SAS ANZ), Rowena Westphalen (Salesforce)

Professor Nicole Gillespie (University of Queensland)

Michael Vardos, Cara Brugeaud, Callum Hamilton



This report provides a deepened understanding of the AI regulatory landscape globally and within Australia and the need to continue to progress a conversation around appropriate regulation.

Stakeholders from across government and industry contributed to the development of this report, including a selection of AIIA members who were interviewed by KPMG. Their collective thoughts and insights on the AI landscape in Australia, and the role of government and industry in this constantly evolving area are shared throughout this report.

## What's it about?

Working with AIIA, KPMG have prepared this report to help you interpret the current AI landscape in order to meet regulatory requirements and societal expectations.



## Who's it for?

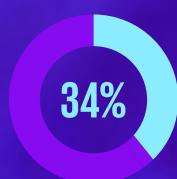
Leaders interested in or tasked with creating policies, governance and oversight of AI technology.



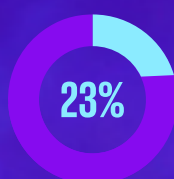
## Key takeaways

1. In the absence of national and authoritative AI regulation and legislation, organisations should self-regulate the design, implementation and running of AI solutions in an informed and transparent manner.
2. Government should be an active enabler and adopter of AI solutions, providing the necessary tools and frameworks to guide the responsible development and application of AI solutions for use in a government context.

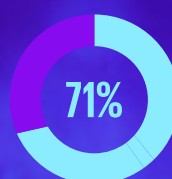
## Why it matters – recent sentiment on AI<sup>1</sup>



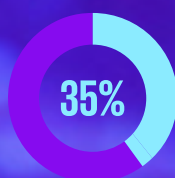
willing to trust



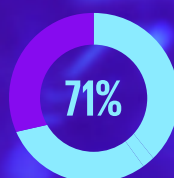
willing to accept



think impact  
is uncertain



agree there are  
enough safeguards



believe regulation  
is necessary

<sup>1</sup> Gillespie, N., Lockett, S., Curtis, C., Pool, J., & Akbari, A. (2023). Trust in Artificial Intelligence: A Global Study. The University of Queensland and KPMG Australia. doi:10.14264/00d3c94

**Right now, there's no authoritative, consistent and overarching definition of AI.**

**It's a fluid and changing space with divergent views, emerging developments and varied user scenarios.**

**Despite this, a consistent frame of reference is required to guide and support AI solution developers, users, regulators, legislators, policymakers and advisers.**

# Contents

## AI – where are we now and where could we go?

– What is AI?	08
– The state of AI in major industries	09
– The view from AIIA members – The state of AI in Australia	11
– What is the value of AI to the Australian economy?	12
– AI and its impact on jobs	13

## AI – the risks and the trust challenge

– Key risks for AI systems	15
– Navigating the trust gap	17
– Do Australians trust AI and should they?	18
– The view from AIIA members – Does government have a role to play in increasing trust in AI?	19

## AI – the regulatory landscape and obligations

– The view from AIIA members – AI regulation	21
– Global legislative and regulatory insights	22
– What does the AI regulatory landscape look like in Australia?	24
– Is self-regulation possible?	26
– The view from AIIA members – The role of industry	27
– Establishing practical AI governance	28

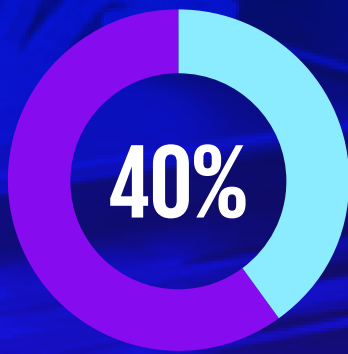
## A checklist for trustworthy AI

– KPMG and UQ’s Trustworthy AI Model and question set	31
– AI – where to from here?	40

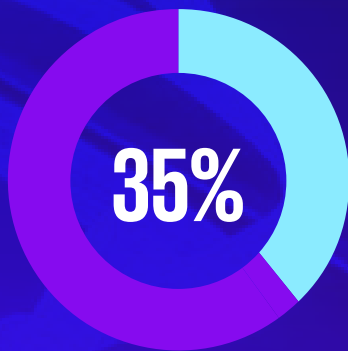
## Appendix

– References	41
--------------	----





**of Australians trust  
the use of AI at work**



**of Australians  
believe there are  
enough safeguards<sup>1</sup>**

## Challenges to successful AI

Leading organisations are addressing AI ethics and governance proactively, rather than waiting for requirements to be enforced upon them by regulators.

The expansion of AI-driven use cases has highlighted both the benefits and the potential risks of AI – notably the issue of trust in technology.

While trust is a defining factor in an organisation's success or failure, the risk of AI goes beyond reputation and customer satisfaction. AI is playing a critical role in shaping the wellbeing and future of individuals and communities around us – even as few fully understand how it works.

Leaders are starting to consider how effective AI governance can help them protect and gain competitive advantage, realise operational efficiencies, and crucially, foster trust among their key stakeholders, including customers.

While regulatory frameworks have been developed to tackle issues related to privacy, there has been little progress towards a more holistic framework that incorporates AI.

A recent University of Queensland and KPMG Australia [study](#) found that only 40 percent of Australians trust the use of artificial intelligence (AI) at work, such as tools like ChatGPT.

Results from the survey also found that only 35 percent of Australians believe there are enough safeguards and current laws or regulations in place to make AI use safe, with no improvement in the adequacy of regulation in the two years since the last survey was conducted. This aligns with previous surveys showing continual public dissatisfaction with the regulation of AI.

Australians expect AI to be regulated with the preferred option indicating the nation needs a dedicated independent regulator to monitor AI usage across a variety of sectors. This highlights the importance of strengthening and communicating the regulatory and legal framework governing AI including data privacy laws.

# AI – where are we now and where could we go?

# What is AI?

**AI is an umbrella term that encompasses a range of interrelated techniques and technologies, including:<sup>2</sup>**

- **Machine learning (ML)** which enables systems to develop models, predictions or insights by ‘learning’ from a training dataset
- **Expert systems** that can learn from and mimic the decision-making abilities of human experts using if–then logical notations
- **Natural language processing (NLP)** which applies computational techniques to the analysis and synthesis of natural language and speech
- **Affective computing** which is a technology to enable systems to identify, process, simulate and respond to human emotions or expressions
- **Computer vision** which enables computers to interpret and understand the visual world
- **Automated decision-making (ADM) systems** that automate the whole or part of a decision-making process.

**As these techniques and technologies become more sophisticated, the market applications of AI are rapidly expanding. For instance, AI is being successfully deployed in areas such as:**

- **Search engines and social media** to provide relevant information and ads to users
- **Logistics** to integrate and optimise freight planning within supply chains
- **Transport and road safety** to build autonomous vehicles and smarter public transport systems
- **Medicine and biotech** for drug discovery and building diagnostic prediction models
- **Smart cities** to monitor energy grid performance and forecast infrastructure maintenance needs
- **Banking, finance and insurance** to detect and prevent fraud as well as to model credit risk
- **Mining and agriculture** to model the environment and automate labour intensive processes.



# The state of AI in major industries

Many industries in Australia and around the world are embracing AI solutions and harvesting business, customer/citizen and economic benefits accordingly.

It is anticipated that broader and deeper acceptance of AI will drive further industry uptake and insight generation.

The sharing of ideation and lessons learned alongside coordinated activity to enable reuse and cross-industry collaboration are key to unlocking greater value and de-risking AI development and adoption.

# The state of AI in major industries

## GOVERNMENT

- Strong uptake of AI across government, especially ADM systems to improve the efficiency of decision-making and the effectiveness of outcomes.
- Automation of tax return processing has reduced the time to refund tax returns to 2 weeks (from 10 weeks manual processing).<sup>3</sup>
- Predictive modelling has integrated environmental data with weather data to help governments anticipate and manage the spread of bushfires.
- AI used in a range of public transport and road safety settings such as identifying drivers using mobile phones, licence plate recognition and to model traffic patterns.
- Applications in law enforcement include predictive policy, surveillance and suspect identification and criminal sentencing.

## MINING

- AI primarily used in mining to optimise processes, enhance decision-making, derive valuable insights from data and improve safety.
- Machine learning supports the discovery of major resource deposits by using training data to identify new targets with similarly high potential for mineralisation or extraction.
- Autonomous mining vehicles are improving mining safety.
- Seismic surveys and modelling can be used to assess the stability of landscapes and reduce risks.
- Sensors can be used to detect and address risks to safety – effectively in real time.<sup>4</sup>

## HEALTHCARE

- AI used to improve the quality of medical research, health service offerings and to support public health.
- AI algorithms helped minimise the spread of COVID-19 with data interpretation to identify hotspots and trace relevant contacts.
- Machine learning promises to transform early detection and diagnosis of diseases, for instance by drawing insights from the trove of data collected by wearables (e.g. smart watches).
- Biotech companies are using AI and machine learning to create efficiencies and reduce costs in the discovery of new drugs.
- AI is being implemented in medical imaging to support physicians to identify and diagnose conditions and promote early interventions.
- Programs have been developed to simulate how healthy cells can be overtaken by viral particles which can help identify the regions of proteins for vaccines to effectively target.<sup>5</sup>

## FINANCIAL SERVICES

- Algorithmic trading systems and platforms are being deployed that can quickly analyse historical data, monitor expert traders' advice and strategies, and automatically execute trades.
- AI can be used to analyse transaction data to uncover fraud trends, detect fraud and automate preventative measures to mitigate customer losses.
- Data analytics provide the opportunity for banks to gain valuable customer insight which can improve financial advice, personalised services and offers, and create efficiencies in lending/loan schemes.
- Banks can use AI to transform the customer experience by enabling always-on service options, such as chatbots.
- AI modelling can optimise insurance offerings and pricings.





# The view from AllA members

## The state of AI in Australia

### Australia – early AI adopters

The opportunity and desire for Australian citizens and businesses to incorporate AI into their daily activities and business operations is rapidly accelerating. Australia is firmly in the early adopter phase of AI, with nearly every new initiative and technology incorporating an element of AI ‘magic’ in it. Concomitantly, there is considerable interest and debate as to what AI is and does, how and where it could and should be used and the risks and benefits that AI can deliver.

### The untapped opportunity

Australia ranks quite highly compared to its APAC neighbours in terms of AI awareness, use and investment. There is evidence of crucial international collaboration and cooperation. However, the reality as a digital and technology enabled nation is that we are underfunded and consequently underutilising our great academic and research resources.

Our challenge is to sustain directional intensity, embrace global and national innovation and advancement in AI solutions and offerings, foster collaborative opportunities for joint industry and government creativity and de-risk the application of AI solutions in an Australian government setting.

### Growing maturity

Many have argued we are over the hype and now in the *trough of disillusionment* with some pockets of maturity. Encouragingly, the tone of the conversations is becoming more pragmatic, industry is maturing and we are moving in a positive direction.

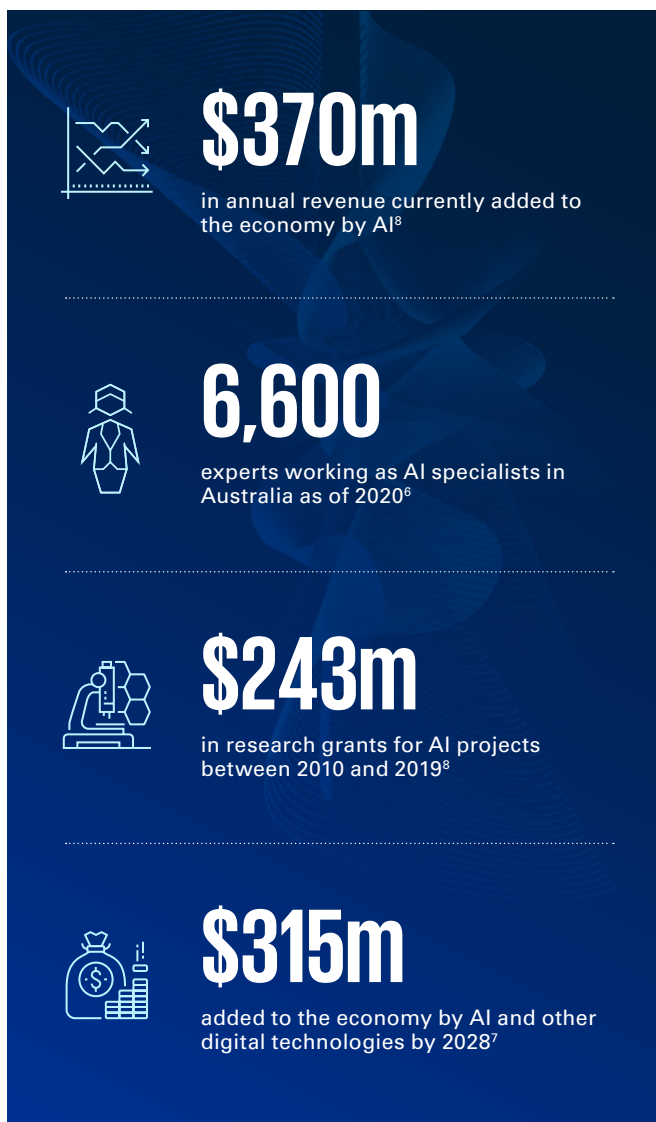
### The common thread

There is concern as to the ethical use of AI and that its application is not detrimental underpins every discussion on this topic.



# What is the value of AI to the Australian economy?

**Given the scale of potential AI-driven economic growth, it is imperative that industry and government investment be optimised and that AI considerations are understood and central to digital economy, digital government planning and delivery decisions.**



Previous thought leadership between KPMG and The American Chamber of Commerce in Australia<sup>8</sup> – *A Prosperous Future: Emerging Tech* – has analysed the opportunities for different Australian industries to contribute to the global AI economy. In terms of AI technology strength, Australia's top industries are:



**Defence**



**Agriculture**



**Financial services**



**Mining**



**Transport**



**Retail**



**Healthcare**



**IT services**



**Logistics**

# AI and its impact on jobs

**AI's potential to drive economic growth will affect the types of work available to people, both positively and negatively. This may also change the demand for different skill sets, knowledge and experience required for professions, industries and society. We are already experiencing the disruptive impact of AI on the job market and the required roles, skills and capabilities for the near and distant future.**



**1.2m**

**new 'technology' jobs in Australia created by 2034<sup>9</sup>**



**5.3m**

**new jobs created by advancements in technology by 2034<sup>10</sup>**



**161k**

**new AI specialists needed by industry before 2030<sup>11</sup>**



**59%**

**of Australians believe AI will eliminate more jobs than it creates<sup>12</sup>**



## ChatGPT – a jobs killer, or a disruptive accelerator?

In recent conversations with government and industry stakeholders, the following themes have emerged:

- curiosity and concern
- an interest to find out more in a government use case context
- uncertainty as to what posture government should take – outright resistance or exploration of opportunity
- appetite for low-risk accelerated learning.

# AI – the risks and the trust challenge



# Key risks for AI systems

While AI promises to deliver enormous economic benefits globally and locally, the emergence and application of AI solutions presents a range of inherent and interrelated risks.

These risks generally relate to data dependencies and the complexity and autonomy of AI solutions which can continue to learn and evolve once deployed in their use case. This often means that AI solutions and their associated risks and impacts are not fully understood at the time they are deployed. Consequentially, it is important for AI developers and users to have strong and responsive governance structures in place to effectively identify and respond to risks and issues across the AI solution lifecycle.

Additionally, the acceptance of risk and the capacity to effectively balance risk and reward in terms of industry reputation, public perception and trust in government and technology will be an essential consideration.

# Key risks for AI systems



## ALGORITHMIC BIAS

Predictions or outputs from an AI system which exhibit an erroneous or unjustified differential treatment between two groups.

- Can be caused by biased training or input data, system design or pre-existing societal inequality.
- Could lead to unfairness or unlawful discrimination.
- Even where sensitive traits are not tracked, AI systems can infer and discriminate by proxy (e.g. gender might correlate with income).



## REAL WORLD COMPLEXITY

Errors can occur if an AI system's limitations are not understood or when it is introduced to uncertainty.

- Data used for training purposes might not reflect real input data.
- Novel situations or rapid changes to the operating environment might confound the AI system.
- Users or stakeholders might interact with the system in unanticipated ways.
- Outputs can be impacted by adversarial attacks.



## INFORMATION PRIVACY

AI challenges established principles of privacy law which presumes sources of personal information are discrete and typically handled by humans.

- Insights gained from big data create a perverse incentive for developers to broadly collect personal information in order to train their models.
- Big data insights can create new personal information about the end user.
- AI enables data to be quickly collected and corroborated in ways that challenge the idea of de-identified data.



## BLACK BOXES

The complexity, autonomy and opacity of AI systems have a 'black box' effect, obscuring how outputs result from inputs and processes.

- Limits stakeholders' ability to make informed decisions about AI systems.
- Risks compounded by automation bias – the tendency to view automated outputs as more objective.
- Gains in model accuracy often come at the cost of model simplicity and explainability.



## LIMITED ACCOUNTABILITY

Governance structures do not support the organisation to manage key risks across the AI system's lifecycle.

- A person responsible for the AI system has not been clearly identified.
- Consideration has not been given to the level of human oversight needed for the system's decision-making.
- Stakeholders do not have ways to contest or provide feedback to outcomes.
- Insufficient documentation and key artefacts for the system across its lifecycle.

As potential risks emerge, a trust gap exists between what AI can do and how users experience it. This gap is exacerbated by two competing narratives – one being that AI will radically and positively change the world, while the other paints a disturbing picture of the damage AI could bring.

# Navigating the trust gap

Leaders across sectors are wrestling with the issue of how to manage AI governance and determine who should take responsibility for AI programs and outcomes. This challenge is occurring amid the widespread adoption of AI technology, which has exposed the risks and intensified the call for prompt AI regulation.

Trust is crucial for AI solutions to be truly transformative. Four key principles – integrity, explainability, fairness, and resilience – are the foundation of trust. These principles can be achieved through proper governance, helping organisations establish transparency, accountability, and trust with their AI solutions.

1. **Integrity** – this means that the algorithms and data used should be valid and appropriate, and their lineage should be traceable.
2. **Explainability** – AI decision-making processes should be transparent, and the reasoning behind them should be easy to understand.
3. **Fairness** – AI systems should be free from bias and prejudice, and any protected attributes should not be used in decision-making.
4. **Resilience** – AI should be technically robust, comply with regulations, and be able to work across different platforms. It should also be resistant to bad actors.



# Do Australians trust AI and should they?

In a recent report<sup>13</sup> produced in partnership with the University of Queensland, KPMG asked Australians how much they trust, accept and support AI in general, as well as specific applications of AI.

Whilst Australians acknowledge the benefits of AI, along with Canada, US, France and the UK we are also the most fearful of it. Significant AI risks identified by people from these countries include cyber security, harmful use, job loss, loss of privacy, system failure, and undermining human rights. Notably, but not surprisingly younger people, specifically Gen Z and Millennials, as well as those with a university education are more likely to trust, accept and support the development of AI. In 2022 more Australians have heard of AI and have more trust in AI than they did in 2020.

Key findings include:

About **3 in 5** indicate they are wary about trusting AI systems, reporting either ambivalence or an unwillingness to trust.

Only **35%** believe there are enough safeguards, laws and regulations in place.

A **third of people** lack confidence in government and commercial organisations to develop, use and regulate AI.

**71%** believe the long-term impact of AI on society is uncertain and unpredictable.

Australians generally accept or tolerate AI, but few approve or embrace it

**> 1,100**

sample group tested

**71%**

believe regulation of AI is necessary

**34%**

trust AI

**23%**

accept AI

# The view from AllA members

## Does government have a role to play in increasing public trust in AI?

### Promoting the opportunity

Government plays a crucial role in building and maintaining public trust. To gain trust, it's important to focus not only on negative aspects but also highlight successes. The government can create awareness about the benefits of AI and its potential impact on society by showcasing examples of how AI is positively impacting key areas such as health, education and environment – as well as lessons learnt when things go wrong. For example, highlighting the success of AI in helping people fill in tax forms.

The other way the government can gain back trust is with transparency – sharing information about actions, decisions and plans for the future.

### The people element

The government also needs to recognise that there are generational differences that impact appetite and sensitivities towards new technologies. It's essential to ensuring there is appropriate representation in government committees and forums so that their concerns, priorities and opinions are heard.

Governments are grappling with opposing tensions. On one hand, as the curators and custodians of people's data, they have an enormous responsibility to protect and ensure the ethical use of that data. On the other hand, citizens want services to be better and more personalised to them. They can work towards balancing these priorities by implementing data protection measures, adopting ethical principles, using data analytics responsibly, and involving citizens in decision-making processes.

### Partnerships with industry and academia

The Australian Government has established the National AI Centre to further develop Australia's AI and digital ecosystem, bringing together partners from government, industry and the research sector to boost exploration and adoption of AI in Australia. This example has a mandate of creating a sense of public trust, awareness and then knowing where and how to go and build on that. We could do so much more. Australians need to be able to trust that AI systems are safe, secure and reliable for us to realise their benefits.

# AI – the regulatory landscape and obligations





# The view from AllA members

## AI regulation

### The ‘for’ views

Government has a key role to play in terms of AI governance and ensuring that benefits are realised and risks are appropriately mitigated. It’s crucial to get the balance right between creating and constraining opportunities for innovation. There is a significant middle ground between complete deregulation and over-regulation where organisations are hesitant to invest and explore AI opportunities for fear of breaching legislative parameters and remaining compliant.

### The challenges – not whether but how

We are at the point now with AI, where there is a need for guidelines, guardrails, and best practice but issues will still arise. Not whether we should use AI, or what sort of AI should be used but how it is used. The key is to learn the lessons and incorporate that into the technology.

Another challenge is that not all AI is the same or has the same implementations. It is hard to see how government regulation could be implemented aside from something vague like: think about it before you do it. Governments are beginning to do what they can with regulation but industry input is crucial here – industry leaders are going to have to lead as best they can, in an ethical way.

Responsible AI isn’t just a regulatory need, it’s good business. In the context of stakeholder capitalism, businesses should consider the interests of all stakeholders. While regulations are crucial to promoting responsible AI, excessive rules may create additional barriers to adoption and hinder the potential benefits that AI can bring to people and their communities.

### The ‘against’ views

The flip side is that AI shouldn’t be regulated. The view is that AI is too amorphous a concept to create rules and regulations. Instead, government should set standards for acceptable outcomes and ensure that the benefits of AI are realised while protecting the fundamental rights of the people it serves.

# Global legislative and regulatory insights

To date, AI regulation has achieved low levels of maturity mainly due to its reliance on voluntary compliance with AI Ethical Principles. Many jurisdictions are following the examples set by the European Union and the OECD in implementing frameworks to develop 'human-centric' AI through self-regulation. Close attention should be given to the AI Act proposed by the EU which, if legislated, could mark a paradigm shift away from laws that address different aspects of AI (e.g. data privacy law) towards comprehensive AI regulations.

## United States

### Legislative maturity:

- The US **has established a National AI Initiative Act**<sup>14</sup> which aims to ensure that American values are integrated into the commercial use of AI.
- This is achieved through investment in AI research and development such as the development of an AI science and technology workforce pipeline.
- The White House released an Executive Order on 9 March 2022 that stipulates a policy for the **Responsible Development of Digital Assets**<sup>15</sup>
- Some US states have begun implementing AI solutions to automate decisions in various industries.

## Canada

### Legislative maturity:

- The Government of Canada has proposed the **Digital Charter Implementation 2022 (Bill C-27)**.<sup>16</sup>
- The Charter seeks to create new rules for the **responsible development and deployment of artificial intelligence (AI)**, as well as regulate other areas of privacy and digitisation.
- The Charter is currently awaiting Senate approval.





## European Union

### Legislative maturity:

The EU's **proposed AI Act**<sup>17</sup> will create a legal framework for trustworthy AI by:

- ensuring that AI ethics and principles are protected.
- improving confidence to embrace AI-based solutions
- encouraging AI development with a risk-based approach to regulation.

The EU is on the cusp of implementing an **AI Liability Directive**,<sup>18</sup> that seeks to:

- lower the risk of harm to vulnerable people

- create an easier process for victims to bring forward civil liability cases
- place the onus on developers.

### Enabling services:

- Proposed **European AI Board** would coordinate the law and oversee a **public database**.
- Developers self-assess conformity or can engage **external auditors**.
- **Market Surveillance Authorities** to enforce law.

## Nordic states

### Legislative maturity:

- The **national AI strategies** of Denmark,<sup>19</sup> Finland,<sup>20</sup> Norway<sup>21</sup> and Sweden<sup>22</sup> have frameworks to guide the development of ethical and trustworthy AI.
- Denmark has **mandatory company legislation for AI and data ethics**,<sup>23</sup> requiring large companies to address data ethics in financial statements.
- A **National Regulation on Automated Decision-making**<sup>24</sup> within public administration is under consideration by Finnish lawmakers.

### Enabling services:

- Norway's **AI regulatory sandbox** is driving innovation in AI development, ethics and business cooperation with regulators.
- The Danish **Data Ethics Council** advises private and public sectors on the ethical use of data in algorithms to achieve socially beneficial solutions.

## Singapore

### Legislative maturity:

- The **Model AI Governance Framework**<sup>25</sup> aims to support the development of ethical AI solutions to promote public understanding and trust in technology.
- The **Implementation and Self Assessment Guide for Organisations**<sup>26</sup> (ISAGO) helps organisations to self-regulate alignment to the Model Framework.

### Enabling services:

- A **Compendium of Use Cases**<sup>27</sup> to demonstrate how organisations have implemented the Model Framework.
- **AI Verify** is an AI Governance testing framework that ensures stakeholders use AI in accordance with the ethical principles in the Model Framework.
- The **Advisory Council on the Ethical Use of AI and Data** advises Government on developing ethical AI.



# What does the AI regulatory landscape look like in Australia?

Australia's AI regulatory landscape is a major obstacle to Australia becoming a leading digital economy and society by 2030. In the absence of specific guidance (policy, legislation and/or consistent frameworks) to regulate the development and use of AI and other similar technologies, inconsistency, apprehension and resistance will continue to challenge broader and accelerated AI adoption.

While existing laws might be applied to impose liability for harms caused by AI solutions, they do not proactively ensure that AI solutions are designed and used safely or help to build trust in responsible innovation.

Given that the existing regulatory landscape was not designed to accommodate emerging AI issues and challenges and is inherently inflexible, regulation will continue to be an impediment to progressing AI development in Australia and specifically in the government domain.

# What does the AI regulatory landscape look like in Australia?

## DATA PRIVACY LAW

### **The Privacy Act 1988**

Liability imposed for interferences with privacy, breach of the *Australian Privacy Principles*, or failure to follow Notifiable Data Breaches scheme.

Regulates some aspects of how personal information (PI) can be collected, used and disclosed by AI systems:

- no data governance requirements for systems using data other than PI
- possibly applies to any insights an AI system has generated about a person
- there is no GDPR equivalent right not to be subject to automated decision-making with legal effects.

### **Australian Privacy Principles (APPs)**

AI systems can process personal information for purposes disclosed in the privacy policy or any other related purpose. Purposes for processing do not need a legitimate interest as is required in some other jurisdictions.

There is no specific requirement to disclose the use of an AI system or explain how it works.

Consent to use of an AI system is only required where it relates to the collection, use or disclosure of sensitive information.

## CIVIL LIABILITY FOR HARM

### **Contractual liability<sup>28</sup>**

Encourages customers to conduct due diligence about how an AI system works and how supplier policies mitigate risks.

Contract terms can be used to:

- limit the scope of liability for uncertain aspects of an AI system's use, such as how intended performance might be impacted by training data or autonomy
- define the scope of an AI system's performance and limitations for the use case it is deployed in
- establish how liability for breach will be allocated
- facilitate monitoring of AI systems, particularly for machine learning algorithms as they 'learn'.

### **Tort of negligence**

- The black box effect makes it difficult to show causation and individual responsibility for harmful outputs.
- Wide range of actors that might contribute to harm, such as developers, programmers, providers, customers, users, and autonomous AI systems.
- Could an AI system's autonomous behaviour be an intervening act that disrupts causation?<sup>29</sup>

### **Discrimination**

- Anti-discrimination rights have been incorporated into federal, state and territory laws, on the basis of protected attributes.
- Algorithmic bias must be protected against in training and implementing an AI system, to prevent the use of flawed datasets which produce decisions that are unreliable or discriminatory.

## CONSUMER PROTECTION AND PRODUCT LIABILITY

### **Australian Consumer Law**

Prohibits users from misleading consumers, including about how their personal information is collected, used or shared, by an AI system.<sup>30</sup>

Requires transparency if the way an AI system delivers outputs or services is influenced by commercial relationships.<sup>31</sup>

Algorithmic bias which causes discrimination in the delivery of products or services might be the basis of unconscionable conduct.

### **Manufacturer's liability**

- Manufacturers will be liable for supplying an AI system with a safety defect where it causes loss or damage to a consumer.
- Developers can limit the risk of action with good practices around AI governance, such as risk management, record-keeping or testing.



# Is self-regulation possible?

**In addition to existing relevant laws, Australia also has a range of ‘soft law’ instruments which are not binding but articulate societal expectations around the development and deployment of AI in ways that can inform best practice. Policy development and law reform are slow to design and implement in a rapidly changing technical environment.**

**To remediate the risk of harm, governments and industry are self-governing AI by producing frameworks to abide by and to self-regulate.**

AI ETHICS FRAMEWORK <sup>32</sup>	ALGORITHMIC BIAS TECHNICAL PAPER <sup>33</sup>	TECHNOLOGY AND HUMAN RIGHTS FINAL REPORT <sup>34</sup>	FACIAL RECOGNITION TECHNOLOGY MODEL LAW <sup>35</sup>	NSW GOVERNMENT AI ASSURANCE FRAMEWORK <sup>36</sup>
<p>Developed by DISR and CSIRO in 2019.</p> <p>Provides a toolkit to help organisations implement Ethical AI Principles including:</p> <ul style="list-style-type: none"> <li>– human, societal and environmental wellbeing</li> <li>– human-centred values</li> <li>– fairness</li> <li>– privacy</li> <li>– reliability</li> <li>– transparency</li> <li>– contestability</li> <li>– accountability.</li> </ul>	<p>Released by Human Rights Commission (AHRC) in 2020.</p> <p>Provides a framework to understand algorithmic bias, including the potential for unfairness and discrimination.</p> <p>Proposes mitigation strategies organisations can take to address algorithmic bias such as:</p> <ul style="list-style-type: none"> <li>– acquire more appropriate data</li> <li>– pre-process the data</li> <li>– increase model complexity</li> <li>– modify the AI system or change the target.</li> </ul>	<p>AHRC policy/regulatory recommendations from 2021, include:</p> <ul style="list-style-type: none"> <li>– an AI Safety Commissioner</li> <li>– human rights impact assessments before automating government decision-making</li> <li>– notice for some ADM systems</li> <li>– establish a human rights approach to AI procurement</li> <li>– anti-discrimination law guidelines for ADM systems.</li> </ul>	<p>Proposed by the Human Technology Institute.</p> <p>Aims to foster responsible innovation while protecting human rights.</p> <p>A risk-based approach to facial recognition tech.</p> <p>Risk ratings include: base-level, elevated and high risk.</p> <p>Imposes requirements, limitations and prohibitions according to the risk rating of a system.</p>	<p>Supports the NSW Government to innovate with AI safely, securely and accountably. Mainly for self-assessment but an AI review body is currently under development.</p> <p>Risk ratings assessed against AI Ethical Principles, advising if the project should:</p> <ul style="list-style-type: none"> <li>– proceed</li> <li>– proceed with additional risk mitigations</li> <li>– stop.</li> </ul> <p>Five mandatory principles:</p> <ol style="list-style-type: none"> <li>1. Community benefit</li> <li>2. Fairness</li> <li>3. Privacy and security</li> <li>4. Transparency</li> <li>5. Accountability</li> </ol>

# The view from Alla members

## The role of industry



It is unlikely that an over-arching set of global AI regulations – a one-size-fits-all approach – will ever be possible nor effective. Regulating AI is as much to do with regulating human values as it is technology, and what may be considered bias or discrimination in one nation might well be the law in another. Getting the right balance between regulation and innovation will be vital for both government and business. Both government and the business community need to come together to create something that is adaptable yet enforceable. This type of cross-society and cross-sector collaboration can also create the building blocks for successful future regulation – whether it be principles of AI adoption or formulating a series of industry standards for AI. The more business is involved in the dialogue with society to shape regulation, the more informed all parties will be.

### **Building trust**

Industry has a role to understand and educate the market. There is an adoption curve, and AI will become less daunting as we move along it. People working in the AI space are going to have to educate those who aren't, to demonstrate the good that can come out of this technology. AI will become widely adopted, every business will need AI on some level, to make their businesses more efficient.

Industry is crucial. Governments are beginning to do what they can with regulation but there is such a wide spread of technologies which means having all the regulations in place will take a long time. Industry leaders are going to have to lead this space as best they can, in a really ethical way.

### **Being transparent**

One of the most interesting conversations in AI over the next few years will be around personal ownership of data. We are starting to see the right to decline to provide information, and companies offering monetary incentives to share data. By promoting transparency and encouraging people to participate, we can empower individuals. When errors happen, it's crucial for organisations to disclose what went wrong – without transparency, people are unlikely to have confidence in these institutions.

Transparency is likely to be the most significant technological factor that industry can prioritise. Ultimately, ensuring data quality is essential, despite the perception that AI can solve data problems.

### **Is self-regulation possible?**

The industry needs to prioritise transparency and governance to ensure convincing outcomes – but self-regulation may not be sufficient as it may not address all necessary measures.

Despite the potential for self-regulation, history has shown that industries driven by profit motives are not successful in regulating themselves. Agreement on regulatory goals is necessary before effective self-regulation can occur – and the EU is leading the way in developing more concrete frameworks to regulate AI. Similar to regulations protecting people from human biases, laws can also protect people from AI biases.

Government and industry have an ongoing responsibility to ensure that AI is designed, developed and implemented to ensure it meets societal expectation and regulatory requirements. This is particularly important as our understanding of AI opportunities, risks, and applications into new fields continues to evolve and grow.



# Establishing practical AI governance

Designing, developing and implementing AI is an ongoing responsibility, to ensure that it is maintained and continually meets societal expectations and regulatory requirements. This is particularly important as our understanding of AI opportunities and risks and application into new fields continues to evolve and grow.

Businesses around the globe find themselves choosing between speed to market with AI-powered solutions and building comprehensive and foundational AI governance capabilities. While being aware of the existential threat that lack of trust in AI poses, organisations find themselves caught in an AI 'space race', whether they are established or new and nimble, companies utilising AI to scale at speed.



# AI system and use case registers

**When building tracking and monitoring mechanisms, consider establishing a register of AI systems and use cases that collects information such as:**

- **A name or label for the system or identified use case**
- **A description of how the system addresses the use case**
- **A person responsible for the system**
- **Project status and any system milestones**
- **Risks or issues and any mitigation actions taken**
- **References to further documentation (e.g. business case, risk assessment, findings, testing, operating logs)**

# Establishing practical AI governance

## 01 | Establishing an AI governance framework

incorporating risk assessments into the system design process and establish clear procedures to escalate risks or issues.

## 02 | Designating a responsible owner for AI governance

in the C-suite and forums to discuss AI governance or any risks or issues associated with AI system.

## 03 | Building mechanisms to track and monitor AI systems and use cases

that are in design, development and that have already been deployed.

## 04 | Responding to appropriately timed reports

for AI systems and facilitating any relevant actions in support of the project teams or AI governance.

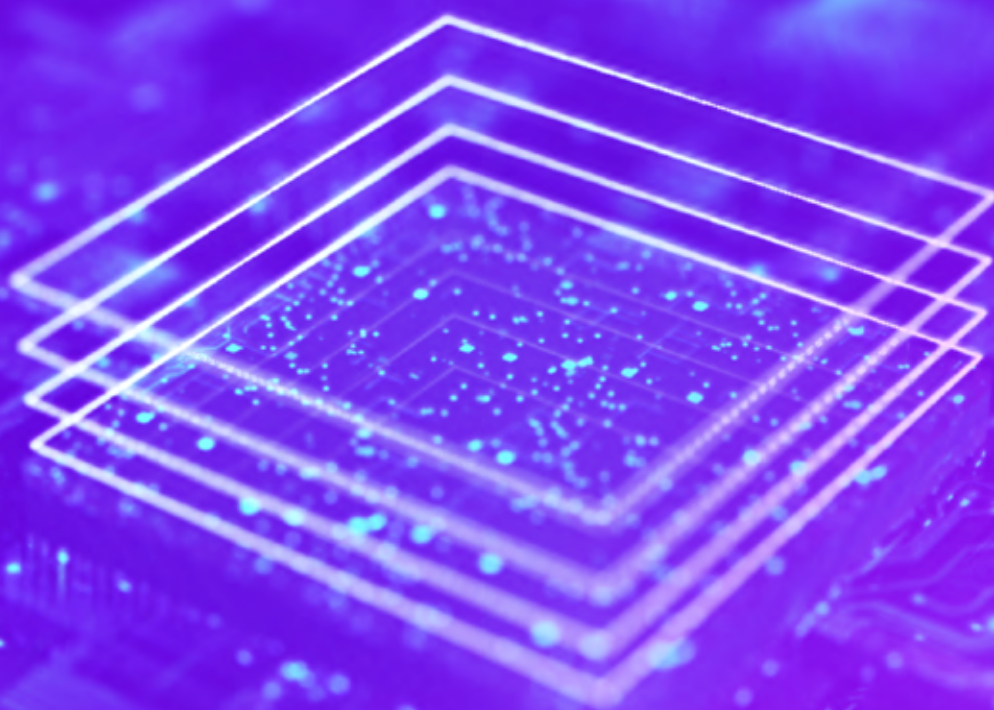
## 05 | Facilitating regular training and knowledge sharing sessions

regarding ethical AI and safe by design principles as well as workshop any risks or issues identified.

## 06 | Implementing routine auditing

of algorithms, involving independent external auditors and a wide range of stakeholders (e.g. system developers, users and end-users).

# A checklist for trustworthy AI

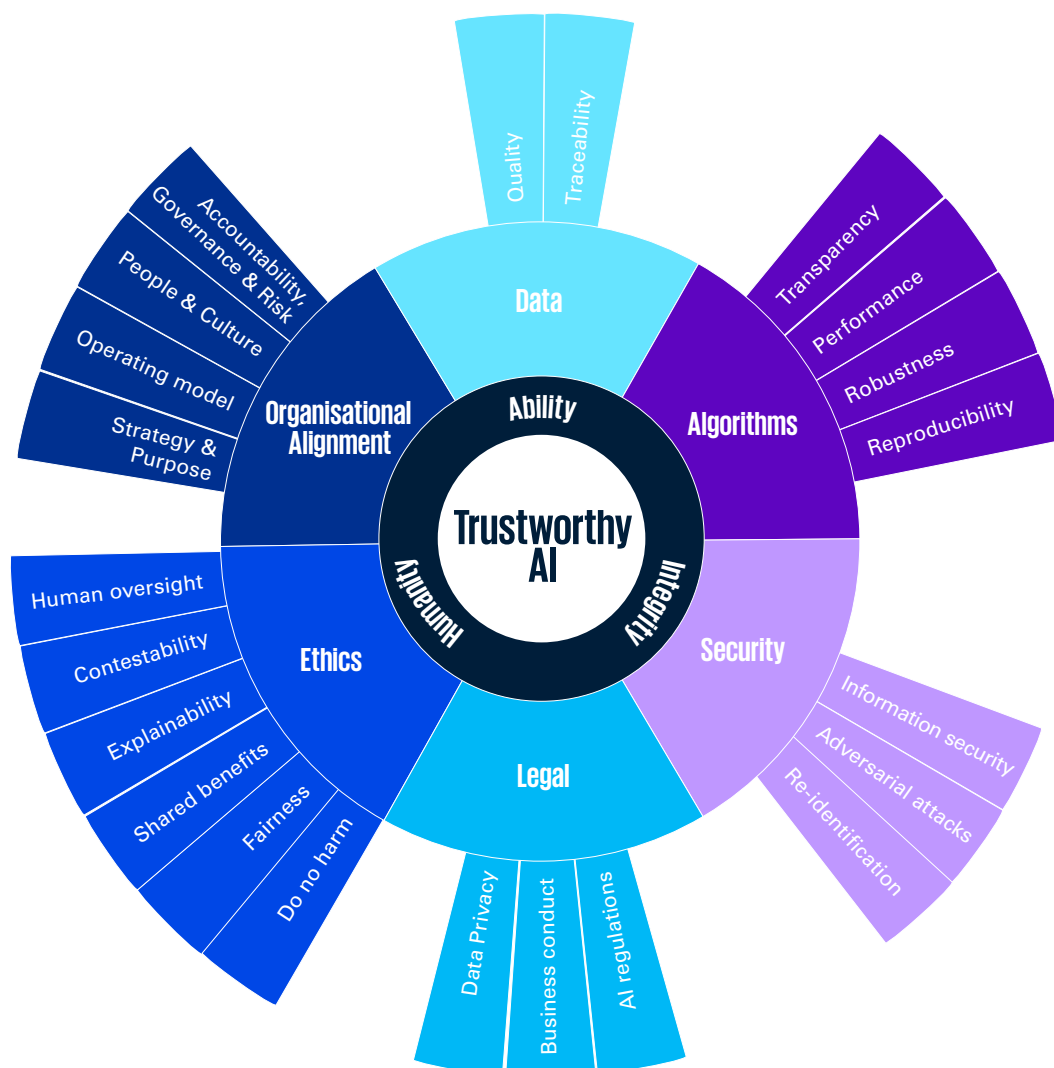


# KPMG and UQ's Achieving Trustworthy AI Model

In 2020 KPMG and the University of Queensland (UQ) created a model for achieving trustworthy AI. The model identifies six dimensions that need to operate in a connected way to ensure Trustworthy AI across the AI lifecycle.

## The Trustworthy AI Model

A model to design, develop, procure, deploy and govern trustworthy data driven systems and their components, including design, data, algorithms and processes.



Gillespie N, Curtis C, Bianchi R, Akbari A and Fentener van Vlissingen R. (2020). [Achieving Trustworthy AI: A Model for Trustworthy Artificial Intelligence](#). KPMG and the University of Queensland Report. doi.org/10.14264/ca0819d



## TRUSTED AI QUESTION SET

# A checklist to provide guidance and use of AI

We have leveraged the proven Trustworthy AI framework to create a checklist to provide guidance around adoption and use of AI within your business, recognising there is no silver bullet for achieving Trustworthy AI.

Organisations need to tailor their approach and ensure any risks presented by the use of AI are proportionate to benefits, including when compared to other possible solutions.

THEME	WHAT DOES GOOD LOOK LIKE?	CHECKLIST
Values and Purpose	AI projects need to <b>align</b> with the <b>organisation's purpose and core values</b> .	<input type="checkbox"/> Does your organisation have a <b>defined and clear purpose</b> in using the identified AI solution (e.g. operational efficiency and/or cost reduction)? <input type="checkbox"/> Has your organisation considered whether the decision to use AI for a specific application/use case is consistent <b>with its core values and/or societal expectations</b> ?
Governance	A governance model (process, people, skills, capability and experience) needs to be able to <b>evaluate whether AI is the right solution</b> for a given problem.	<input type="checkbox"/> Does your organisation have the right governance processes and the skills in place to confirm that <b>AI is the optimal technology solution</b> for a given problem?
Responsible and Trusted	Purpose, risks and benefits of a solution are <b>clearly understood</b> and can <b>inform decisions</b> to implement responsible AI solutions.	<input type="checkbox"/> Has your organisation considered conducting an assessment on whether the <b>expected benefits</b> of implementing the identified AI solution in a responsible manner <b>outweighs the expected costs and risks</b> ? <input type="checkbox"/> Has your organisation considered if there is a <b>less complex alternative than AI</b> for the specific application/use case?

# Data

THEME	WHAT DOES GOOD LOOK LIKE?	CHECKLIST
<b>Traceability</b>	<p><b>Data strategy</b> and <b>governance functions</b> are established as a part of the <b>executive management team</b> (e.g. Chief Data Officer). If not, appropriate mechanisms for functions to engage with the executives and the board are in place.</p> <p>The source and lineage of data within the system is <b>known, documented, traceable</b> and <b>auditable</b>.</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Does your organisation have an <b>enterprise data strategy</b> that aligns to the corporate strategy?</li> <li><input type="checkbox"/> Has your organisation identified a <b>senior leader</b> who is <b>responsible</b> and <b>accountable</b> for <b>data</b> across the organisation and have you defined, catalogued, prioritised, owned, and controlled data assets?</li> <li><input type="checkbox"/> Do you have <b>data stewards</b> who have deep understanding of the data (e.g. purpose, lineage, quality, analysis and reporting) for which they're responsible, and actively manage issues relating to these data assets through to resolution?</li> </ul>
<b>Quality</b>	<p>Processes are in place to <b>measure data quality</b> to <b>avoid unintended consequences</b>. Tolerance levels are established to identify where data quality limits are breached.</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Is your organisation able to verify the <b>accuracy of the dataset</b> in terms of how well the values in the dataset match the true characteristics of the entity described by the dataset?</li> <li><input type="checkbox"/> Has your organisation taken steps to <b>mitigate unintended biases</b> in the dataset used for the AI model, especially omission bias and stereotype bias?</li> <li><input type="checkbox"/> Is the <b>dataset used complete</b> in terms of attributes and items, and fully representative of the actual data or environment the AI solution may function in?</li> <li><input type="checkbox"/> Do you <b>periodically review and update the datasets</b> to ensure accuracy, quality, currency, relevance and reliability?</li> </ul>
<b>Data Risk<sup>1</sup></b>	<p>The organisation's <b>data function</b> is <b>aligned</b> with the <b>risk function</b>. Risk controls and risk tolerances are in place to enable effective monitoring. Risk controls and tolerance are dynamic to respond to change.</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Has your organisation <b>identified and evaluated all risks</b> relating to critical data across the data life cycle and lineage?</li> </ul>
<b>Data Innovation<sup>1</sup></b>	<p>The data function has an <b>eye on the future</b> and measures weak signals for new innovations.</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Does your organisation <b>monitor developments and innovations in data</b> and adopt new practices where relevant?</li> <li><input type="checkbox"/> Does your organisation have a <b>futuristic view</b> to the data by identifying and collecting important data that might be required in the future?</li> </ul>

<sup>1</sup> Themes not included in original Trusted AI Model, however complement the original themes.

# Algorithms

THEME	WHAT DOES GOOD LOOK LIKE?	CHECKLIST
Transparency	Algorithms and outputs can be <b>explained to humans</b> with consideration of varying levels of understanding.	<ul style="list-style-type: none"> <li><input type="checkbox"/> Has your organisation <b>documented the technical features of the algorithm</b> and its design to enable understanding of how the end-to-end process works, and how it arrives at its outcomes?</li> <li><input type="checkbox"/> Do you have a clear <b>explanation of the algorithm's behaviour</b> and the logic behind its design?</li> <li><input type="checkbox"/> Are the stakeholders clearly informed of <b>what data is being collected</b> from them and what processes are automated?</li> <li><input type="checkbox"/> Are <b>performance results accessible</b> to stakeholders?</li> </ul>
Performance	Design (including documentation of design outcomes) needs to <b>ensure that responsible and ethical outcomes are achieved</b> .	<ul style="list-style-type: none"> <li><input type="checkbox"/> Does your organisation perform <b>specialised tests</b> to ensure outcomes are free from <b>unfair bias</b>?</li> <li><input type="checkbox"/> Are datasets reflective of the data in production?</li> <li><input type="checkbox"/> Does your organisation set <b>appropriate performance targets</b> based on the sensitivity and use of the AI solution?</li> <li><input type="checkbox"/> Prior to deployment and use, has your organisation set <b>effective performance metrics</b> to ensure targets are achieved?</li> </ul>
Robustness	<b>Continuous feedback from the outcomes of the algorithms</b> and warning signals if some of the base parameters change and the solution delivers unintended results.	<ul style="list-style-type: none"> <li><input type="checkbox"/> Does your organisation <b>regularly test outcomes</b> and processes to ensure that the same performance that was established and confirmed during AI solution development is upheld, despite possible changes in the environment that might occur during AI solution operations?</li> <li><input type="checkbox"/> What ongoing mechanisms are in place to continue <b>validating the algorithm does no harm</b>?</li> </ul>
Reproducibility	Reproducibility in machine learning means that you can repeatedly run your algorithm on certain datasets and <b>obtain the same</b> (or similar) <b>results</b> on a particular project.	<ul style="list-style-type: none"> <li><input type="checkbox"/> Does your organisation have a process in place to <b>track the collection of new data</b> and/or <b>results</b> to ensure an appropriate level of quality control?</li> <li><input type="checkbox"/> For systems that produce similar results, but not exact results from testing, does your organisation have a <b>risk management process</b> in place to monitor additional deviations?</li> <li><input type="checkbox"/> Does your organisation have the <b>appropriate governance mechanism</b> to discuss the risk of reproducibility for AI solutions?</li> </ul>



# Security

THEME	WHAT DOES GOOD LOOK LIKE?	CHECKLIST
<b>Information Security</b>	Robust and clear <b>information security</b> and access protocols are in place to ensure the <b>confidentiality</b> , integrity, access and availability of data is protected throughout the data and AI life cycle.	<ul style="list-style-type: none"> <li><input type="checkbox"/> Where data is provided by a <b>third party</b>, does your organisation assess and <b>manage</b> the <b>risks</b> of using these <b>datasets</b>?</li> <li><input type="checkbox"/> Has your organisation <b>determined and documented</b> the locations, confidentiality, integrity and availability requirements of the <b>systems and information</b>?</li> <li><input type="checkbox"/> Has your organisation completed appropriate <b>back-ups</b> on the data on a regular and proven basis?</li> </ul>
<b>Adversarial Attacks</b>	Robust <b>cyber security measures</b> are in place to identify and prevent <b>adversarial machine learning attacks</b> , hacking and other types of cyber attacks that may compromise the performance of the AI solution, breach human and legal rights, and result in unfair outcomes.	<ul style="list-style-type: none"> <li><input type="checkbox"/> Does your organisation take controls to <b>minimise</b> the chance of <b>unauthorised access</b> to training data?</li> <li><input type="checkbox"/> Has your organisation considered the potential <b>adversarial vulnerabilities during the design stage</b> and thought of solutions to prevent this when developing the algorithms?</li> <li><input type="checkbox"/> Does your organisation have the resources to constantly monitor and periodically reassess algorithms?</li> <li><input type="checkbox"/> Have you developed a <b>threat model</b> for your AI solution?</li> <li><input type="checkbox"/> Is the <b>solution proofed</b> against evasion, poisoning and inference attacks?</li> </ul>
<b>Re-identification</b>	The risk of <b>malicious actors</b> re-identifying individuals by combining anonymised data with other sources is <b>effectively identified and managed</b> .	<ul style="list-style-type: none"> <li><input type="checkbox"/> Has your organisation <b>utilised anonymised or model datasets</b> where practical, and utilised personal data only where absolutely necessary?</li> <li><input type="checkbox"/> Do your <b>data-handling practices</b> consider where data deals with <b>personal identifiable information</b>, and the risk profile and remediation if datasets were combined?</li> </ul>

# Legal

THEME	WHAT DOES GOOD LOOK LIKE?	CHECKLIST
<b>AI Regulations</b>	Local and global, soft and hard regulations and <b>legislative frameworks</b> relating to data and AI are <b>understood and consistently adhered</b> to across the organisation. Changes are dynamically monitored.	<ul style="list-style-type: none"> <li><input type="checkbox"/> Have you got <b>documentation of the solution</b> (e.g. what data is being used, what problem the solution is trying to address, and what are the solution outputs?) that enables <b>constant monitoring</b> of changes to the solution and any impacts of changing legislation?</li> <li><input type="checkbox"/> Does your organisation have a function in place to <b>dynamically monitor new or changing risks</b> and compliance with Australian law (e.g. AI governance committee, dedicated legal assurance function, auditability)?</li> <li><input type="checkbox"/> Does your organisation have <b>consistent ethical principles</b> you abide by? Is there alignment with government published materials (e.g. Australia's Artificial Intelligence Ethics Framework or the NSW Government AI Assurance Framework)?</li> <li><input type="checkbox"/> Has your organisation <b>assessed the AI solution</b> against other <b>policy makers</b> (e.g. Standards Australia, or the Human Rights Commission)?</li> </ul>
<b>Data Privacy</b>	Privacy impact assessments and procedures are in place to ensure <b>legal compliance</b> and stakeholders' <b>ethical privacy expectations</b> are met.	<ul style="list-style-type: none"> <li><input type="checkbox"/> Has your organisation identified why your solution requires <b>specific access to data</b> and can you reduce the amount of data that you require access to?</li> <li><input type="checkbox"/> Do you have a privacy impact assessment mechanism you can apply to AI solutions?</li> <li><input type="checkbox"/> Do you have the appropriate <b>data governance controls</b> in place?</li> <li><input type="checkbox"/> Do you understand the <b>possible vulnerabilities</b> of the data you are using (e.g. individual data) and is there user-sensitive personal information?</li> <li><input type="checkbox"/> Is there a way to dynamically <b>monitor the insights and personal information</b> being generated by the solution about the individual over the life cycle of the solution?</li> </ul>
<b>Business Conduct</b>	Business conduct regulations are proactively identified to ensure <b>AI solutions are compliant</b> .	<ul style="list-style-type: none"> <li><input type="checkbox"/> How is AI governance factored into your <b>corporate governance</b>?</li> <li><input type="checkbox"/> Do you have a <b>dedicated responsible owner for AI</b> solutions in your organisation?</li> <li><input type="checkbox"/> Are there mechanisms for people in the business and from the public to <b>question AI risk</b>?</li> <li><input type="checkbox"/> How will your use of <b>AI be impacted by existing conduct</b> of business obligations?</li> </ul>

# Ethics

THEME	WHAT DOES GOOD LOOK LIKE?	CHECKLIST
<b>Do No Harm</b>	The risks, unintended <b>consequences</b> and potential for <b>harm</b> of an AI solution are fully <b>assessed and mitigated</b> prior to, and during, its deployment. Particular care is given to human rights and vulnerable stakeholders.	<ul style="list-style-type: none"> <li><input type="checkbox"/> Has your organisation conducted a <b>human rights assessment</b> on the AI solution?</li> <li><input type="checkbox"/> Are <b>safe-by-design principles</b> followed? Are vulnerable groups accounted for, and protected in the design of the solution?</li> <li><input type="checkbox"/> Has your organisation mapped the possibility of the solution producing <b>harmful outputs</b> and are the appropriate <b>risk mitigations</b> in place?</li> <li><input type="checkbox"/> Can your organisation justify that there is a <b>positive value</b> or impact produced by the AI solution?</li> </ul>
<b>Fairness</b>	The <b>outcomes</b> of AI solutions are <b>regularly monitored</b> to ensure that they are fair, free of unfair bias and discrimination, and designed to be inclusive of diverse stakeholders.	<ul style="list-style-type: none"> <li><input type="checkbox"/> Has your organisation taken the relevant steps to <b>mitigate unintended bias</b> in the datasets used in the model? <b>See: Design and Test Principles &amp; Performance (Algorithms)</b></li> <li><input type="checkbox"/> Has your organisation considered how <b>vulnerable groups</b> could be impacted by the solution?</li> <li><input type="checkbox"/> Is your organisation's AI solution compliant with <b>Australian anti-discrimination laws</b>?</li> </ul>
<b>Shared Benefits</b>	The AI solution is designed to benefit a range of stakeholders, including customers, employees and end users. Organisations are held accountable to a benefits management and realisation plan to measure the <b>benefits of the AI solution</b> .	<ul style="list-style-type: none"> <li><input type="checkbox"/> Does the use of <b>individual data deliver an improvement</b> in the quality of the solution and quality of services?</li> <li><input type="checkbox"/> Has your organisation conducted <b>user testing</b> to ensure that the output from the solution would be of <b>benefit to consumers</b>?</li> <li><input type="checkbox"/> Is the implementation of an AI solution <b>human-centric</b> and have employees been involved and supported to understand that the solution will <b>supplement their capabilities</b> and not replace them?</li> </ul>
<b>Explainability</b>	The purpose of the AI solution, how it functions and how data is used and managed is <b>transparently explained</b> and <b>understandable to a variety of stakeholders</b> .	<ul style="list-style-type: none"> <li><input type="checkbox"/> Are your organisation's decision-makers able to sufficiently <b>explain the solution</b> to a range of different stakeholders at appropriate levels of detail?</li> <li><input type="checkbox"/> Are there resources available for employees and for customers to <b>understand each component of the AI solution</b> and how a decision has been made?</li> <li><input type="checkbox"/> If explainability cannot be achieved to an adequate level, are there alternative methods to <b>explain that the solution is functioning</b> in the way it was supposed to (i.e. repeatability testing)?</li> <li><input type="checkbox"/> Can the AI solution and its risks be <b>explained</b> and easily understood when sold by a vendor?</li> <li><input type="checkbox"/> Are the appropriate resources provided to ensure that solutions and <b>systems are not mishandled</b>?</li> </ul>



# Ethics (continued)

THEME	WHAT DOES GOOD LOOK LIKE?	CHECKLIST
<b>Contestability</b>	Any impacted user or stakeholder is able to <b>challenge the outcomes of an AI</b> solution via a fair and accessible human review process, with clear mechanisms for remediation where appropriate.	<ul style="list-style-type: none"> <li><input type="checkbox"/> Do users know where to go to make a <b>complaint or contest a decision</b>?</li> <li><input type="checkbox"/> Is there a <b>fair mechanism</b> in place for users to be able to <b>contest a decision</b> with (e.g. with human oversight)?</li> <li><input type="checkbox"/> Is there a <b>function to escalate</b> where decisions have been overturned to ensure that the AI solution doesn't continue to <b>learn from an incorrect decision</b>?</li> <li><input type="checkbox"/> Are users able to provide <b>feedback</b> to the solution/ its <b>administrator to improve the quality</b> of its outputs?</li> <li><input type="checkbox"/> Is there an alternate <b>manual process</b> if use of automated process is being contested?</li> </ul>
<b>Human Oversight</b>	There is appropriate <b>human oversight</b> and control of AI solutions and their impact on stakeholders by people with <b>sufficient knowledge and AI literacy</b> to ensure informed engagement, decision-making and risk management. Standard operating procedures and frameworks are defined and documented to support human oversight.	<ul style="list-style-type: none"> <li><input type="checkbox"/> What degree of <b>human oversight</b> is in place (e.g. when is the decision made: human in-the-loop, human out-of-the-loop, and human over-the-loop)?</li> <li><input type="checkbox"/> What contributed to the <b>decision to utilise the above</b> approach to human oversight?</li> <li><input type="checkbox"/> Has your organisation developed a method to <b>monitor and mitigate risk</b> from the solution (i.e. changes in how systems operate, and the impact of external changes on the solution)?</li> </ul>

# Organisational Alignment

THEME	WHAT DOES GOOD LOOK LIKE?	CHECKLIST
<b>Strategy and Purpose</b>	The purpose, design and use of AI systems <b>align with the organisation's strategy</b> , purpose and values, and are <b>designed to engender trust</b> .	<ul style="list-style-type: none"> <li><input type="checkbox"/> Has your organisation defined a <b>clear purpose</b> in using the identified AI solution (e.g. operational efficiency and cost reduction)?</li> <li><input type="checkbox"/> Has your organisation established a set of <b>ethical principles</b> that is in line with or can be incorporated into your organisation's vision and mission statement?</li> <li><input type="checkbox"/> Is the use of AI for a specific application/use case consistent with your <b>core values and/or societal expectations</b>?</li> </ul>
<b>Operating Model</b>	Resourcing, processes, policies and operational systems are <b>developed and updated</b> to execute the organisation's <b>AI strategy</b> .	<ul style="list-style-type: none"> <li><input type="checkbox"/> Does your organisation have an <b>existing governance structure</b> that can be leveraged to <b>oversee the use of AI</b>?</li> <li><input type="checkbox"/> Has your organisation considered specific requirements or adjustments to ensure the <b>governance structure is fit for purpose</b> (e.g. legal, ethical and commercial considerations? Centralised vs decentralised decision-making)?</li> <li><input type="checkbox"/> Is the governance structure in place appropriately resourced with <b>clarity on responsibilities and approach for how an AI project would be designed and delivered</b>?</li> </ul>
<b>People and Culture</b>	The right people, capabilities, knowledge and diversity, and cultural practices are in place to <b>achieve trustworthy AI</b> .	<ul style="list-style-type: none"> <li><input type="checkbox"/> Is everyone involved in decision-making, design, or use of the AI solution <b>aware of their responsibilities</b>, appropriately trained, and equipped with the right tools and resources to perform their role?</li> <li><input type="checkbox"/> Are the people dealing with the AI solution <b>properly trained</b> to understand the model output and decisions, and detect and manage bias in data?</li> </ul>
<b>Accountability Governance and Risk</b>	The chain of <b>accountability and responsibility</b> for the AI solution (including governance of data and algorithms) across key stages of its life cycle are <b>clearly defined</b> and understood.	<ul style="list-style-type: none"> <li><input type="checkbox"/> Do those with accountability for the AI solution have <b>clarity on their responsibilities and controls</b> throughout design, development and deployment? Is this appropriately documented and available to all staff?</li> <li><input type="checkbox"/> Does your organisation evaluate whether your accountability, governance structures, and risk management processes are <b>in line with local and international developments</b>?</li> </ul>

# AI – where to from here?

How do we create a future where people reap the benefits of AI while trusting in the organisations who build and regulate it?

Currently, one-third of individuals lack trust in governments, technology, and commercial organisations to responsibly develop and use AI.

However, through implementing mechanisms such as regular accuracy and reliability monitoring, AI codes of conduct, independent ethics reviews, and adhering to international standards, organisations can earn and maintain trust while demonstrating their commitment to responsible AI use.



# Appendix: References

1. Gillespie N, Lockey S, Curtis C, Pool J, Akbari A. (2023). Trust in Artificial Intelligence: A Global Study. The University of Queensland and KPMG Australia. doi:10.14264/00d3c94
2. Office of the Victorian Information Commissioner. (August 2018) [Artificial Intelligence and Privacy: Issues and Challenges](#), OVIC.
3. Law Council of Australia (June, 2022) [2022 06 03 – S – Automated Decision Making and AI Regulation Issues with attachments.pdf \(lawcouncil.asn.au\)](#), Digital Technology Taskforce, Department of the Prime Minister and Cabinet.
4. Ghosh R (2022) [Applications, Promises and Challenges of Artificial Intelligence in Mining Industry: A Review \(techrxiv.org\)](#) TechRxiv.
5. Hajkowicz SA, Karimi S, Wark T, Chen C, Evans M, Rens N, Dawson D, Charlton A, Brennan T, Moffatt C, Srikanth S, Tong KJ (2019) Artificial intelligence: Solving problems, growing the economy and improving our quality of life. CSIRO Data61, Australia.
6. Department of Industry, Science and Technology. (June 2021) [Australia's AI Action Plan](#), DISR.
7. Hajkowicz SA, et al. Artificial intelligence (n 5)
8. KPMG and AmCham. (July 2022) [A Prosperous Future: Emerging Tech – Opportunities for Australia-US Trade in Digital Economy, Artificial Intelligence, and Quantum Science](#), KPMG and AmCham.
9. Faethm. (2020) Technology Impacts on the Australian Workforce, Australian Computer Society; cited by Department of Industry, Science and Technology. (June 2021) Australia's AI [Action Plan](#), DISR.
10. Ibid.
11. Hajkowicz SA, et al. Artificial intelligence (n 5)
12. Lockey S, Gillespie N and Curtis C. (October 2020) [Trust in Artificial Intelligence: Australian insights 2020](#), The University of Queensland and KPMG Australia, doi:10.14264/b32f129
13. Gillespie N, et al. Trust in Artificial Intelligence (n 1)
14. [National AI Initiative Act of 2020](#) (DIVISION E, SEC. 5001).
15. [Executive Order on Ensuring Responsible Development of Digital Assets](#) (9 March 2022).
16. Bill C-27 (44-1) [An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts](#). (22 November 2021, to present) Sponsored by the Hon. François Philippe Champagne, Minister of Innovation, Science and Industry.
17. European Commission, [Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence and Amending Certain Union Legislative Acts](#) (Doc No 2021/0106 (COD), 21 April 2021).
18. European Commission, [Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence \(AI Liability Directive\)](#) (Doc No 2022/0303 (COD), 28 September 2022).
19. Danish Ministry of Finance and Ministry of Industry, Business and Financial Affairs. (March 2019) [National Strategy for Artificial Intelligence](#).
20. Ministry of Economic Affairs and Employment of Finland. (2020) [Artificial Intelligence 4.0 First interim report: From start-up to implementation](#).
21. Norwegian Ministry of Local Government and Modernisation. (January 2020) [National Strategy for Artificial Intelligence](#).
22. Government Offices of Sweden. (2018) [National approach to artificial intelligence](#).
23. Larsen F W. (June 15 2020). [Denmark: an independent council and a labelling scheme to promote the ethical use of data](#). OECD AI Policy Observatory.
24. Finnish Ministry of Justice and Ministry of Finance. (2020) [National Regulation on Automated Decision-Making](#).
25. Singapore Digital, Info-communications Media Development Authority and Personal Data Protection Commission. (January 2020, 2nd ed.) [Model Artificial Intelligence Governance Framework](#), PDPC.
26. World Economic Forum, Singapore Digital, Info-communications Media Development Authority and Personal Data Protection Commission. (January 2020) [Companion to the Model AI Governance Framework – Implementation and Self-Assessment Guide for Organisations](#), WEF.
27. Singapore Digital, Info-communications Media Development Authority and Personal Data Protection Commission. (2020) Compendium of Use Cases: Practical Illustrations of the Model AI Governance Framework ([Volume 1](#)); ([Volume 2](#)), PDPC.
28. Kelly P, Walsh M, Wyzkiewicz S and Alls S. (October 2021) [Artificial Intelligence and the Law: Practical measures to mitigate legal risk](#), DLA Piper.
29. Chesterman, S (2021) [We, The Robots: Regulating Artificial Intelligence and the Limits of Law](#), Cambridge University Press, p. 90-91, ISBN: 9781316517680
30. Cantatore F and Marshall B. (2021) [Safeguarding Consumer Rights in a Technology Driven Marketplace](#), Adelaide Law Review 46(2) p. 468-500.
31. ACCC. (8 October 2020) [iSelect to pay \\$8.5 million for misleading consumers comparing energy plans](#).
32. CSIRO Data61 and DISR. (November 2020), [Australia's AI Ethics Principles](#).
33. Australian Human Rights Commission. (2020) [Using artificial intelligence to make decisions: Addressing the problem of algorithmic bias](#).
34. Australian Human Rights Commission. (2021) [Human Rights and Technology Final Report](#).
35. Nicholas Davis, Lauren Perry and Professor Edward Santow. (September, 2022) [Facial Recognition Technology – Towards a model law](#). Human Technology Institute.
36. NSW Government. (March 2022) [NSW Artificial Intelligence Assurance Framework](#).



**Dean Grandy**  
**Lead Partner, Technology –  
Infrastructure, Government &  
Healthcare**  
KPMG Australia



**Simon Bush**  
**Chief Executive Officer**  
Australian Information  
Industry Association

**KPMG.com.au**

The information contained in this document is of a general nature and is not intended to address the objectives, financial situation or needs of any particular individual or entity. It is provided for information purposes only and does not constitute, nor should it be regarded in any manner whatsoever, as advice and is not intended to influence a person in making a decision, including, if applicable, in relation to any financial product or an interest in a financial product. Although we endeavour to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation.

To the extent permissible by law, KPMG and its associated entities shall not be liable for any errors, omissions, defects or misrepresentations in the information or for any loss or damage suffered by persons who use or rely on such information (including for reasons of negligence, negligent misstatement or otherwise).

©2023 KPMG, an Australian partnership and a member firm of the KPMG global organisation of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved.

The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organisation.

Liability limited by a scheme approved under Professional Standards Legislation.

March 2023. 1059218486IGH