Contents

Executive Summary Human rights in the digital age 2

Section A
KPMG Australia’s recommendations for human rights in the digital age .................. 7
1. Future-proofing the debate on human rights and emerging technologies .............. 7
2. Human rights, safety and the environment: ensuring AI works for us ................... 9
3. Australia is a net importer of emerging technologies – does this challenge our ethical standards? ........................................................................................................... 10
4. Human rights and community benefits: practical decision making for business ........................................................................................................... 11
5. Privacy considerations and challenges in the digital age ........................................ 12
6. How to deal with the unintended consequences of new technologies ................. 14
7. Empowering human understanding and contestability of AI ................................. 15

Section B
Responses to the Australian Human Rights Commission Proposals and Questions.... 18
Part A: Introduction and framework ............................................................................ 18
Part B: Artificial intelligence ........................................................................................ 20
Part C: National leadership on AI .............................................................................. 28
Part D: Accessible technology .................................................................................... 30
Executive Summary
Human rights in the digital age

Technology holds great potential to help solve complex and interconnected sustainable development challenges such as human rights, climate change, access to health care, and inequality.\(^1\)

Australia’s digital innovation roadmap is forecast to deliver up to $315 billion in gross economic value over the next decade. High tech advancements in a number of domains - such as artificial intelligence (AI), human augmentation, blockchain, Internet of Things (IoT), quantum computing and green, nano and neuro-technologies, to name a few - as well as the increasingly exponential volumes and new forms of data (e.g. genomic and other bio data, video, voice) being exchanged, are anticipated to deliver benefits across sectors and geographies.

Notwithstanding the benefits of technological breakthroughs, there is a risk Australia will not fully realise its digital potential. In fact, negative impacts from new technologies on individuals, communities, and the environment can undermine their intended benefits. High profile incidents can also result in mistrust within the community, stalling the development of breakthrough technologies.

In our view, digital innovation and protection of human rights should not be treated as competing outcomes but as two sides of the same coin. We want to balance risk and reward enabling rights-based, ethical approaches which release competitive and breakthrough innovation in Australia.

These risks mainly arise because the pace of technological advancements far outweighs the ability of policy makers and regulators to provide clear guidelines on how to balance the benefits of technological innovation with legal, ethical and human rights considerations.

Our recommendations are centred on seven key themes:

1. Future-proofing the debate on human rights and emerging technologies to the greatest extent possible by conducting research on human rights implications of emerging technologies and ensuring that any regulatory frameworks are flexible and adaptable.

Recommendation 1.1:

Conduct early research on the human rights impacts of emerging technologies and their interoperability - in order to inform the development of guidelines and standards which are future-proof and forward thinking.

---

\(^1\) Source: https://assets.2030vision.com/files/resources/resources/state-of-play-report.pdf
Recommendation 1.2:
Adopt a flexible, ‘values based’, use-case driven framework which embeds human rights in the design, development and deployment of emerging technologies.

Recommendation 2:
Broadening the scope of the conversation on responsible innovation to embed considerations on sustainability and generating benefits to the planet.

Recommendation 2:
Human rights are fundamentally important and to truly achieve shared prosperity - any framework needs to ensure sustainability is embedded in the design, development, and deployment of emerging technologies to ensure that benefits are for the planet and well as for all humans.

Recommendation 3:
Understanding that Australia will often be a net importer of emerging technologies and any framework must ensure we remain competitive in sourcing and accessing cutting-edge and lower-cost technologies worldwide.

Recommendation 3:
Introduce a system to test and govern AI and emerging technologies that we import against our human rights data protection laws and ethical standards. This should be based on an internationally endorsed accreditation system developed by a recognised international standards body, where practical.

Recommendation 4:
Recognising the indivisibility of rights, encourage practical guidance on how to balance individual rights and community benefits to assist organisations in managing instances where the two seem to compete.

Recommendation 4:
Provide practical guidance on how to balance human rights and community benefits to assist organisations in managing instances of competing priorities. This should include consideration of protection of vulnerable people and communities, as a minimum.

The contributing team for this submission reflects our multi-disciplinary approach to the strategy, design, development and implementation of new and emerging technologies, as well as our industry leading experience in five key domains: data and technology; human rights and organisational trust; business strategy; data ethics and data management; regulation and privacy.
5. Understanding the challenges and limitations of the current Australian privacy framework vis-à-vis technological breakthroughs.

**Recommendation 5:**

Undertake additional consultation and release guidance on:

- The human rights impact of using specific types of data (e.g. sensitive information) to develop AI solutions.
- Data related concepts such as data integrity and quality, data ownership, data collection, anonymisation, de-identification, re-identification, encryption and their role in the context of AI and protection of human rights.
- The definition of “personal information” given the increasingly diverse types of data (e.g. video, voice, biometric, genomic, etc.) that could possibly trigger human rights violations.

6. Balancing innovation and regulation as the multiple positive applications of emerging technologies far outweigh the small proportion of misuse or negative unintended consequences. However in any trade-off, the most vulnerable should be protected.

**Recommendation 6.1:**

Ensure that the framework provides a balance between innovation and regulation and can accommodate ongoing adaptation. The multiple positive applications and usages far outweigh the small proportion of misuse or negative unintended consequences.

**Recommendation 6.2:**

Establish an ‘AI Commissioner’ or ‘AI Council’ with the remit to focus equally on safety and innovation in the sector.

7. Empowering trust by introducing productive transparency, broadening the scope of explainability and introducing a clear accountability framework.

**Recommendation 7.1:**

Introduce the concept of ‘productive transparency’ mandating informing people when they have agency in seeking an alternative process.

**Recommendation 7.2:**

Investigate the concept of ‘human in the loop’ noting that this is attracting support internationally.

**Recommendation 7.3:**

Include a principle of explainability which includes the requirement to disclose the nature and use of the information to be provided.
Recommendation 7.4:

Consider the adoption of a human rights due diligence approach which would enable a fair apportioning of accountabilities throughout the supply chain and allow organisations to confidently innovate while effectively protecting human rights.

Our submission is designed to help inform the ongoing development of policy and regulations to promote and embed the respect of human rights in emerging technologies to achieve mutually beneficial outcomes for both society and organisations. We begin by exploring our seven key recommendations in Section A: KPMG Australia’s recommendations for human rights in the digital age. In Section B we make specific responses to the Australian Human Rights Commission’s proposals and questions.

KPMG Australia is committed to actively participating in and collaborating with industry leaders and policy makers to explore the real-life applications and implications of emerging technologies, for both society and businesses in the near and long-term future.

We are also committed to responsible innovation which is driven by our Human Rights Policy and our approach to ethical AI.
Section A

KPMG Australia’s recommendations for human rights in the digital age
Section A

KPMG Australia’s recommendations for human rights in the digital age

1. Future-proofing the debate on human rights and emerging technologies

To future-proof our approach to understanding and addressing the intersection of human rights and technology we need flexible frameworks focused on impact. While future proofing may be exceedingly difficult in a sector that is changing daily, policy makers must aim to future proof any regulatory framework.

The AHRC has positioned the debate on ethical and human rights impact of emerging technologies around Artificial Intelligence. Notwithstanding the importance of addressing the challenges presented by AI, we welcome further consultation on the broader spectrum of current and emerging technological breakthroughs and how these might interoperate (e.g. IoT, blockchain and virtual reality) as these may have significant social and human rights impacts that have not been fully investigated and understood. For example, a technology on its own may be benign but in tandem with others may represent new risks and opportunities.

A use case framework may help inform and shape a “human rights by design” approach to new applications of technology from the earliest stage. This approach would help address, the challenges currently faced by more mature technologies that now require a retrospective assessment of their human rights and ethical adequacy.

Additionally, we acknowledge that there are many emerging AI related concepts that are still in their infancy and that the future is hard to predict. Therefore, in order to stay relevant, the human rights approach and framework would need to be flexible, to be iterated over time, and anchored to the context in which these technologies are being used. Even adopting a use-case driven framework has limitations as emerging technologies will have use cases that no one could have predicted. Policy makers must be mindful that frameworks, even based on use cases, must be able to adapt and flex over time.

Policy makers should consider whether a ‘values based’ framework should be implemented in consultation with ethical specialists and technologists. A conversation will need to be started around whether machines can be taught human rights and ethical behaviours.

Overall, we recommend to:
**Recommendation 1.1:**

Conduct early research on the human rights impacts of emerging technologies and their interoperability - in order to inform the development of guidelines and standards which are future-proof and forward thinking.

**Recommendation 1.2:**

Adopt a flexible, ‘values based’, use-case driven framework which embeds human rights in the design, development and deployment of emerging technologies.

This approach - which has already been explored in other jurisdictions (e.g. United Kingdom Data Ethics and Innovation Commission) - presents two key benefits:

- It anchors principles and guidelines to how these technologies are being used in various contexts rather than to the technology itself. This approach addresses the challenges associated with the different impact (i.e. positive vs, negative vs unintended consequences) that the same technology can have depending on how it is used.
- It results in a higher resiliency of the framework to technological advancements and breakthroughs.

Examples of current use-cases include risk profiling, tracking and monitoring, surveillance, manipulation of behaviours, health monitoring.
2. Human rights, safety and the environment: ensuring AI works for us

In an era of heightened awareness of the risks posed by climate change, international policy makers are broadening the scope of the conversation on responsible innovation to embed considerations on sustainability and climate footprint of emerging technologies.

For example, the OECD Council Recommendations on Artificial Intelligence\(^2\) recommends that AI “benefit people and the planet by driving inclusive growth, sustainable development and well-being”. The AI Now Institute recommends governments to “mandate public disclosure of the AI industry’s climate impact.” \(^3\)

In this context we recommend that policy makers consider the following:

---

**Recommendation 2:**

**Human rights are fundamentally important and to truly achieve shared prosperity - any framework needs to ensure sustainability is embedded in the design, development, and deployment of emerging technologies to ensure that benefits are for the planet and well as for all humans.**

---

Alongside guidance on climate change and sustainability, international policy makers also stress the importance that emerging technologies are secure and safe for human use. As an example, the OECD Council noted that “AI systems should be robust, secure and safe throughout their entire lifecycle so that, in conditions of normal use, foreseeable use or misuse, or other adverse conditions, they function appropriately and do not pose unreasonable safety risk.”

**Environmental, safety and security impacts are critical to the protection of human rights.** We propose that these considerations are reflected in the “human rights by design” and “human rights prioritisation” approaches outlined in this Discussion Paper.

---

**CASE STUDY**

_In a study conducted by the University of Massachusetts, Amherst, found that the process to train several common large AI models can emit more than 626,000 pounds of carbon dioxide equivalent—nearly five times the lifetime emissions of the average American car, including manufacture of the car itself. As electricity grids across the world decarbonise, the impact will be reduced, but should remain front of mind._

---


3. Australia is a net importer of emerging technologies – does this challenge our ethical standards?

Australia will likely be a net importer of AI and emerging technologies, given that international technology firms already own and power most commercial AI systems locally. In the ongoing quest to remain competitive, businesses will continue to source and access cutting-edge and lower-cost technologies internationally. This may result in ethical trade-offs where other countries may not have equivalent human rights protections in place.

Cloud based technology platforms also blur the boundaries of international borders. It may be unclear to companies about where an AI product’s decisions or processing is based. We need to be mindful of human rights trade-offs when sourcing technology but also don’t want to limit innovation and progress domestically.

Australian organisations deploying AI and emerging technology solutions will face significant challenges applying a “human rights by design” approach to technologies developed in jurisdictions with diverging and potentially conflicting human rights values and standards.

We recommend that policy makers:

**Recommendation 3:**

*Introduce a system to test and govern AI and emerging technologies that we import against our human rights data protection laws and ethical standards. This should be based on an internationally endorsed accreditation system developed by a recognised international standards body, where practical.*
4. Human rights and community benefits: practical decision making for business

The design, development and deployment of emerging technologies must manage instances where the respect for human rights seems to compete with a potential community-wide benefit.

"Business needs practical guidance on how to apply human rights principles to decisions on the use of technology."

This is a calculation governments must make all the time, such as determining the net human rights impact of public surveillance via technology, balancing the right to privacy with national security and the safety of the community.

Business already grapple with this in everyday decisions too. For example, using technology to identify particular community groups or specific demographics can be invaluable for designing responses to enhance access. However, it can also lead to direct or indirect discrimination, and exclusion of vulnerable groups considered unfit for a particular product or service.

We believe that further guidance on the topic would be beneficial in order to provide organisations with practical tools (e.g. decision-making guidelines, use-cases, etc.) to navigate the ethical dilemmas associated with privileging some human rights over others, or the provision of a benefit to one group that may lead to negative impacts on another. We recommend policy makers:

Recommendation 4:

Provide practical guidance on how to balance human rights and community benefits to assist organisations in managing instances of competing priorities. This should include consideration of protection of vulnerable people and communities, as a minimum.
5. Privacy considerations and challenges in the digital age

The Australian Privacy Principles apply to the collection and handling of individuals’ personal information including sensitive data such as health and biometric data. The federal Privacy Act has some legislative gaps related to employee records and small business exemptions.

The exemptions in the current data privacy legislation expose employees to the risk of discrimination and more generally the risk of unregulated tech start-ups developing AI based services that have a large privacy impact.

These gaps must be addressed in order to deal with new and emerging technology. These include de-identified data which is increasingly important in the development of autonomous decision making systems; the misuse of data which results in detrimental outcomes for consumers as well as employees; and where technological breakthroughs and innovations are often driven by smaller firms.

Data ownership is also a critical evolving concept that will need to be closely considered. Ownership drives much of the debate around consent and control.

The impact of the use of different types and combinations of data by AI, in particular sensitive information (such as health records, sexual orientation etc.) must be adequately addressed. The quality of the data and the ethics around how the data is collected is also paramount. The data set may be legally obtained in one jurisdiction, but its use in another is either illegal or unethical. This is relevant in the context of identifying the legal boundaries and limitations of the use of machine learning, AI and the higher impact decision-making outcome.

Further, the reliance on notice and consent needs revisiting – imposing a higher threshold of consent is not a panacea (see the Digital Platform Inquiry’s recommendations which include a review of the Privacy Act). Consent needs to be reconsidered in light of the functionalities and capabilities of technology in specific contexts of use, which are not static.

The latest international research and policy recommendations suggest that the scope of data protection regulations should be revisited to cover:

- **All types of information and information owners**, including employee data and information about legal persons (2019 Germany Data Ethics Commission recommendations, 20194).
- **All organisations, irrespective of their turnover** (as reflected in General Data Protection Regulation - GDPR).
- **De-identified data**. (Hong Kong Monetary Authority5 Ethical Accountability Framework).

In this context, we recommend policy makers:

---

4 Source: https://www.bmjv.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN.pdf?__blob=publicationFile&v=2
5 Source: https://www.pcpd.org.hk/misc/files/Ethical_Accountability_Framework.pdf
Recommendation 5:

Undertake additional consultation and release guidance on:

- The human rights impact of using specific types of data (e.g. sensitive information) to develop AI solutions.

- Data related concepts such as data integrity and quality, data ownership, data collection, anonymisation, de-identification, re-identification, encryption and their role in the context of AI and protection of human rights.

- The definition of “personal information” given the increasingly diverse types of data (e.g. video, voice, biometric, genomic, etc.) that could possibly trigger human rights violations.
6. How to deal with the unintended consequences of new technologies

Technology breakthroughs often present unexpected and unintended consequences which are difficult to predict at the design stage.

CASE STUDY

An example of the benefits associated with facial recognition is technology underdevelopment that may be used to better assess non-verbal patients with Dementia and similar conditions for levels of pain. This information is shared with their caregivers, potentially allowing for more effective pain management.

An example of these challenges is represented by facial recognition which is increasingly being used by public and private companies in Australia.

The misuse of facial recognition technology poses a threat to human rights, however some applications can promote human rights by helping bring criminals to justice. 3D Printing has revealed similar threats (e.g. printing weapons) and benefits (e.g. printing personalised artificial bones for reconstructive surgery).

Policy makers should also consider the role of ‘device-less interfaces’ and human embedded technologies and devices. Companies are moving to ‘frictionless’ interactions where phones and devices are replaced with facial and other recognition.

In general, we recommend that policy makers:

**Recommendation 6.1:**

Ensure that the framework provides a balance between innovation and regulation and can accommodate ongoing adaptation. The multiple positive applications and usages far outweigh the small proportion of misuse or negative unintended consequences.

**Recommendation 6.2:**

Establish an ‘AI Commissioner’ or ‘AI Council’ with the remit to focus equally on safety and innovation in the sector.
7. Empowering human understanding and contestability of AI

A critical success factor in the deployment and acceptance of AI, and any future new technology, will be “empowering humans” to understand (1) how the technology works, (2) the logic behind its decisions or actions and (3) the information that is used as input to train and develop these machines.

For example, people should know if AI is deciding a loan application for them and the basis on which the decision has been made. However, it simply would not be productive for passengers on a plane to be told every time a pilot engages the autopilot function. If we do not inform people about the use of AI when they have no alternative process to choose, we risk eroding trust in proven systems and violating human rights.

Human empowerment is driven by:

**Transparency:** we agree with the principle of transparency outlined by the Australian Human Rights Commission. We welcome further guidance on the type of information and level of detail to be provided to individuals, noting that full transparency may not be practical in all cases. Further, the source code may not be able to fully explain how a decision is made, it may be beneficial to instead focus on the data that has driven the end outcome.

The concept of ‘human-in-the-loop’ is also gaining momentum internationally. The concept empowers people to use AI to do their jobs better and ensures a computer-human interface. Having a human-in-the-loop will also promote understanding and acceptance within the community.

In particular, we recommend policy makers:

**Recommendation 7.1:**

*Introduce the concept of ‘productive transparency’ mandating informing people when they have agency in seeking an alternative process.*

**Recommendation 7.2:**

*Investigate the concept of ‘human in the loop’ noting that this is attracting support internationally.*

**Explainability and contestability:** we agree with the principle of explanation outlined by the Human Rights Commission. We also recognise that AI will increase the information asymmetry between what organisations know and what is explained to their customers regarding their decisions, systems and services. Therefore, we recommend that policy makers:
Recommendation 7.3:
Include a principle of explainability which includes the requirement to disclose the nature and use of the information to be provided.

As an example, an explanation about the inputs and rules an algorithm uses to make decisions should be easily accessible to those impacted by it. We also welcome further guidance on the following points:

- The process to access explanations
- Timeframes and format of the explanation
- The process to challenge explanations
- If there would be an independent body that would be responsible for reviewing, assessing and resolving disputes
- Situations where explanations may be withheld due to the need to protect the community’s interests and rights (e.g. public safety)
- Remediation of harm that may be caused.

Accountability: we agree with the concept of accountability outlined by the Australian Human Rights Commission. Emerging technology applications and systems can involve a range of different agents with complex interrelationships, which could lead to a complex system of apportioning accountability. In this context, we recommend policy makers:

Recommendation 7.4:
Consider the adoption of a human rights due diligence approach which would enable a fair apportioning of accountabilities throughout the supply chain and allow organisations to confidently innovate while effectively protecting human rights.

“Ultimately, we believe that a solution needs to be developed that maintains trust in AI and is focussed on the rights of the individual”

As part of the due diligence, organisations would be responsible as a minimum for ensuring proper impact and risk assessments, safeguards to prevent harm from the decision, as well as appropriate plans for remedy should harm occur.
Section B

Responses to the Australian Human Rights Commission Proposals and Questions
Section B
Responses to the Australian Human Rights Commission Proposals and Questions

This section summarises our feedback in relation to proposals and questions submitted by the Human Rights Commission.

Part A: Introduction and framework

1. The Australian Government should develop a National Strategy on New and Emerging Technologies. This National Strategy should:
   a. Set the national aim of promoting responsible innovation and protecting human rights
   b. Prioritise and resource national leadership on AI
   c. Promote effective regulation—this includes law, co-regulation and self-regulation
   d. Resource education and training for government, industry and civil society.

KPMG Australia Response:

We welcome further clarification on:

- The current strategic strategies that are relevant in the context of the development of Australian National Strategy on New and Emerging Technologies (e.g. Data61 AI Ethics Framework; Digital Transformation Agency Vision 2025).
- The scope of the strategy, and specifically, if it will cover all types of emerging technologies or solely AI.
- The process to qualify if a technology is “new” and “emerging” and the process to dynamically identify those.
- The inclusion of existing technologies where new and emerging applications may result in detriment of human rights (e.g. use of social media in the age of Big Data).
- The net importer challenge and how to address it.

2. The Australian Government should commission an appropriate independent body to inquire into ethical frameworks for new and emerging technologies to:
   a. Assess the efficacy of existing ethical frameworks in protecting and promoting human rights
   b. Identify opportunities to improve the operation of ethical frameworks, such as through consolidation or harmonisation of similar frameworks, and by giving special legal status to ethical frameworks that meet certain criteria.
KPMG Australia’s Response:

"Ethical frameworks” or “codes of practice” will not achieve much unless there is also an enforcement mechanism to back them up

An effective enforcement strategy of existing regulations should be considered. There are multiple legislative reforms currently being considered and planned, such as the Consumer Data Right, Data Sharing and Release Reforms, Digital Platforms Inquiry (DPI), Access & Assistance Act (A&A Act) to address access and sharing of private and public data. These should not happen in isolation.

While “principles-based” regulation can take a long time for the industry (and the regulators) to work out what amounts to good compliance, it does have a role to play but means it is even more important to have a strong regulator enforcing the new arrangements. Clear regulatory guidance on their application as well as some more prescriptive requirements would be more useful.

In addition, in the process of constructing a regulatory framework for AI attention should also be given to Australian legislation that may require further review following feedback by some privacy and civil liberty stakeholders, such as the A&A Act. Further, Australia’s regulatory and legal approach to AI should not be inconsistent with what is expected from other national and international regulatory bodies.

Notwithstanding the benefits of technological breakthroughs, there is a risk Australia will not fully realise its digital potential.
Part B: Artificial intelligence

**Question A:** The Commission’s proposed definition of ‘AI-informed decision making’ has the following two elements: there must be a decision that has a legal, or similarly significant, effect for an individual; and AI must have materially assisted in the process of making the decision.

Is the Commission’s definition of ‘AI-informed decision making’ appropriate for the purposes of regulation to protect human rights and other key goals?

**KPMG Australia’s Response:**
We agree with this definition and welcome the opportunity to conduct research on more advanced forms of AI in order to inform the development of principles and frameworks.

**Proposal 3:** The Australian Government should engage the Australian Law Reform Commission to conduct an inquiry into the accountability of AI-informed decision making. The proposed inquiry should consider reform or other change needed to:

a. Protect the principle of legality and the rule of law

b. Promote human rights such as equality or non-discrimination.

**KPMG Australia’s Response:**
We agree with this proposal in principle, however we suggest that the Australian Law Reform Commission partner with international standards bodies where possible.

**Proposal 4:** The Australian Government should introduce a statutory cause of action for serious invasion of privacy.

**KPMG Australia’s response:**
A statutory tort for serious invasion of privacy is a serious cause of action that will have social and organisational impacts on both government and private sector. We recommend that the Government approach this issue with caution. We note that in December 2019 the Government through the DPI acknowledged the ACCC’s recommendation to establish a statutory cause of action for serious invasion of privacy. The Government stated that this recommendation would need to be considered through the review of the Privacy Act 1988.

**Proposal 5:** The Australian Government should introduce legislation to require that an individual is informed where AI is materially used in a decision that has a legal, or similarly significant, effect on the individual’s rights.

**KPMG Australia’s Response:**
We welcome further guidance on the definition of “legal” and “similarly significant” effects. For example, it is unclear whether an impact on an individual’s reputation, health (physical and psychological), and financial, professional or legal position would be encompassed within the term ‘impact’.

The approach taken in the European Union’s GDPR may be instructive. Under Article 22, the GDPR requires that, for any automated decision with “legal effects or similarly significant effects” such as employment, credit, or insurance coverage, the person affected has recourse to a human who can review the decision and explain its logic. This incorporates a human component and an element of due process that provide a check on anomalous or unfair outcomes. This safety mechanism partly addresses Proposal 8 and should be available for algorithmic decisions that have a material impact on individuals’ lives.
Because of its ambiguity, however, the term “legal effects or similarly significant effects” will need further explanation and probably additional guidance from the regulators as there are no commonly accepted rules as to what kind of decisions have significant effects on individuals.

The right to object should also be considered in the case of AI. Once informed, individuals should have the right to object to their personal information being processed by an AI-informed decision-making system if the operator cannot demonstrate compelling, legitimate grounds for the processing. These grounds must be sufficiently compelling to override the interests, rights and freedoms of the individual, such as to establish, exercise or defend against legal claims. This aligns with the human-focused approach adopted by the European Union’s GDPR.

**Proposal 6:** Where the Australian Government proposes to deploy an AI-informed decision-making system, it should:

a. Undertake a cost-benefit analysis of the use of AI, with specific reference to the protection of human rights and ensuring accountability

b. Engage in public consultation, focusing on those most likely to be affected

c. Only proceed with deploying this system, if it is expressly provided for by law and there are adequate human rights protections in place.

**KPMG Australia’s Response:**

We agree with this proposal. We highlight the importance of ensuring that:

- There is sufficient scope for testing and learning within agreed parameters so as not to create a disincentive to experiment (safely);
- That any system is informed by international best practice so that the Australian marketplace is not disadvantaged by a unique Australian system;
- Learnings are gathered before full scaled implementation that might require full analysis and consultation; and
- The government conduct a human rights impact assessment on any proposed deployment of AI informed decision making as part of the public consultation to identify any mitigation measures that might be required to protect vulnerable people or provide alternative pathways for the AI-informed decision making system, noting that special consideration may need to be granted for activities that fall under existing national security legislation.

**Proposal 7:** The Australian Government should introduce legislation regarding the explainability of AI-informed decision making. This legislation should make clear that, if an individual would have been entitled to an explanation of the decision were it not made using AI, the individual should be able to demand:

a. A non-technical explanation of the AI-informed decision, which would be comprehensible by a lay person, and

b. A technical explanation of the AI-informed decision that can be assessed and validated by a person with relevant technical expertise.

In each case, the explanation should contain the reasons for the decision, such that it would enable an individual, or a person with relevant technical expertise, to understand the basis of the decision and any grounds on which it should be challenged.

**KPMG Australia’s Response:**

Please refer to the **Empowering human understanding and contestability of AI section** of this submission paper.
Proposal 8: Where an AI-informed decision-making system does not produce reasonable explanations for its decisions, that system should not be deployed in any context where decisions could infringe the human rights of individuals.

KPMG Australia’s Response:

The intent of this proposal is appropriate, however, we seek further guidance on how to determine when an explanation is deemed to be reasonable.

Explainability may require (1) identifying algorithmic decisions, (2) deconstructing specific decisions, and (3) establishing a channel by which an individual can seek an explanation. Reverse-engineering algorithms based on machine learning can be difficult, and even impossible in some cases. This difficulty increases as machine learning becomes more sophisticated.

Explainability therefore may entail a significant regulatory burden and constraint on use of AI algorithmic decision-making and, in this light, should be concentrated in the breath of its application, as the EU has done (at least in principle) with its “legal effects or similarly significant effects” threshold as mentioned in Proposal 5 above.

Question B: Where a person is responsible for an AI-informed decision and the person does not provide a reasonable explanation for that decision, should Australian law impose a rebuttable presumption that the decision was not lawfully made?

KPMG Australia’s Response:

Regulators should be careful with how liability regimes for AI are designed. The European Commission’s Liability for Artificial Intelligence and Other Emerging Digital Technologies Report (released in November 2019) proposed that if the operation of some technology that includes AI, for example, is legally permissible, presuming that the developer made use of state-of-the-art knowledge at the time the system was launched. Any subsequent choices made by the AI technology independently may not necessarily be attributable to some flaw in its original design. The question therefore arises whether the choice to admit it to the market, or implement the AI system in an environment where harm was subsequently caused, in itself is a breach of the duties of care applicable to such choices. The more complex the circumstances leading to the victim’s harm are, the harder it is to identify relevant evidence and the root cause.

We acknowledge that this is a complex topic which requires further consultation and including:

- Clarifications on whether this principle would equally apply to government and public agency decisions
- On what grounds the presumption could be rebutted
- The related consequences.

Proposal 9: Centres of expertise, including the newly established Australian Research Council Centre of Excellence for Automated Decision-Making and Society, should prioritise research on how to design AI-informed decision-making systems to provide a reasonable explanation to individuals.

KPMG Australia’s Response:

We agree with the proposal and suggest to:

- Prioritise the scope of the research on how to submit, manage and resolve situations where individuals challenge the explanations provided.
• Clarify the definition of “deployment” specifically in relation to circumstances where organisations deploy AI-informed decision making systems developed by third parties.

• Mirror best practice in other jurisdictions where practical.

Proposal 10: The Australian Government should introduce legislation that creates a rebuttable presumption that the legal person who deploys an AI-informed decision-making system is legally liable for the use of the system.

KPMG Australia’s Response:

For our recommendations on this proposal please refer to section Empowering human understanding and contestability of AI.

Question C: Does Australian law need to be reformed to make it easier to assess the lawfulness of an AI-informed decision-making system, by providing better access to technical information used in AI-informed decision-making systems such as algorithms?

KPMG Australia’s Response:

We recognise that this is a complex issue which requires further consultation. We also recognise that in the case of more advanced forms of AI – e.g. deep learning – the access to the technical information used in the AI-informed decision making system may not result in increased transparency and explainability. Other aspects to consider include what data is being used to train the system and the validation procedures to which the system has been subjected.

Question D: How should Australian law require or encourage the intervention by human decision makers in the process of AI-informed decision making?

KPMG Australia’s Response:

As technology advances we are seeing some AI applications outperforming human reasoning and decision making. In this context, our view is that guidance should be provided to encourage the embedding of humans in the loop - where there is a clear benefit - as well as to navigate where human intervention is warranted.

Proposal 11: The Australian Government should introduce a legal moratorium on the use of facial recognition technology in decision making that has a legal, or similarly significant, effect for individuals, until an appropriate legal framework has been put in place. This legal framework should include robust protections for human rights and should be developed in consultation with expert bodies including the Australian Human Rights Commission and the Office of the Australian Information Commissioner.

KPMG Australia’s Response:

We welcome further consultation on this topic and suggest to consider the introduction of a process (in lieu of a legal moratorium) which requires organisations to discuss the use case for applying facial recognition with a designated regulator/AI Commissioner to seek approval or licensing for the use case.

Proposal 12: Any standards applicable in Australia relating to AI-informed decision making should incorporate guidance on human rights compliance.

KPMG Australia’s Response:

We agree with this proposal following a cost benefit analysis finding net benefit. We note that business will benefit from practical guidance, which is rooted in the UN Guiding Principles on Business and Human Rights and encourages a human rights due diligence approach to identify, assess and mitigate potential negative human rights impacts.
Proposal 13: The Australian Government should establish a taskforce to develop the concept of ‘human rights by design’ in the context of AI-informed decision making and examine how best to implement this in Australia. A voluntary, or legally enforceable, certification scheme should be considered. The taskforce should facilitate the coordination of public and private initiatives in this area and consult widely, including with those whose human rights are likely to be significantly affected by AI-informed decision making.

KPMG Australia’s Response:

We agree with this proposal in principle. The taskforce should comprise a broad range of stakeholders – technologists, academia, legal profession, ethicists, human rights advocates, professional services firms and other relevant stakeholders and thought-leaders in this space. In line with the UN Guiding Principles on Business and Human Rights, we would expect to see targeted consultation with rights-holders and vulnerable populations differentiated by use-case.

Proposal 14: The Australian Government should develop a human rights impact assessment tool for AI-informed decision making, and associated guidance for its use, in consultation with regulatory, industry and civil society bodies. Any ‘toolkit for ethical AI’ endorsed by the Australian Government, and any legislative framework or guidance, should expressly include a human rights impact assessment.

KPMG Australia’s Response:

We agree with this proposal in principle. We note however, that toolkit approaches to human rights impact assessments (HRIAs) without the input of subject matter experts or capability building, can lead to tick-box or compliance oriented approaches to assessment that will fail to surface negative human rights impacts, or design appropriate mitigation measures.

We suggest that government has a dedicated resource for its own HRIAs and provide pathways for support for business. Methodologies that have been applied to conduct HRIAs in other sectors will need to be tested and refined for AI-informed decision making.

Question E: In relation to the proposed human rights impact assessment tool in Proposal 14:

a. When and how should it be deployed?

b. Should completion of a human rights impact assessment be mandatory, or incentivised in other ways?

c. What should the consequences be if the assessment indicates a high risk of human rights impact?

d. How should a human rights impact assessment be applied to AI-informed decision-making systems developed overseas?

KPMG Australia’s Response:

Human rights impact assessments should be supplemented by assessment of the broader environmental, sustainability and safety impacts of the use of any particular technology. Like a privacy risk assessment (PIA), a human rights impact assessment should focus on the potentially affected individuals, vulnerable groups or communities and the impact on their human rights.

Risk assessments for algorithmic decision-making may provide an opportunity to anticipate potential biases in design and data as well as the potential impact on individuals’ human rights. The level of risk assessment should be appropriate to the significance of the AI-informed decision-making in question, which depends on the consequences of those decisions, the
number of people and volume of data potentially affected, and the novelty and complexity of algorithmic processing.

Consumer data rights have become an essential part of the consumer and privacy rights discussion in Australia and it would be appropriate to consider the impacts on consumer rights and human rights at large, as well as the risks of failing to comply with its standards.

Whether or not human rights impact assessments are made mandatory needs to be considered in the context of whether or not human rights due diligence is made mandatory. Assessment is only one component of the identification, assessment, monitoring and remediation required to manage human rights impacts and risks.

Proposal 15: The Australian Government should consider establishing a regulatory sandbox to test AI-informed decision-making systems for compliance with human rights.

KPMG Australia’s Response:
We agree with the proposal and highlight the importance of:

- Broadening the scope of the testing activities enabled by the sandbox to assessing compliance with other relevant regulations as well as ethics principles
- Making the sandbox available to public and private organisations
- Potentially broadening the scope of the sandbox to include other emerging technologies that have relevant impact on human rights.

Question F: What should be the key features of a regulatory sandbox to test AI-informed decision-making systems for compliance with human rights? In particular:

a. What should be the scope of operation of the regulatory sandbox, including criteria for eligibility to participate and the types of system that would be covered?

b. What areas of regulation should it cover e.g., human rights or other areas as well?

c. What controls or criteria should be in place prior to a product being admitted to the regulatory sandbox?

d. What protections or incentives should support participation?

e. What body or bodies should run the regulatory sandbox?

f. How could the regulatory sandbox draw on the expertise of relevant regulatory and oversight bodies, civil society and industry?

g. How should it balance competing imperatives e.g., transparency and protection of trade secrets?

h. How should the regulatory sandbox be evaluated?

KPMG Australia’s Response:
We welcome the initiative of introducing a sandbox, in line with local (e.g. ASIC sandbox) as well as international (e.g. U.K. Financial Conduct Authority sandbox).

Our view is that the regulatory sandbox should focus on validating the efficacy of AI-informed decision making systems and compliance with AI-ethics and Human Rights frameworks covered in this Discussion Paper.

To further assist with validating efficacy of AI systems, the sandbox should comprise approved datasets that organisations can use to develop and test their AI applications. It is essential that an organisation’s intellectual property is protected to ensure that the sandboxes are utilised.
In any event, an ethical and human rights approach for the regulatory sandbox and how to evaluate it should be founded in and start with a clear and accurate explanation of the law (including its foundational concepts). This approach needs to address the regulatory limitations of AI-informed decision making and the need for human intervention, control and responsibility. For example, AI can be useful in pre-filtering—such as for rough matching (e.g., fuzzy logic matching) but should be relied on as the only basis for making critical decisions (those in the scope of producing “legal effects or similarly significant effects”).

In regard to what bodies should run the regulatory sandbox, it is essential to have regulators that are prepared to have a strong approach to enforcement. Regulators that are seen as “too friendly”, “too under resourced” or simply too scared (or inexperienced) to take a proactive approach might not be appropriate to lead the regulatory sandbox.

The OAIC has the tools to police activities but it has not yet used its power to prosecute serious interferences with privacy (that were introduced in 2014) and now it is being given powers to impose higher penalties. The question is whether it will actively use these powers. On the other hand, the ACCC has also been given more powers to pursue consumer-related offences, such as those given as a result of the Data Platform Inquiry.

The ACCC is seeking to expand its powers to also include a prohibition against “unfair trading practices” which would capture a wider range of conduct. However, in practice, the ACCC is only likely to use these powers against larger players, so it is questionable whether these powers would be sufficient to address industry-wide ethical issues in AI or the use of technology. The introduction of the Consumer Data Right regime is a testament that the co-regulatory model is on the rise and regulators like the ACCC and the OAIC can be part of the solution. They, however, cannot answer all ethical issues. No one approach, framework, law or regulator can address these issues. It will require a re-think in the regulatory approach and an overarching framework reflecting a range of laws that will apply.

Proposal 16: The proposed National Strategy on New and Emerging Technologies (see Proposal 1) should incorporate education on AI and human rights. This should include education and training tailored to the particular skills and knowledge needs of different parts of the community, such as the general public and those requiring more specialised knowledge, including decision makers relying on AI data points and professionals designing and developing AI-informed decision-making systems.

KPMG Australia’s Response:

We agree with this proposal and highlight the importance to adopt a use-case based approach to training to facilitate the understanding of the real life applications of emerging technologies and to include in the education programme other topics – such as relevant data regulations, ethical principles and standards – which are equally important to inform the design, development and deployment of AI solutions which respect and protect human rights.

Proposal 17: The Australian Government should conduct a comprehensive review, overseen by a new or existing body, in order to:

a. Identify the use of AI in decision making by the Australian Government
b. Undertake a cost-benefit analysis of the use of AI, with specific reference to the protection of human rights and ensuring accountability
c. Outline the process by which the Australian Government decides to adopt a decision-making system that uses AI, including any human rights impact assessments
d. identify whether and how those impacted by a decision are informed of the use of AI in that decision-making process, including by engaging in public consultation that focuses on those most likely to be affected

e. Examine any monitoring and evaluation frameworks for the use of AI in decision-making.

KPMG Australia’s Response:
We agree with the proposal. Any new regulation should be informed by a cost benefit analysis and relevant best practice in overseas jurisdictions.

Proposal 18: The Australian Government rules on procurement should require that, where government procures an AI-informed decision-making system, this system should include adequate human rights protections.

KPMG Australia’s Response:
Further guidance is needed on “adequate” human rights protections especially when technology solutions are sourced from countries with less stringent ethical and legislative frameworks around data protection, human rights and ethics. Typically, for business, a human rights due diligence approach would require demonstrated practice from suppliers in high-risk jurisdictions that the supplier had met international human rights standards, rather than simply complied with local law. This could be achieved through an international accreditation model.

Special consideration needs to be applied to situations where AI-informed decision making systems are procured in the interests of national security.
Part C: National leadership on AI

Proposal 19: The Australian Government should establish an AI Safety Commissioner as an independent statutory office to take a national leadership role in the development and use of AI in Australia. The proposed AI Safety Commissioner should focus on preventing individual and community harm, and protecting and promoting human rights. The proposed AI Safety Commissioner should:

a. Build the capacity of existing regulators and others regarding the development and use of AI
b. Monitor the use of AI, and be a source of policy expertise in this area
c. Be independent in its structure, operations and legislative mandate
d. Be adequately resourced, wholly or primarily by the Australian Government
e. Draw on diverse expertise and perspectives
f. Determine issues of immediate concern that should form priorities and shape its own work.

KPMG Australia’s Response:

In principle, we agree with the proposal which is aligned with similar initiatives undertaken by some countries. However we do not support ‘safety’ as the primary focus of the Commissioner. Safety and innovation should be balanced equally.

As for point a) consideration should be given on developing an AI curriculum for governmental officials to build capability on AI, emerging technologies, benefits and opportunities as well as risks. This program could be shaped in collaboration with national or international universities and other subject matter experts. Additionally, further consideration should be given to the harmonisation of the role of other regulatory bodies and agencies (e.g. ACCC, D61, OAIC) and the role of an AI Commissioner. In the interim while the technology matures, it may be prudent for the OAIC to take on the role as Commissioner to conserve resources and expertise.
Part D: Accessible technology

KPMG Australia’s Response:

KPMG Australia has carefully considered proposals 20 – 25 which focus on particular levers for enhancing accessibility and monitoring compliance. We support measures that increase the accessibility of emerging and current technologies.

In our view, comprehensive review of existing mechanisms for establishing and maintaining accountability for the development of accessible technology is an important building block. We urge, however, that accessibility questions are not resolved simply through these existing mechanism and that the principled based approach to ensuring that the most vulnerable are not disadvantaged in pursuit of community-wide benefits discussed above is central to any frameworks or requirements.

A cost benefit analysis will be a useful step in the development of any Standard or any new requirement to anticipate potential outcomes and mitigate costs, while at the same time ensuring that people with disabilities are given agency and voice. However, it is important to remember that a cost benefit analysis should only make up one part of any final review.
Key contacts

Dr Sanjay Mazumdar
Chief Data Officer, KPMG Australia
E: skmazumdar@kpmg.com.au
T: +61 478 403 462

Rossana Bianchi
Data Ethics Lead, KPMG Australia
E: rbianchi2@kpmg.com.au
T: +61 438 141 093

Dr Meg Brodie
Director, Human Rights and Social Impact, KPMG Banarra
E: megbrodie@kpmg.com.au
T: +61 400 931 275

Veronica Scott
Director, Privacy, Technology & Data Protection, KPMG Law
E: vscott1@kpmg.com.au
T: +61 411 206 248

Sophie Finemore
Associate Director, Regulatory Affairs
E: sfinemore@kpmg.com.au
T: +61 409 414 634

The primary authors also acknowledge the contributions of the following KPMG staff - James Mabbott, Vanessa Wolfe-Coote, Jon Stone, Ben Wallace, Zoe Willis, Kate Marshall, Rita Fentener van Vlissingen, David Evans, Leo Zhang, Richard Boele, Renuka Kimber, Christine Wilcox, Spencer Hulme, Michael Sutton and Zuong Dang.