Australian Human Rights Commission

Human Rights and Technology

DISCUSSION PAPER • DECEMBER 2019

© Australian Human Rights Commission 2019.

We encourage the dissemination and exchange of information presented in this publication and endorse the use of the **Australian Governments Open Access and Licensing Framework** (AusGOAL).

All material presented in this publication is licensed under the **Creative Commons Attribution 4.0 International Licence**, with the exception of:

- photographs and images
- the Commission's logo, any branding or trademarks
- where otherwise indicated.

To view a copy of this licence, visit **http://creativecommons.org/licenses/by/4.0/legalcode**.

In essence, you are free to copy, communicate and adapt the publication, as long as you attribute the Australian Human Rights Commission and abide by the other licence terms.

Please give attribution to: © Australian Human Rights Commission.

Human Rights and Technology • Discussion Paper • 2019

ISBN 978-1-925917-15-4

For further information about copyright in this publication, please contact:

Communications Unit
Australian Human Rights Commission
GPO Box 5218
SYDNEY NSW 2001
Telephone: (02) 9284 9600
TTY: 1800 620 241
Email: **communications@humanrights.gov.au**

Design and layout: Dancingirl Designs
Cover image: iStock
Internal photography: iStock, Alamy, AAP, Disability Images

The Australian Human Rights Commission is Australia's National Human Rights Institution. It is an independent statutory organisation with responsibility for leading the promotion and protection of human rights in Australia. Further information about the Commission can be found at **www.humanrights.gov.au/about-commission**.

Authors: Sophie Farthing, John Howell, Katerina Lecchi, Zoe Paleologos, Phoebe Saintilan and Edward Santow.

Acknowledgements:

- President Rosalind Croucher, National Children's Rights Commissioner Megan Mitchell and Disability Discrimination Commissioner Ben Gauntlett.
- Australian Human Rights Commission staff: Partha Bapari, Darren Dick, Connie Kwan, Lauren Perry, Liz Stephens and Leon Wild.
- Consultant editor: Bruce Alston.
- Expert Reference Group for their peer review and advice in the preparation of this Paper: Amanda Alford, Her Excellency The Honourable Margaret Beazley AO AC, Professor Genevieve Bell, Peter Dunne, Dr Tobias Feakin, Dr Alan Finkel AO, Verity Firth, Peter Leonard, Brooke Massender, Sean Murphy, Kwok Tang, Myfanwy Wallwork and Professor Toby Walsh.
- Major project partners for contributing expertise and resources: the Australian Government Department of Foreign Affairs and Trade, Herbert Smith Freehills, LexisNexis, The University of Technology Sydney.
- The Australian Government Digital Transformation Agency and Lisa Webber Corr; World Economic Forum.

Major project partners

Australian Government
Department of Foreign Affairs and Trade

UTS UNIVERSITY OF TECHNOLOGY SYDNEY

HERBERT SMITH FREEHILLS

LexisNexis®

Human Rights and Technology

Discussion Paper

December 2019

Australian Human Rights Commission

Australian
Human Rights
Commission

Contents

Commissioner's foreword 7

PART A : INTRODUCTION AND FRAMEWORK 9

Chapter 1: Introduction 9

1.1 Background to the Project 11
1.2 Public consultation 11
1.3 Other issues 13

Chapter 2: Human rights framework 15

2.1 Introduction 15
2.2 What are human rights? 16
2.3 How are human rights engaged by new and emerging technologies? 22
2.4 Applying a human rights approach 25
2.5 Accountability and responsibility 29
2.6 The Commission's preliminary view 31

Chapter 3: Regulation 33

3.1 Introduction 33
3.2 Regulation and new technologies 34
3.3 The Commission's preliminary view 38
3.4 National regulation and governance 42
3.5 The Commission's preliminary view 45

Chapter 4: Ethical frameworks 47

4.1 Introduction 47
4.2 Emergence of ethical frameworks 48
4.3 Impact of ethical frameworks 52
4.4 The Commission's preliminary view 55

PART B: ARTIFICIAL INTELLIGENCE 59

Chapter 5: AI-informed decision making 59

5.1 Introduction 59
5.2 What is Artificial intelligence? 60
5.3 What is AI-informed decision making? 61
5.4 How human rights are engaged by the use of AI 65
5.5 AI and regulation 68
5.6 The Commission's preliminary view 71

Chapter 6: Accountable AI-informed decision making 75

6.1 Introduction 75
6.2 The elements of accountability 76
6.3 Lawfulness 76
6.4 Transparency 82
6.5 Explainability 82
6.6 Responsibility and liability 86
6.7 Human intervention in AI-informed decision making 89
6.8 The Commission's preliminary view 89

Chapter 7: Co- and self-regulatory measures for AI-informed decision making 107

7.1 Introduction 107
7.2 Design, standards and impact assessments 109
7.3 The Commission's preliminary view 114
7.4 Regulatory sandboxes 118
7.5 The Commission's preliminary view 119
7.6 Education and training 121
7.7 The Commission's preliminary view 123
7.8 Oversight of government use of AI 124
7.9 The Commission's preliminary view 128

PART C: NATIONAL LEADERSHIP ON AI 131

Chapter 8: National leadership on AI 131

8.1 Introduction 131
8.2 Is there a need for leadership? 131
8.3 The Commission's preliminary view 136
8.4 What should the AI Safety Commissioner look like? 137
8.5 The Commission's preliminary view 143

PART D: ACCESSIBLE TECHNOLOGY 147

Chapter 9: The right to access technology 147

9.1 Introduction 147
9.2 Right to access technology as an enabling right 153
9.3 The Commission's preliminary view—functional accessibility 160
9.4 Obtaining Digital Technologies 164
9.5 The Commission's preliminary view—obtaining Digital Technologies 166

Chapter 10: Design, education and capacity building 169

10.1 Introduction 169
10.2 Human rights by design 170
10.3 The Commission's preliminary view 175

Chapter 11: Legal protections 181

11.1 Introduction 181
11.2 Submissions and consultations 181
11.3 The Commission's preliminary view 184

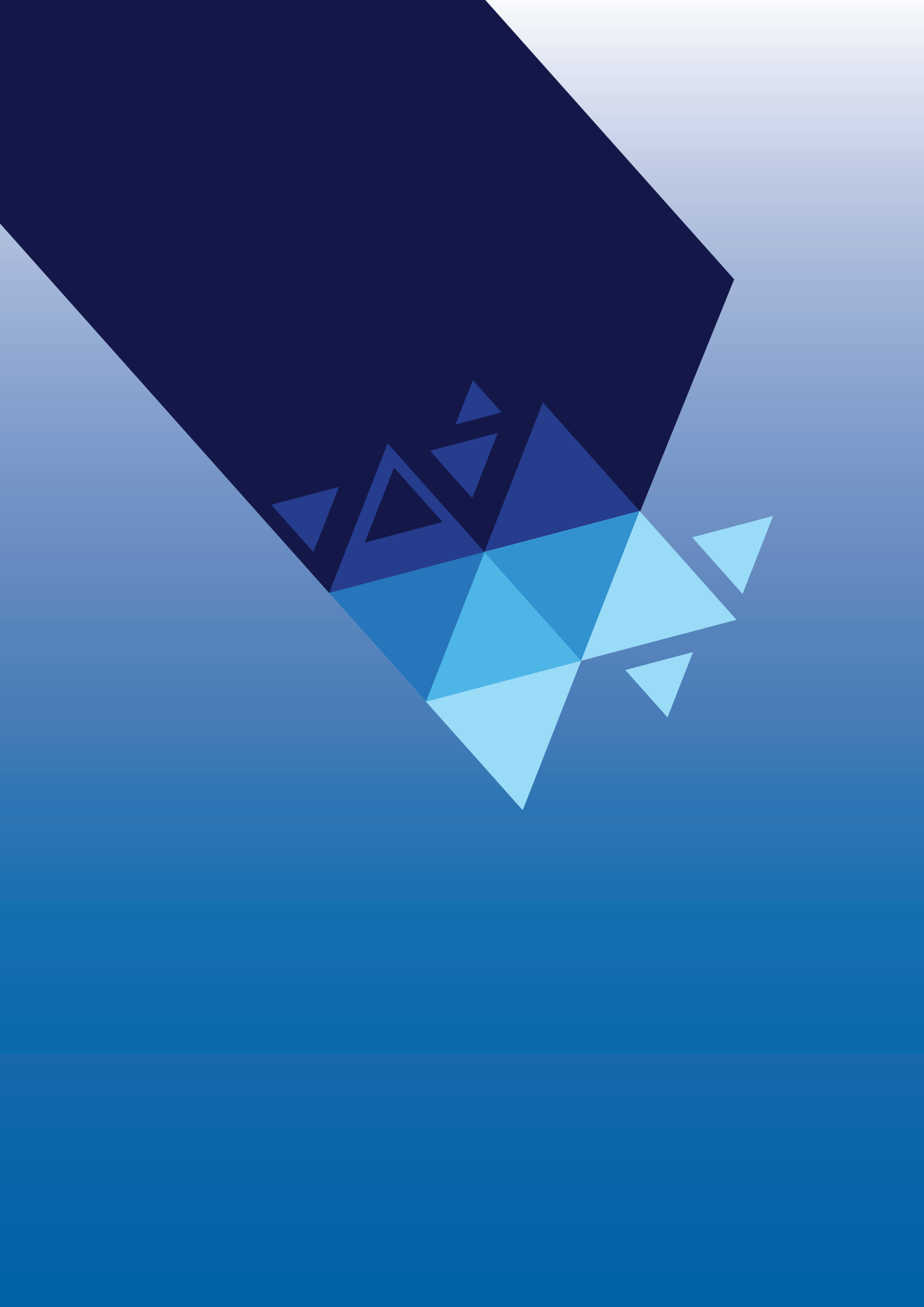
PART E: CONSULTATION 189

Chapter 12: Proposals and questions 189

Chapter 13: Making a submission 195

Appendices

Appendix A – List of submissions 197
Appendix B – Acronyms 202



Edward Santow

Human Rights Commissioner
Australian Human Rights Commission

Commissioner's foreword

Foreword from the Australian Human Rights Commissioner

This Discussion Paper sets out the Australian Human Rights Commission's preliminary views on protecting and promoting human rights amid the rise of new technologies.

New technologies are changing our lives profoundly—sometimes for the better, and sometimes not.

We have asked a fundamental question: how can we ensure these new technologies deliver what Australians need and want, not what we fear?

We recently completed the first phase of the Commission's public consultation.

We heard how Australian innovators are building our economy, revolutionising how health care and other services are delivered, and using data to make smarter decisions.

But we also saw how artificial intelligence (AI) and other new technologies can threaten our human rights. Time and again people told us, 'I'm starting to realise that my personal information can be used against me'.

For example, AI is being used to make decisions that unfairly disadvantage people on the basis of their race, age, gender or other characteristic. This problem arises in high-stakes decision making, such as social security, policing and home loans.

These risks affect all of us, but not equally. We saw how new technologies are often 'beta tested' on vulnerable or disadvantaged members of our community.

So, how should we respond?

Australia should innovate in accordance with our liberal democratic values. We propose a *National Strategy on New and Emerging Technologies* that helps us seize the new economic and other opportunities, while guarding against the very real threats to equality and human rights.

The strongest community response related to AI. The Commission proposes three key goals:

1. AI should be used in ways that comply with human rights law
2. AI should be used in ways that minimise harm
3. AI should be accountable in how it is used.

Sometimes it's said that the world of new technology is unregulated space; that we need to dream up entirely new rules for this new era.

However, our laws apply to the use of AI, as they do in every other context. The challenge is that AI can cause old problems—like unlawful discrimination—to appear in new forms.

We propose modernising our regulatory approach for AI. We need to apply the foundational principles of our democracy, such as accountability and the rule of law, more effectively to the use and development of AI.

Where there are problematic gaps in the law, we propose targeted reform. We focus most on areas where the risk of harm is particularly high. For example, the use of facial recognition warrants a regulatory response that addresses legitimate community concern about our privacy and other rights.

Government should lead the way. The Discussion Paper proposes strengthening the accountability protections for how it uses AI to make decisions.

But the law cannot be the only answer. We set out a series of measures to help industry, researchers, civil society and government to work towards our collective goal of human-centred AI.

Education and training will be critical to how Australia transitions to a world that is increasingly powered by AI. The Discussion Paper makes a number of proposals in this area. We also propose the creation of a new AI Safety Commissioner to monitor the use of AI, coordinate and build capacity among regulators and other key bodies.

Finally, innovations like real-time live captioning and smart home assistants can improve the lives of people with disability. But as technology becomes essential for all aspects of life, we also heard how inaccessible technology can lock people with disability out of everything from education, to government services and even a job.

The Commission makes a number of proposals to ensure that products and services, especially those that use digital communications technologies, are designed inclusively.

We thank the many representatives of civil society, industry, academia and government, as well as concerned citizens, who participated in the Commission's first phase of public consultation. This input has been crucial in identifying problems and solutions.

We also pay tribute to the Commission's major partners in this Project: Australia's Department of Foreign Affairs and Trade; Herbert Smith Freehills; LexisNexis; and the University of Technology Sydney (UTS). In addition, we thank the Digital Transformation Agency and the World Economic Forum for their significant support. The Commission acknowledges the generosity of its Expert Reference Group, who provide strategic guidance and technical expertise.

The Commission sets out here a template for change, but it is written in pencil rather than ink. We warmly invite you to comment on the proposals and questions in this Discussion Paper. We will use this input to shape the Project's Final Report, to be released in 2020.

Edward Santow
Human Rights Commissioner

December 2019

PART A : INTRODUCTION AND FRAMEWORK

CHAPTER 1
Introduction

New technologies are emerging at extraordinary speed, with unprecedented social and economic consequences. Change is so pervasive that this era has been termed the 'Fourth Industrial Revolution'.[1]

The Human Rights and Technology Project (Project) considers the implications for our human rights, and how we should respond to protect and promote those rights.

Informed by extensive public consultation, this Discussion Paper uses a human rights approach to identify and analyse the challenges and opportunities for human rights protection and promotion in the context of new and emerging technologies. It considers what is needed in terms of regulation, governance and leadership, and makes concrete proposals to reform laws, policy and practice.

The Discussion Paper is published at a critical time for Australia. There are several inquiries and consultations currently being conducted, with a view to shaping Australia's response to the Fourth Industrial Revolution. There is a general consensus that this era presents risks and opportunities, and Australia needs to address both.[2]

The great promise of new and emerging technologies is likely to be realised only if there is social trust in their development and use. That in turn requires innovation that is consultative, inclusive and accountable, with robust safeguards for human rights protection and promotion. This accords with Australia's binding obligations under international human rights law to respect, protect and fulfil human rights across all areas of life.

The Discussion Paper is divided into five parts.

In each part, the Commission sets out proposals and questions to guide further public consultation in 2020.

1.1 Background to the Project

(a) Project overview

The Australian Human Rights Commission is Australia's national human rights institution. The Commission is independent and impartial. It aims to promote and protect human rights in Australia.

The Commission launched the Project on 24 July 2018 at an international human rights and technology conference. The Project has two phases of public consultation.[3] The first phase took place in the second half of 2018 and early 2019. It was guided by two documents: an Issues Paper, published on 24 July 2018; and a White Paper on Artificial Intelligence governance and leadership, co-authored with the World Economic Forum, published in January 2019.[4]

The second phase of consultation will be guided by this Discussion Paper, which sets out a roadmap for reform. Details of how to contribute to this phase of public consultation can be found in Chapter 13.

A Final Report is due in 2020, following which the Commission will work with the Australian Government and other stakeholders to support implementation of the report's final recommendations.

(b) Project in context: domestic and international processes

There are a number of concurrent processes considering how to make the most of the opportunities provided by new and emerging technologies, as well as minimising risk.

In Australia, there are processes examining how to support the growth of ethical AI, promote the digital economy, and provide for security in the design of Internet of Things (IoT) devices.[5]

At the transnational level, bodies such as the United Nations (UN), European Union, Organisation for Economic Cooperation and Development (OECD), G7 and G20 are all heavily involved in shaping an international response to new and emerging technologies.[6] Nation states, industry and civil society are all engaged in a similar task.[7]

Some overseas initiatives have applied international human rights law. The OECD Principles on AI, for example, which were also recently endorsed by the G20,[8] state that the development of AI technologies should be 'innovative, trustworthy and respect human rights and democratic values'.[9]

1.2 Public consultation

A human rights approach prioritises a broad, inclusive process of public consultation.

At the time of writing, the Commission is the only independent government body in Australia that has applied a human rights approach with an inclusive public consultation to understand and address the social impact of new and emerging technologies like AI. The Commission has also sought to collaborate with, and contribute to, concurrent inquiry processes both in Australia and overseas.

(a) Consultation methodology

The Commission's consultation process is designed to be participatory and inclusive.

The Commission established an Expert Reference Group for this Project.[10] Constituted by a diverse group of leading experts, the Expert Reference Group advises the Commission on the Project's strategy, consultation process and on its findings and proposals. All views are the Commission's, and the Commission is responsible for this Discussion Paper and other Project outputs and statements.

The Commission has consulted a wide range of other experts, stakeholders and the community. The Commission has received written submissions, and held roundtable meetings and interviews with key informants. This has involved representatives from industry, civil society, affected communities, academia, Australian and overseas governments and intergovernmental organisations.

The Commission is particularly interested in the experiences of those most affected by new technologies and will prioritise listening to these people in the next phase of consultation.

(b) First phase of public consultation

From July 2018 to March 2019, the Commission consulted on the questions raised in the Issues Paper and White Paper.

The Issues Paper asked questions related to Australia's regulatory framework for new and emerging technologies; the use of AI in decision making; and accessible technology for people with disability. The White Paper asked six questions regarding AI governance and leadership in Australia.

The Commission received 119 written submissions on the Issues Paper from civil society, academia, government and industry stakeholders.[11] Around 100 of these were from organisations, including member-based and representative organisations and those with specialist expertise.

The Commission received 63 written submissions in response to the White Paper from civil society, academia, government and industry stakeholders.

The Commission also conducted face-to-face roundtables. Around 380 invited stakeholders attended these roundtables in Sydney, Canberra, Melbourne, Brisbane, Perth and Newcastle. The stakeholders who attended were from civil society, academia, government and industry, particularly technology companies. The Human Rights Commissioner and Disability Discrimination Commissioner also held a dedicated roundtable for disability advocacy organisations.

A symposium on the White Paper questions, co-hosted with the World Economic Forum, was held at the University of Technology Sydney on 6 March 2019, with 65 experts and senior decision makers attending from across civil society, academia, government and industry.

Formal and informal interviews were held with human rights and technology experts, government officials and other key stakeholders.

Finally, the Human Rights Commissioner participated as a keynote speaker or panel member talking about the human rights impact of new technologies at 70 events reaching over 13,000 participants in Sydney, Brisbane, Melbourne, Newcastle, Perth, Canberra; and, in Canada, Italy, Singapore and the United States of America.

These events ranged from community forums, to large conferences and intergovernmental meetings, including the G7 Multi-stakeholder Conference on Artificial Intelligence in Montreal, and the Future of Human-Centered AI event hosted by the Stanford Centre on Democracy, Development and the Rule of Law.[12] Senior Project Team members have also participated as speakers in public forums.

(c) Second phase of consultation on the Discussion Paper

The Commission now invites stakeholders to comment on the proposals and questions in this Discussion Paper. This second phase of consultation will be the final public consultation phase in this Project.

In addition to written submissions, the Commission also intends to:

- hold public roundtables with stakeholders from industry, civil society, academia and government
- seek the views and experience of people who are particularly affected by new technologies, especially those from at-risk and vulnerable groups
- conduct interviews with experts and key decision-makers
- continue engaging with other review processes on areas related to this Project.

Details of how to comment on the Discussion Paper proposals and questions are set out in Chapter 13 below.

1.3 Other issues

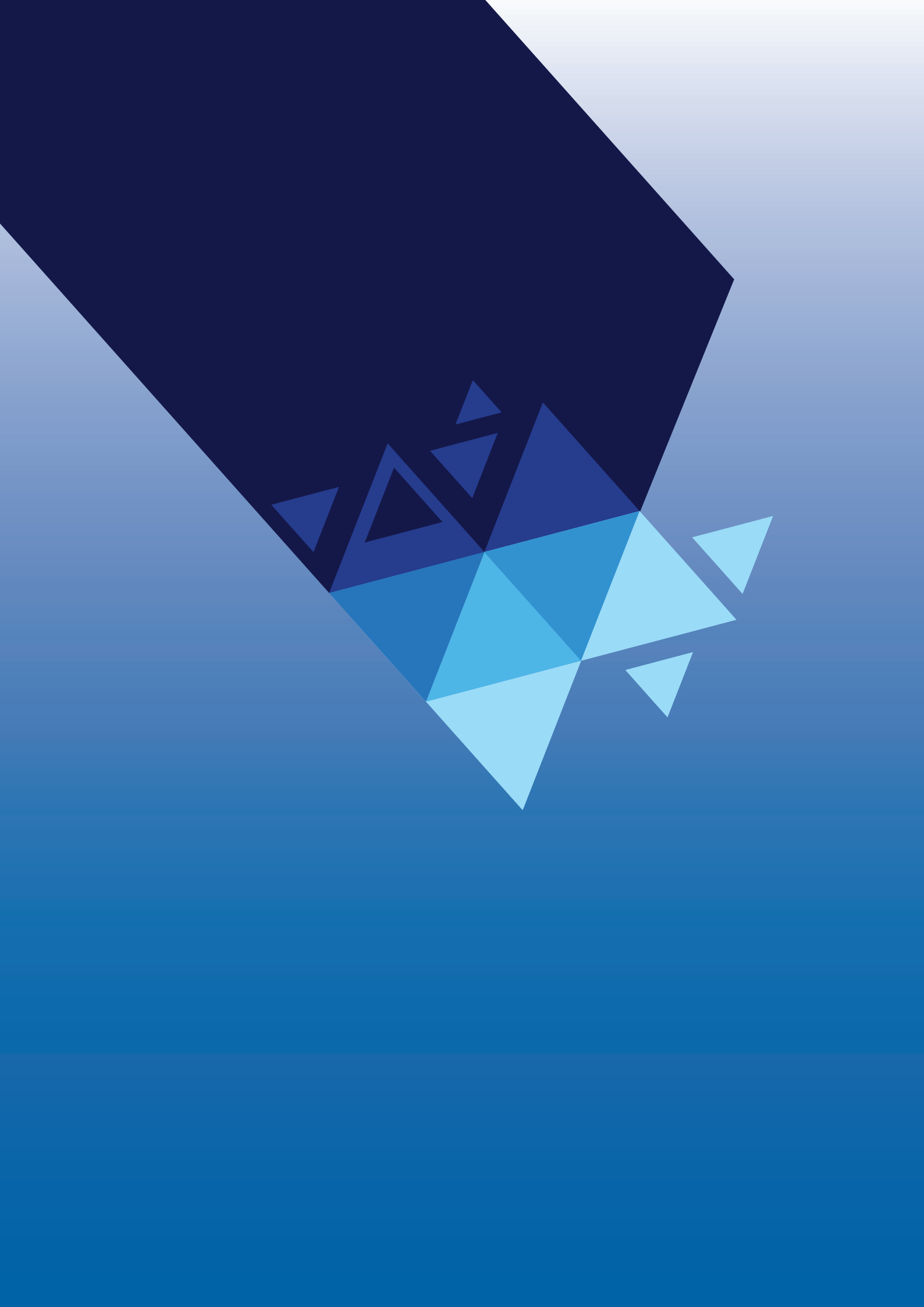
During the first phase of the Commission's consultation, stakeholders raised a significant number of human rights issues relating to new and emerging technologies, and identified the different impacts on different groups of people.

For the purposes of this Project, the Commission has chosen to undertake an in-depth analysis of a selection of these issues, recognising that other organisations are conducting relevant projects related to specific areas or interest groups.

Stakeholders were asked to comment on three substantive issues in the Issues Paper: regulation; accessibility; and AI informed decision-making. These issues have emerged as the most pressing and with the widest implications for human rights.

Nevertheless, the Commission emphasises that there are many other areas that would benefit from dedicated research, analysis and consultation. These include, in particular:

- **The future of work in Australia and the impact of automation on jobs**, including reviewing the impact on social inequality.[13]
- **The impact of connectivity**, including the IoT, 'smart cities', and the convergence of IoT development and data governance.[14]
- **Digital inclusion**, including exclusion from reliable internet access, due to such factors as poor infrastructure, socio-economic status and geographical location. This is a significant issue for many vulnerable and marginalised groups and was raised by a number of stakeholders. The issue is considered in relation to the specific impact on people with disability in Part D. It is also being considered by other bodies (see, for example, the *Australian Digital Inclusion Index Reports*).[15]
- **Technology-facilitated, gender-based violence and harassment, and online hate**, including the adverse consequences of technology, particularly for victims of domestic violence, and people from minority backgrounds.[16] Online safety, particularly for women and children, forms the focus of the work of the eSafety Commissioner.[17]
- **The regulation of social media content**, including the impact of social media platforms on democratic processes.[18] This issue is being considered by other bodies, both in Australia and overseas.[19]
- **Human-centred, or 'anthropological', questions about the implications of AI**, focusing on our current and future relationship with technology, the impact on our collective futures and the impact of smart technology such as intuitive robotics.
- **Digital literacy education** (including issues relating to AI) for adults and children.



CHAPTER 2
Human rights framework

2.1 Introduction

Human rights are set out in international and domestic laws. While the roots of the human rights movement can be traced to ancient philosophical writings on natural law, the modern human rights framework has its origins in the formation of the United Nations (UN) in 1945.[20]

At the core of the international human rights framework is the principle that 'all human beings are born free and equal in dignity and rights'.[21]

Put simply, human rights are about what it means to be human; to have autonomy, personal freedom and value regardless of ethnicity or race, religious belief, gender or other characteristics.

Human rights law accepts that human rights come into tension with each other, and with other interests. It provides ways of resolving these tensions to accommodate multiple rights and interests. Human rights law can help us answer the big questions posed by the rise of new technologies, including how to protect humans in a digital world.[22]

Stakeholders contributing to the Commission's consultation process have reported many ways in which new technologies engage our human rights, and they have been broadly supportive of focusing more on human rights to analyse the social impact of new technologies.[23]

Drawing on this broad stakeholder support, the Commission considers a human rights approach to analysing the impact of new and emerging technologies is vital in understanding and responding to the risks and opportunities.

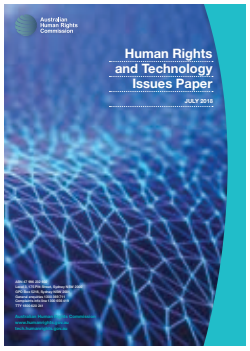
Applying a human rights framework starts with considering how new and emerging technologies affect humans. The potential human rights impact is enormous and unprecedented. AI, for example, can have far-reaching and irreversible consequences for how we protect privacy, how we combat discrimination and how we deliver health care—to name only three areas.

Human rights is already used as a source of law and an established set of norms, to analyse the impact of new technologies.[24] Support for a human rights approach is also increasingly reflected in the actions and statements of:

- United Nations organs, such as the UN High Commissioner for Human Rights and the Special Rapporteurs on privacy and freedom of expression[25]
- multilateral international bodies, such as the OECD[26] and G20[27]
- comparable liberal democracies, such as Canada and France[28]
- leading multinational corporations, such as Microsoft.[29]

This Discussion Paper responds to and builds on that growing body of work. In this chapter, the Commission:

- summarises how international human rights law operates generally, and in the specific context of the rise of new technologies
- applies a human rights approach to new technologies, with particular reference to some examples raised during the consultation process.



2.2 What are human rights?

We are all entitled to enjoy our human rights for one simple reason—that we are human. We possess human rights regardless of our background, age, gender, sexual orientation, political opinion, religious belief or other status.

Human rights are centred on the inherent dignity and value of each person, and they recognise humans' ability to make free choices about how to live.

Australia is a signatory to the nine core human rights treaties, which cover the full spectrum of human rights, including civil and political rights, and economic, social and cultural rights. Accordingly, Australia has voluntarily agreed to comply with human rights standards and to integrate them into domestic law, policy and practice.

Human rights are *universal*, meaning that they apply to everyone.

They are *indivisible*, meaning that all human rights have equal status.

They are *interdependent* and *interrelated*, meaning the improvement of one human right can facilitate the advancement of others. Likewise, the deprivation of one right can also negatively affect other human rights.

Accordingly, while there are sometimes complex inter-relationships between different rights, governments must ensure everyone's human rights are protected.[30]

Table 1 provides an overview of how human rights and new technology can intersect. These examples show that new technologies can advance or restrict one or more human rights, and sometimes offer both possibilities at once.

Table 1: Overview of technology advancing and restricting human rights[31]

Right to equality and non-discrimination

Applications that use AI, and especially machine learning, must be 'trained' using data. Where that data incorporates unfairness, such as discrimination, this can be replicated in the new application.

Where training data is collected and used well, new technologies such as AI can enable better service delivery, especially for vulnerable groups.

Unequal access to critical new technologies can exacerbate inequalities, especially where access is affected by factors such as socio-economic status, disability, age or geographical location.

Freedom of expression

New technologies can enable wide-scale surveillance online and in the physical world, which can deter individuals from legitimately sharing their views.

New technologies can aid freedom of expression by opening up new forms of communication.

New technologies can assist vulnerable groups by enabling new ways of documenting and communicating human rights abuses.

Hate speech, 'fake news' and propaganda can be more readily disseminated.

Right to benefit from scientific progress

New technologies can improve enjoyment of human rights such as access to food, health and education.

Ensuring accessibility across all sectors of the community can be difficult.

Freedom from violence, harassment and exploitation

Access to new technologies can provide greater protections from violence and abuse, and the ability to document abuse, but can also facilitate other forms of abuse (such as image-based abuse or covert surveillance).

Greater access to information and support through technology can make support for survivors of violence and abuse more affordable and accessible.

Accessibility

New technologies can provide new ways to deliver services, thereby increasing accessibility for people with disability and others.

Reduced cost of services through affordability of new technology can promote equality for people with disability by ensuring progressive realisation is achieved faster and reasonable adjustments are more affordable.

New technologies can increase barriers for people with disability if they are used in products and services in ways that are not accessible.

Protecting the community and national security

New technologies can increase government's capability to identify threats to national security.

Use of such technologies for surveillance purposes can be overly broad and, without appropriate safeguards, can impinge unreasonably on the privacy and reputation of innocent people.

Right to privacy

The ease of collecting and using personal information through new technologies such as facial recognition can limit the right to privacy.

Personal data can flow easily and quickly, across national and other borders. This can make privacy regulation and enforcement more difficult.

It can be difficult to 'correct' or remove personal information once communicated.

The ease of communicating and distorting personal information (eg, through 'deep fakes') can lead to reputational damage and other harms.

Right to education

New technologies can improve the availability and accessibility of education.

Lack of access to technology can exacerbate inequality, based on factors such as age, disability, Indigenous status, and rural or remote location.

Access to information and safety for children

Online environments provide children with the opportunity to access a wealth of information, but also pose challenges for their wellbeing.

New technologies create different settings for harassment and bullying that are sometimes challenging to moderate.

Digital technology can also facilitate the exploitation of children.

(a) Obligations of states to protect human rights

International human rights law requires Nation States to respect, protect and fulfil human rights.[32] In particular:

- The obligation to **respect** means that Nation States must refrain from interfering with or curtailing the enjoyment of human rights. In other words, governments themselves must not breach human rights.
- The obligation to **protect** requires Nation States to protect individuals and groups against human rights abuses. In other words, laws and other processes must protect against breaches of human rights by others, including non-state actors.
- The obligation to **fulfil** means that Nation States must take positive action to facilitate the enjoyment of human rights.

As noted above, Nation States must recognise that the breach of one human right might affect another; that is, all human rights are 'indivisible and interdependent and interrelated'.

(b) How are human rights protected in Australia?

In order to be fully enforceable in Australia, international human rights law must be incorporated into domestic Australian law through legislation, policy and other arrangements.

Where international law is so incorporated, this creates rights, obligations and accountability mechanisms *under Australian law*, which apply to individuals, public and private organisations.

Human rights are protected in Australia in a number of ways including:

- **in Australia's Constitution**—Australia has no federal bill of rights, but a small number of rights are protected directly or indirectly in the Australian Constitution—most particularly, the right to freedom of political communication.[33]

- **in legislation**—there are protections for certain human rights in Australia's federal, state and territory laws, including laws relating to privacy, discrimination, criminal activity and social security law. Victoria, Queensland and the Australian Capital Territory have enacted more general human rights acts.[34]
- **by the common law**—sometimes known as judge-made law, the common law protects a range of human rights, such as the right to due process or procedural fairness, which aims to ensure people receive a fair hearing.[35] There is also a well-established common law principle of statutory interpretation that Parliament is presumed not to intend to limit fundamental rights, unless it indicates this intention in clear terms.[36]
- **by specialist government bodies**—such as the Commission. The Commission has special responsibility for protecting human rights in Australia, including through a conciliation function (in respect of alleged breaches of federal human rights and anti-discrimination law), education and policy development.[37] Other specialist bodies also have regulatory and broader functions in respect of specific rights, such as the Office of the Australian Information Commissioner, which is responsible for privacy and freedom of information.
- **through Australian democratic processes**—in particular, when a law is introduced into Parliament, the relevant member of Parliament (usually a government Minister) must also present a statement of compatibility, which considers how the draft law engages human rights. The Parliamentary Joint Committee on Human Rights scrutinises draft laws, including by reference to statements of compatibility, in order to advise the Australian Parliament on whether they are consistent with international human rights law.[38]
- **by participating in UN review processes**—the Australian Government participates in the various review processes conducted by UN bodies that report on Australia's compliance with its human rights obligations. Individuals are also able to make a complaint to certain UN bodies, on the ground that the Australian Government is in breach of its obligations under one of its treaty commitments. In addition, the UN appoints special rapporteurs and other mandate holders to report on human rights conditions in countries including Australia.

(c) Safeguarding human rights

Under international human rights law, governments can limit or restrict enjoyment of most human rights.[39] There are some 'absolute' rights that can never be limited or restricted. These include the right to life, the right to be free from torture, freedom from slavery and servitude, and the right to recognition before the law.[40]

As the vast majority of human rights are not absolute, international law has developed principles for how a human right may be limited or restricted.

Generally speaking, a limitation on a human right can be contemplated only if this pursues a legitimate aim (eg, protecting national security). The limitation must be necessary, reasonable and proportionate in pursuing that legitimate aim.[41] It must involve the least restrictive limitation on another human right that is possible. This framework also can be applied to resolve any apparent tension between different human rights in a particular situation.

Some specific requirements also apply in the context of limiting particular human rights. For example, any limitation on freedom of expression must be necessary to 'respect the rights or reputation of others', or to protect national security, public order, public health or morals.[42]

Some rights are 'derogable'. A government may temporarily suspend the application of such a right in an exceptional situation, such as an emergency.[43] Examples of derogable rights include the right to liberty of movement,[44] and the right to security of person and freedom from arbitrary detention.[45] Significantly, any measures by governments to derogate from these rights cannot involve discrimination solely on the ground of race, colour, sex, language, religion or social origin.[46]

Applying a human rights analysis to new and emerging technologies has two advantages.

The first, and most obvious, advantage is that it promotes compliance with international and domestic law. Secondly, this analysis provides a principled basis to consider the claimed benefits of the development or use of a particular technology, as against any impingement on human rights.

For non-absolute rights, such as the right to privacy, human rights law has developed mechanisms to ensure that rights can be effectively accommodated when other legitimate aims are pursued. For example, by asking whether an impingement on privacy is proportionate to the legitimate aim being pursued.

Courts have begun to grapple with this type of question in the context of new and emerging technologies.[47] One such example is set out in Box 1 below.

Box 1:

When might limiting a human right be lawful?

Catt v the United Kingdom (European Court of Human Rights, Application No 43514/15, 24 January 2019)

The applicant, Mr Catt, had been active in peace movements and public demonstrations in the United Kingdom. In March 2010, Mr Catt made a request to the police under the UK's *Data Protection Act 1998* for any information held by the police relating to him. The information in the police database concerned 'domestic extremism' and was contained in records of other individuals. Mr Catt requested the police delete these entries, but the police refused.

The European Court of Human Rights affirmed that the collection of personal data to prevent crime and disorder was a lawful and legitimate purpose.[48] However, the Court found that in the present case, the retention of personal data by the police without scheduled review and beyond established limits was disproportionate and unnecessary.[49] The Court held that Mr Catt's right to respect for privacy and family life had been violated, considering the nature of the personal data, Mr Catt's age and the fact he had no history or prospect of committing acts of violence.[50]

2.3 How are human rights engaged by new and emerging technologies?

While public discussion of the implications of new technologies has tended to focus on the right to privacy and non-discrimination, other human rights also are engaged.

As noted by one submission, in the context of large scale data analytics in particular, there is significant risk that government and companies will 'lose sight of the fundamental dignity of human beings when datasets flow far beyond the initial point of collection, treating human beings as data points rather than individuals'.[51]

Other submissions identified various positive and negative ways in which human rights can be engaged by new and emerging technologies. Some of these examples are discussed here.

(a) Different experience of technologies within the community

The rapid and widespread application of new technologies is impacting a range of human rights. This impact is not uniformly experienced across the community, or even within communities.

Research indicates that individuals' experience of new technologies can differ in at least three ways:

- groups that already experience higher levels of discrimination or socio-economic disadvantage before a new technology is implemented are more vulnerable to experiencing negative human rights impacts from that technology
- groups that were not experiencing significant levels of discrimination or socio-economic disadvantage were more likely to be positively affected by a new technology
- the same technology, and even the same product or service, can have a positive human rights impact for some people and a negative impact for others.[52]

There is also a diversity of experiences both within and between groups of people sharing a characteristic such as race, gender, disability status, age or socio-economic status. This can be affected by a range of factors such as digital literacy.

For people who experience multiple forms of disadvantage, the negative experience of new technologies can be compounded. For example, an Aboriginal person with a disability may experience intersectional disadvantage if this person faces barriers to accessing technology by reason of being Aboriginal and further barriers because of their disability. This can exacerbate existing exclusion and discrimination.[53]

(b) Protection and promotion of human rights

New and emerging technologies can improve human rights protections and create new opportunities for human rights promotion in diverse contexts and settings. Examples that were identified during the Commission's consultation include the following:

- Many stakeholders referred to the use of new and emerging technologies to support advocacy efforts, improve access to information and the ability to document and communicate human rights violations more effectively.[54] Digital communications platforms, including social media, can enable like-minded people to work together on such issues and can facilitate peaceful protest.

  This, in turn, can support freedom of expression,[55] which is predicated on the ability to exchange ideas and opinions freely. It can also advance the right to freedom of association.[56]

- As set out in greater detail in Chapter 6, where designed and used appropriately, AI can promote diversity and reduce bias and discrimination, especially in the workforce. For example, AI has been used to reduce human bias and prejudice in recruitment. One company reported that its recruitment tool uses data-driven trait analysis to match a person based on their skills rather than personal or demographic information.[57] This aims to eliminate some of the biases associated with more conventional recruitment methods, which focus heavily on assessing the applications of job candidates.

  If such AI-powered tools can successfully achieve their aims, by reducing or eliminating discrimination in recruitment on grounds such as race, gender or disability, this would promote the right to equality and non-discrimination.[58] It would also promote the right to work.[59]

- AI and related technologies can be used to advance the right to health care through, for example, more accurate diagnosis of some forms of cancer.[60] One stakeholder identified how medical technologies with improved diagnostic capabilities could help close the health gap for Aboriginal and Torres Strait Islander people.[61] Similarly, geospatial imagery and sensors are starting to be used to develop evidence-based policy for healthy urban environments.[62]

  These examples engage the right to 'the highest attainable standard of health conducive to living a life in dignity' and the right to scientific benefit.[63] The attainment of the right to health is indispensable to the fulfilment of many other human rights such as the right to education, work, non-discrimination, equality, privacy and life.[64] Where new technologies can be used to address particular health-related disadvantage, such as that experienced by some Aboriginal and Torres Strait Islander people, this would advance key elements of the Declaration on the Rights of Indigenous Peoples and other similar international instruments.[65]

(c) Adverse impacts on human rights

Stakeholders also identified how new and emerging technologies can adversely affect various human rights. Examples include the following.

- Unequal access to digital technologies can entrench existing inequality. People with disability, for example, will be excluded from a range of online services if they are inaccessible, such as health information websites that cannot be read by screenreaders, used by people with a vision impairment.[66]

  This limits human rights such as the right to non-discrimination and equality, the right to health, and more specific rights under the Convention on the Rights of Persons with Disabilities, including the right of people with disability to 'access aspects of society on an equal basis' and the right to live independently in the community.[67]

- As explored in greater detail in Chapter 6, recruitment tools that use machine learning may 'learn' from a historic data set that applicants with a shared characteristic (such as gender, age or race) are a preferred applicant, resulting in skewed recruitment that may directly or indirectly discriminate.[68] For instance, Amazon reportedly stopped using an AI-powered recruitment tool once it became apparent during testing that it was skewing its recruitment disproportionately towards male applicants.[69]

  Similarly, targeted advertisements for jobs or housing may exclude groups with a shared characteristic. An historic settlement with Facebook early in 2019 led to a change in company guidelines, so it will no longer be possible for advertisers to select who sees their advertisements on the basis of their age or race.[70]

  Such use of technology can limit the rights to non-discrimination and equality, to privacy, and to work.

- Automated risk assessment tools, many of which use AI, are increasingly being used in the criminal justice context. Such tools can be used to target policing resources to high-risk areas or towards potential offenders, and to assist judges in bail and sentencing decisions. They can have a disproportionate, negative impact on groups, such as young people and Aboriginal and Torres Strait Islander people.[71] There is also a growing body of research that suggests these can entrench existing and historic inequality in the criminal justice context.[72]

  Where these have these unintended effects, they can infringe human rights including the rights to a fair trial and to be free from arbitrary arrest, privacy, non-discrimination and equality. Where these tools unfairly disadvantage particularly vulnerable groups, based on characteristics such as their race or being children, this can engage further human rights.

2.4 Applying a human rights approach

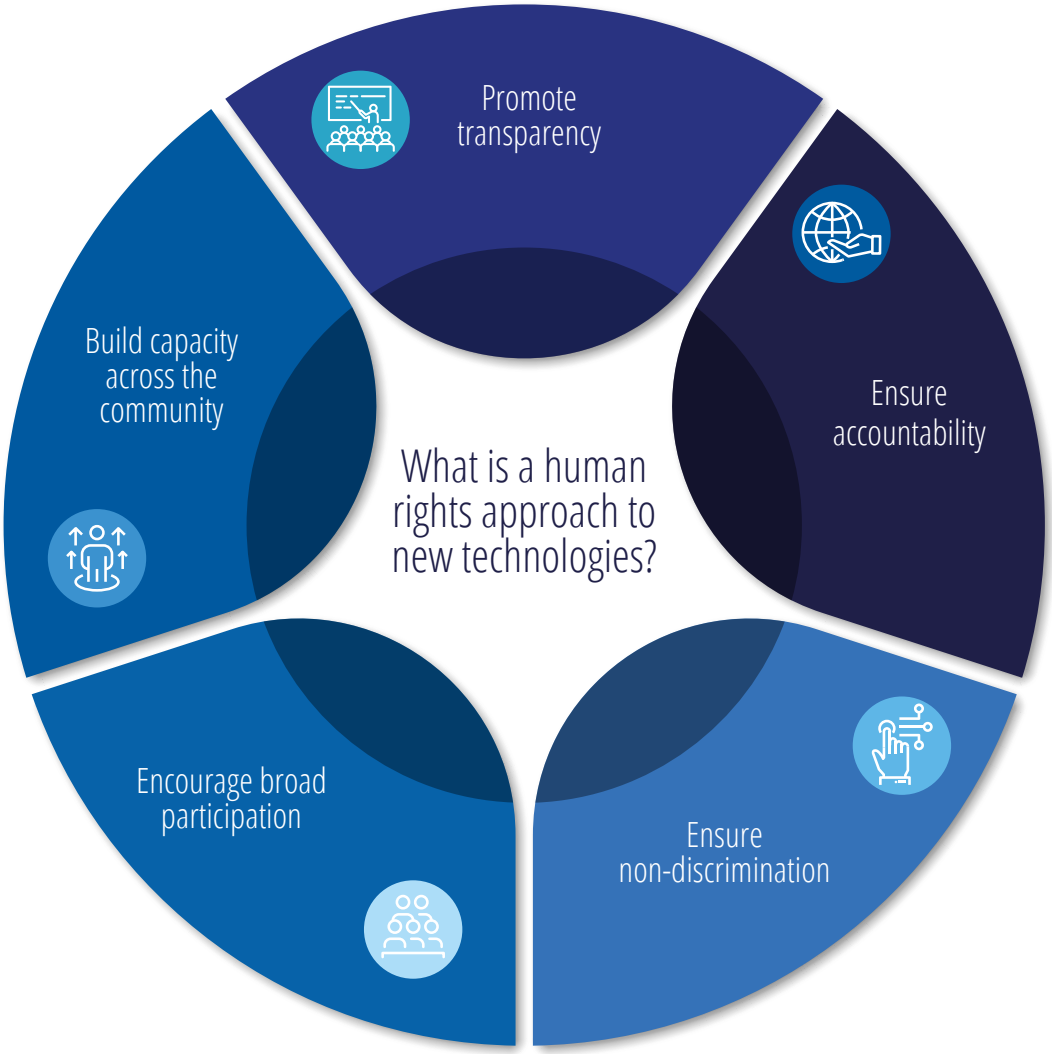
A human rights approach to the development and use of new technologies is increasingly common internationally,[73] with a growing number of experts emphasising the importance of human rights law in analysing the social impacts of new and emerging technologies.[74]

Some international initiatives use human rights as one of a number of 'lenses' through which to view the development and use of new technologies.[75] Given that the technologies are new, human rights law jurisprudence in this area is still developing.[76]

(a) What is a human rights approach to new technologies?

Adopting a human rights approach means building human rights into all aspects of policy development and decision making. Some key features of a human rights approach include that it will:

- **Promote transparency in government decision making**. Governments should consider the impact of policy and governance for new technologies on people's human rights, and build in safeguards for human rights protection.
- **Ensure accountability of:**
  - **government,** by putting in place measures to influence government use of new technologies, such as through procurement rules, and clearly defining state and non-state obligations in a human rights compliant regulatory framework
  - **non-state actors,** with responsibility and liability defined by the government's framework. This should also promote business respect for human rights in accordance with the United Nations Guiding Principles on Business and Human Rights (UN Guiding Principles), discussed in further detail below.
- **Ensure that new technologies are non-discriminatory in their application**. For example, anti-discrimination law principles should be applied to the development and use of new and emerging technologies. It also means taking into account the specific needs of people who are vulnerable, such as people with disability who may face exclusion by inaccessible technologies.
- **Require participatory approaches**. This aims to ensure the voices of all stakeholders impacted by new and emerging technologies are heard, including the general public, particularly affected groups and civil society, as well as experts and decision-makers.
- **Build capacity across the community**. The community needs to understand the impact of new technologies on their lives, such as the impact of automation in decision-making processes for essential services, and have knowledge of, and access to, a review process and/or remedy.[77]



(b) Stakeholder support for a human rights approach to new technologies

A number of stakeholders expressed support for the application of a human rights approach to the issues raised by new and emerging technology.[78] The Australian Centre for Health Law Research, for example, noted:

> The fundamental principles underpinning all human rights are respect for the dignity, autonomy and equality of every individual. These principles ought to guide the design, development and implementation of new technologies, helping to embed a human rights-based approach to technology regulation.[79]

Similarly, Kingsford Legal Centre commented:

> Technology companies that pride themselves on innovation and 'disruptive technologies' often rally behind the motto of 'move fast and break things'. This approach, combined with the concerning issue of workforce diversity in Silicon Valley, threatens a human rights approach to technology. A systemic lack of diversity and unconscious bias may lead to a lack of due consideration of key issues such as equity and respect for human rights in the design of technologies.[80]

The Castan Centre for Human Rights Law also commented on the benefit of using human rights law to challenge 'power asymmetries' in the development of AI:

> Many of the risks relating to AI concern the entrenchment of the disadvantage experienced by marginalised and vulnerable groups. A human rights analysis requires the identification of duty bearers, and empowers rights holders with helpful analytical, normative and institutional tools to hold them to account.[81]

The University of Technology Sydney advocated that a human rights approach be supplemented with a transdisciplinary approach that takes into account the interaction between technologies and can take into account systemic influences.[82]

(c) Case study: A human rights approach to facial recognition technology

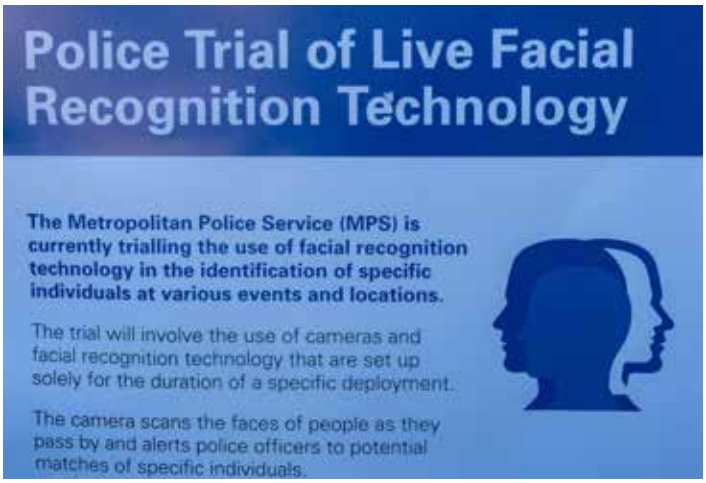
(i) How facial recognition technology is used

The deployment of facial recognition technology in public places was raised in a number of submissions.[83] Facial recognition technology has already been deployed by governments and non-government organisations in Australia.[84] Its use overseas, most notably as part of 'social credit score' systems in China, shows how this technology can enable mass surveillance.[85]

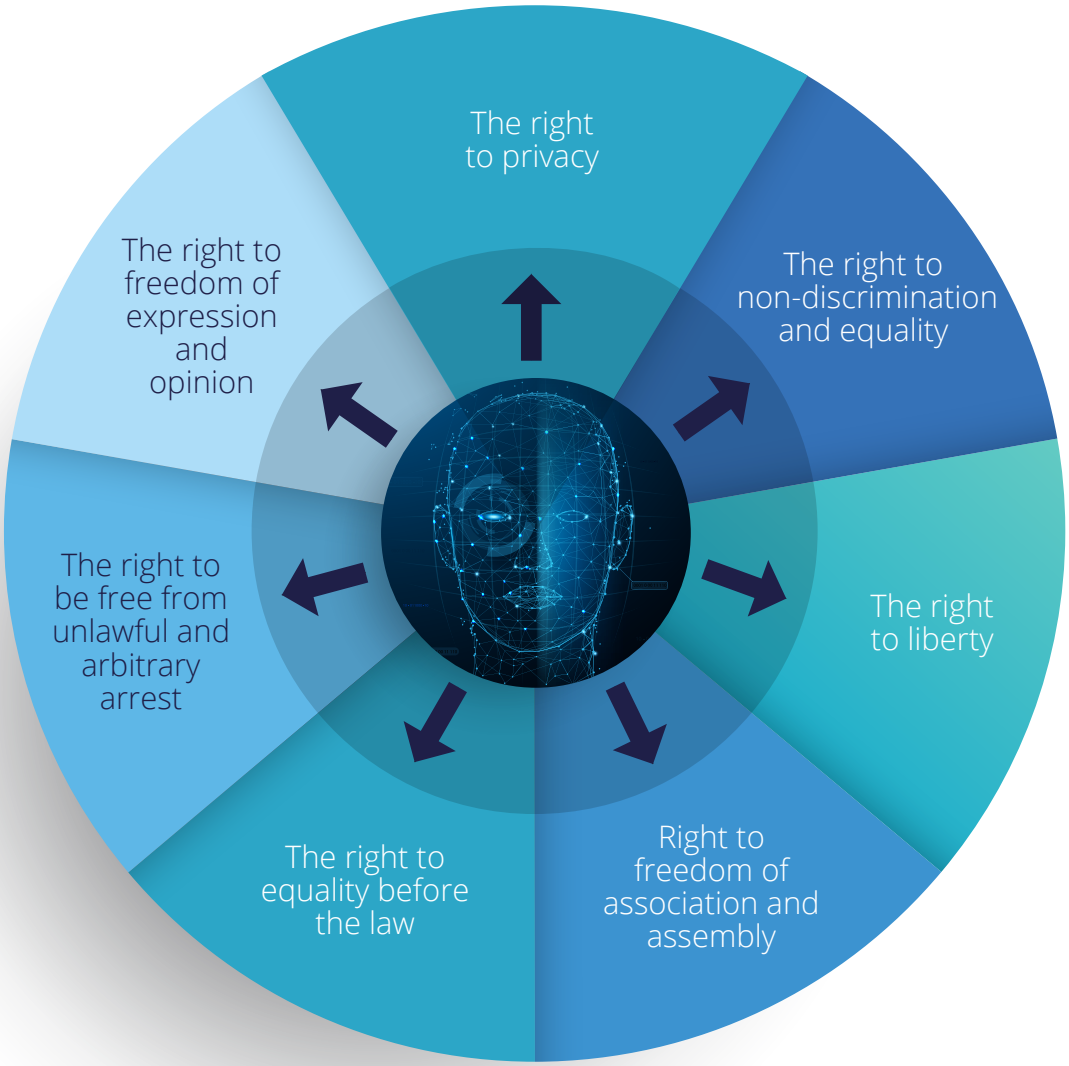
Use of facial recognition technology is growing in Australia. Perhaps most significantly, the Australian Government introduced the Identity-matching Services Bill 2019 (the Bill), which would establish a 'National Facial Biometric Matching Capability'.[86] If passed, this law would enable both one-to-one and one-to-many applications of the technology, and it would support a centralised database of passport, visa, citizenship and driver's licence photos.[87] The Commission has analysed the Bill in detail as part of the parliamentary scrutiny process,[88] and the Parliamentary Joint Committee on Intelligence and Security recently expressed strong concerns about this Bill.[89]

(ii) What human rights are engaged by use of facial recognition technologies?

Biometric identification technologies, including facial recognition, are extremely powerful tools. Examples of biometric technologies are wide and varied, ranging from fingerprint scanning and voice recognition to keystroke dynamics and body odour measurement.[90] Such technologies can be used in ways that impinge significantly on human rights.[91]



From 2016 to 2019, the Metropolitan Police Service in the United Kingdom trialled live facial recognition technology during police operations.



The human rights engaged by facial recognition technology will differ depending on the context in which it is deployed. For example:

- Biometric data that is collected from an individual in one setting and for one purpose may be collected and merged with personal data from other surveillance mechanisms such as drone footage, satellite imagery and encrypted communications. This type of surveillance will affect the right to privacy, and may engage other rights such as the right to non-discrimination and the right to liberty.
- Where a person is under surveillance in certain contexts, they may fear potential consequences from participating in lawful democratic processes such as protests and meetings with individuals or organisations, including increased surveillance or scrutiny by police.[92] This engages the right to freedom of association and assembly and freedom of expression and opinion, as well as, potentially, the right to be free from unlawful and arbitrary arrest.

- Merged data of this kind may be used to draw inferences about an individual which are shared with third parties, without any meaningful consent from the affected individual.[93] Sensitive personal information may be extracted or inferred from biometric identifiers, including in relation to the person's age, race, sex and health. This can be used to undertake 'profiling' activity, where intrusive action is undertaken by reference to people's age, race, sex or other characteristics. Examples of this kind of profiling, based on race, have emerged in other jurisdictions.[94] This engages the right to non-discrimination or equality. Depending on the context in which the technology is deployed, it is also likely to engage other human rights as well. For example, if such technology is used for racial profiling in police identity and other checks,[95] this could also limit rights such as equality before the law.
- There is also emerging evidence that facial recognition technology generally is less accurate when identifying women and people from minority ethnic and racial groups.[96]

(iii) Protection for human rights

There is growing public concern about the human rights impact of the use of facial recognition technology by government, especially where it leads to large-scale surveillance. Several American jurisdictions, for example, have passed or are considering laws banning use of facial recognition software where there is potential for harm.[97]

This Discussion Paper considers how Australia should respond in law, and more broadly, to these challenges in Chapter 6.

2.5 Accountability and responsibility

Not only must governments adhere to human rights in their own activities, governments also must establish a regulatory framework requiring non-state actors (such as companies) to adhere to human rights.

The human right to security of the person, for example, requires states to pass legislation that criminalises murder.

In the context of new and emerging technologies, the traditional lines between public and private accountability are becoming increasingly difficult to navigate. Private companies are developing new forms of technology that can have significant positive and negative human rights impacts. Companies outside the technology sector, and even governments, are integrating these new developments into their products and services. Many of these new technologies rely on personal information, with large databases of personal information now held outside government, including a small number of unprecedentedly large holdings.[98]

The challenge of assigning accountability, liability and responsibility for human rights protection in this context was identified by several stakeholders.[99] One answer to this challenge lies in the evolution of the international framework in business and human rights.

(a) UN Guiding Principles on Business and Human Rights

Since 2011, the UN Guiding Principles have provided an authoritative global standard for preventing and addressing the risk of adverse human impacts linked to business activity.[100]

The UN Guiding Principles set out the respective roles of government and businesses in protecting against the adverse impact on human rights in the context of business activities, and give guidance to businesses about steps to ensure they respect human rights.[101] The Guiding Principles are being incorporated into laws and soft law standards for specific business sectors—for example, legislation to combat 'modern slavery'.[102]

The UN Guiding Principles clarify that states have a duty to protect against human rights abuses by businesses in their territory or jurisdiction (which may include companies operating overseas), including through laws, regulation, policy and adjudication,[103] and recommend that businesses:

- respect human rights, which involves avoiding infringing human rights, and addressing any adverse human rights impacts, including by remediation. Businesses should also seek to 'mitigate adverse human rights impacts directly linked to their operations, products or services by their business relationships, even if they have not contributed to those impacts'[104]
- make a policy commitment to human rights protection, carry out human rights due diligence processes to identify, mitigate and account for the business' human rights impacts, and have processes to enable remediation of adverse human rights impacts caused by them or to which they have contributed[105]
- establish grievance mechanisms to facilitate access to an effective remedy for victims of adverse human rights impacts.[106]

Submissions identified the importance of the UN Guiding Principles in addressing the regulatory challenges posed by new technologies.[107] The Law Council of Australia, for example, recommended that the Australian Government develop guidance for businesses regarding effective human rights due diligence in accordance with the principles.[108]

(b) UN Guiding Principles and new technologies

International bodies have begun to apply the UN Guiding Principles to the development and use of new technologies.[109] The UN High Commissioner for Human Rights recently announced that her office is consulting with global stakeholders to develop guidance on this issue.[110]

Human rights law—and the UN Guiding Principles specifically—assist in answering some of the challenging questions raised by the social impact of new technologies.

For example, in 2018, the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, applied the UN Guiding Principles to the context of AI.

His report called on States to ensure that human rights are central to private sector design, deployment and implementation of AI systems.[111]

The Special Rapporteur described the approach businesses should take in the development, deployment and use of AI, in accordance with the UN Guiding Principles:

> Human rights law gives companies the tools to articulate and develop policies and processes that respect democratic norms and counter authoritarian demands. This approach begins with rules rooted in rights, continues with rigorous human rights impact assessments for product and policy development, and moves through operations with ongoing assessment, reassessment and meaningful public and civil society consultation.[112]

Submissions identified the significant challenges of responsibility associated with new technologies.[113] For example, who is responsible for the use of technology that adversely affects human rights, when the designer of the technology did not foresee that use?

Experts are starting to respond to this question by reference to the UN Guiding Principles. For instance, Business for Social Responsibility, a global non-government organisation (NGO), stated:

> The [UN Guiding Principles] offer an elegantly simple answer. Every enterprise in the value chain has responsibility to exercise due diligence and use their leverage to avoid negative outcomes for people, regardless of whether that business is designing, developing, marketing, selling, buying, using, or somehow benefiting from that technology. The [UN Guiding Principles] do differentiate between what different actors are expected to do based on the degree of their involvement in the adverse impact, but the underlying premise does not waiver.[114]

As noted above, the UN Guiding Principles also offer guidance on establishing effective grievance mechanisms.

Principle 31 provides that to be effective, non-judicial grievance mechanisms (such as company grievance mechanisms, or other state-based grievance mechanisms) should be, among other things: legitimate, in the sense of being trusted by stakeholder groups; accessible and known to affected stakeholder groups; predictable, providing a clear and known procedure and timeframe for resolution; and transparent.

Facebook is a technology company that is currently laying the foundations for an Oversight Board which will provide a new way to appeal Facebook's content decisions.[115] The effectiveness of company grievance mechanisms should be judged against the criteria outlined in Principle 31 of the UN Guiding Principles.

2.6 The Commission's preliminary view

(a) A human rights approach

International human rights law offers the most widely accepted framework for protecting individual dignity and promoting the flourishing of communities. These human rights norms are the core values agreed to by the international community, and by individual Nation States such as Australia.

In the following chapters, the Commission outlines how new and emerging technologies are having an impact on human rights protection and promotion, with particular focus on AI, and the impact of digital communications technologies on people with disability. Given the pace of technological change, it will be a significant challenge to ensure our regulatory system provides effective accountability where technology is used in ways that infringe human rights.

Community trust in new and emerging technologies has been decreasing.[116] Building confidence that our human rights are protected in this new era will be important in creating an environment that supports responsible technological innovation and the growth of Australia's digital economy.[117]

(b) Adopting an international approach

Stakeholders stressed the importance of connecting Australian and international approaches to new and emerging technologies.[118]

New and emerging technologies are transnational, in the sense that they are developed and operate across national boundaries. Nation States, such as Australia, must respond to this phenomenon when crafting policy and regulation in this area.

In Australia, we may also be affected by regulation from elsewhere.

The European Union's General Data Protection Regulation, for example, applies to any business that processes the personal data of an EU citizen, regardless of where that business is domiciled.

In the following chapters, the Commission considers several regulatory initiatives overseas, which seek to meet the challenges posed by new and emerging technologies. The Commission considers that Australian policy development should be underpinned by engagement in relevant international processes.

Specifically, the Commission urges the Australian Government to take an active role in promoting an international rules-based system, in a way that embeds human rights protections in international approaches to new and emerging technologies.

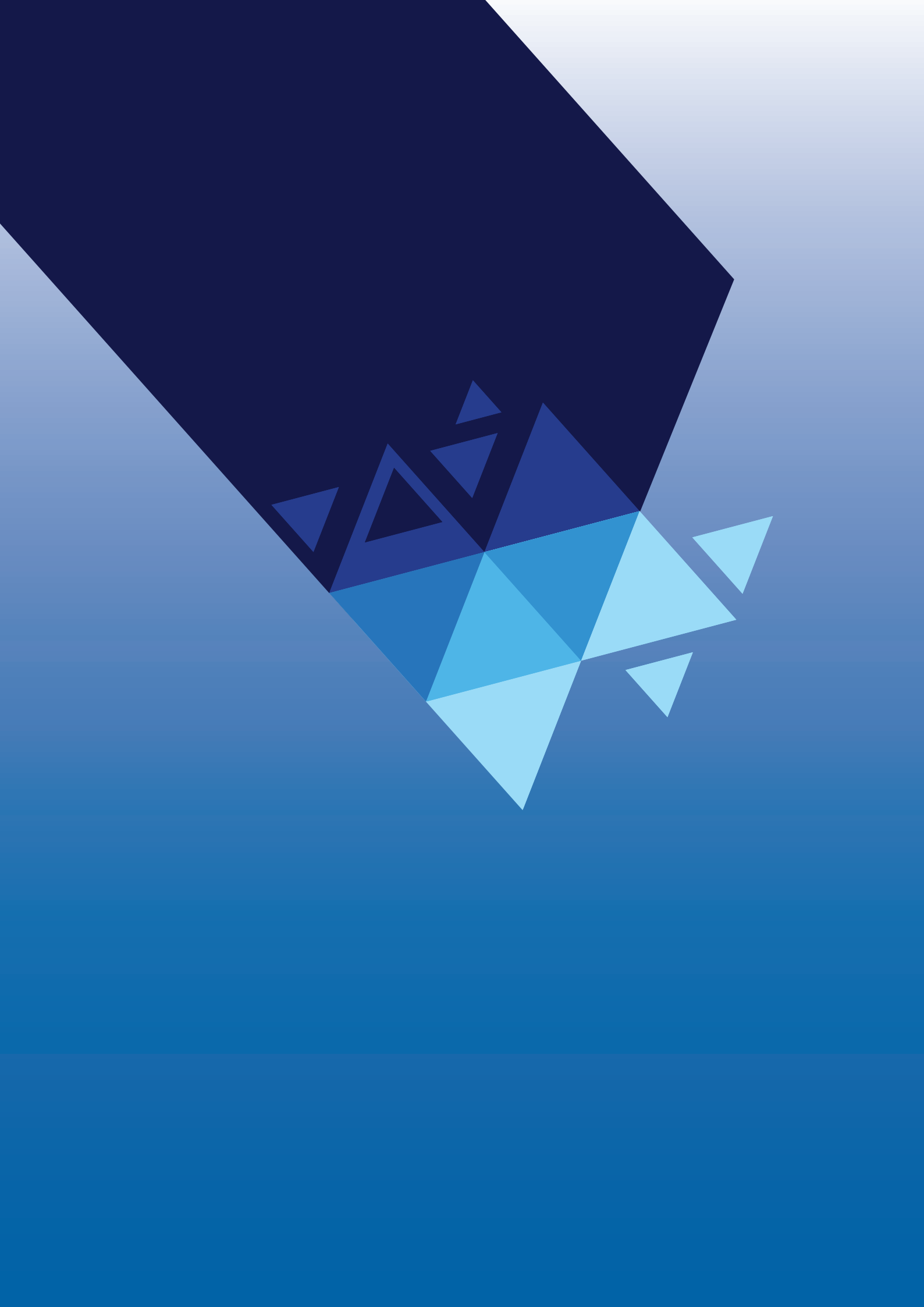
The Commission commends the Australian Government for progress it has already made in this regard. In May 2019, for example, the Australian Government signed the OECD's voluntary guiding principles for the design, development and use of AI, including to protect human rights, along with 41 other nations.[119]

In addition, the Commission notes that the UN High Commissioner for Human Rights commenced a major project on human rights and technology in June 2019.[120] An expert panel appointed by the Human Rights Council has been asked to consider how technology companies should incorporate well-established human rights principles into 'workable company practices'.

The High Commissioner noted the fragmented nature of self-regulatory approaches adopted by technology companies, and the comparative benefit of the global business and human rights legal framework to guide the public and private sector response to the opportunities and risks posed by new and emerging technologies.[121]

The UN Committee on the Rights of the Child is also currently drafting a General Comment on children's rights in relation to the digital environment. This Comment will offer guidance to governments on how to realise children's rights in the context of digital technology.[122]

The Commission will continue to engage through its consultation process in relevant regional and international processes.



CHAPTER 3
Regulation

3.1 Introduction

This chapter discusses regulation of the development and use of new and emerging technologies in Australia.

In its Issues Paper, the Commission invited comment on how Australian law and practice should protect human rights in the development, use and application of new technologies. Feedback was sought on gaps in the law, international examples of good practice and the principles that should guide regulation in this context. The Commission also asked how the Australian Government, the private sector and others should protect and promote human rights in the development of new technology.

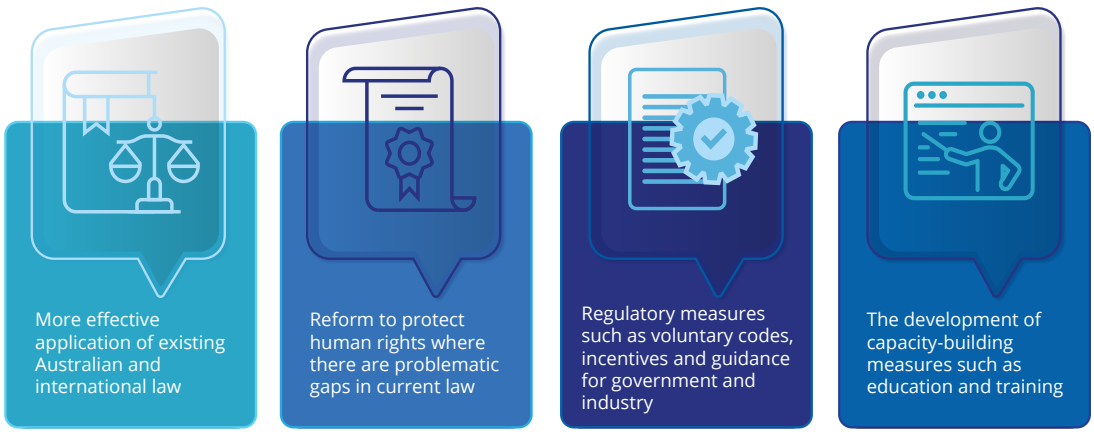
The Commission proposes that the Australian Government develop a national strategy for the protection of human rights in the development and use of new and emerging technologies. This national strategy should set a multi-faceted regulatory approach—including law, co-regulation and self-regulation—that protects human rights while also fostering technological innovation.[123]

In particular, the Commission proposes:

- more effective application of existing Australian and international law
- reform to protect human rights where there are problematic gaps in current law
- co- and self-regulatory measures such as voluntary codes, incentives and guidance for government and industry
- the development of capacity-building measures such as education and training.

This approach accords with international human rights law, including the UN Guiding Principles, which urge states to introduce a mix of regulatory measures to foster business respect for human rights.[124]

Some of the general principles discussed in this chapter are applied in greater detail in later parts of this Discussion Paper dealing with AI and accessible technology.



3.2 Regulation and new technologies

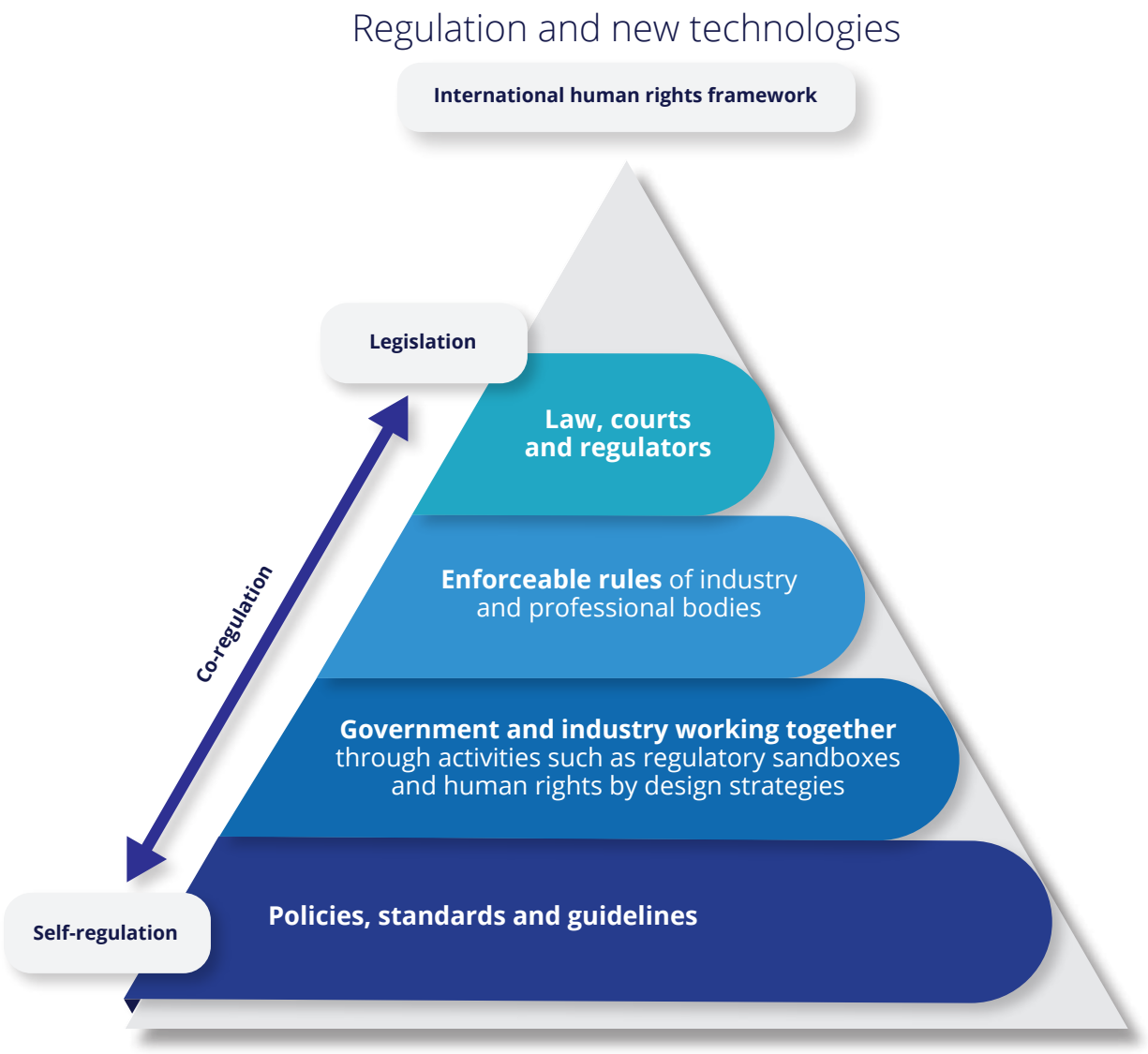
(a) Scope of regulation and governance

This section outlines the meaning of some key terms used in this Discussion Paper.

Regulation includes rules and other processes that aim to moderate individual and organisational behaviour, in order to achieve identified objectives. Broadly speaking, there are three types of regulation:

- *Legislation*—includes rules contained in Acts of Parliament, as well as subordinate legislation made under those Acts.
- *Co-regulation*—when industry develops its own rules, codes or standards. Legislation can give co-regulatory measures a particular legal status, and can enable enforcement.
- *Self-regulation*—when an industry or sector develops and voluntarily commits itself to its own rules, codes or standards. This is also referred to as 'soft law' and includes industry standards, guidelines and policies that are generally not legally binding.

This Discussion Paper sometimes uses the terms 'regulatory framework' and 'national framework', which comprise all three types of regulation. The regulatory framework may also include relevant international human rights instruments.



Governance is a broader term. The Commission uses it here to refer to the overarching system that supports regulation and oversight, including regulators, dispute resolution bodies, courts, industry and professional bodies, among others. Governance also refers to the underpinnings of the regulatory and oversight system in law, policy and education.

(b) The regulatory object

New technology can be affected by regulation in several ways. Rules can apply directly to how technology is designed or made, how it functions, or how and when it may be used. Regulation can apply also to fields of activity in which new technologies are used. Such regulation can set rules for how technology may be used in those fields, or how the technology must operate for the regulated activities to be compliant.[125]

For example, there are rules on the use of drones by members of the public. The rules stipulate permitted distances drones may be used from people and places.[126] The object of the regulation is the technology—the drone.

Alternatively, a particular technology may be used in several different ways, engaging human rights to a greater or lesser extent depending on the specific context. For example, facial recognition technology can be used as a security tool to unlock a smartphone, or this same technology can be used in policing to identify potential offenders.

The reliability of the underlying technology might be affected by many of the same sorts of factors (eg, an individual's age or skin colour), but the human rights engaged in each situation will differ. This is most apparent where the technology fails. If facial recognition fails, causing a delay in an individual 'unlocking' their smartphone, this situation is materially different from a person who is wrongly accused of a crime because the police have made an error in their use of facial recognition to identify a suspect.

What is most important here is the *outcome* of using the relevant facial recognition tool. Hence, generally it is most effective to target regulation towards the relevant outcome. In particular, if an individual is unfairly disadvantaged because of factors such as those mentioned above (age or skin colour), government should ensure this is prohibited in all circumstances.

It is already unlawful to discriminate on grounds such as age and skin colour. But existing laws are being engaged in new ways by new technologies. A major challenge is to ensure that our laws, including laws that protect human rights, continue to be applied effectively and rigorously to the use of new technology.

The level of reliability needed from a single type of technology may vary depending on the context in which it is used and its impact. There may be a need for more specific regulation in particular contexts—for example, with nuclear technology and aviation. But, on the whole, regulating an entire technology, as a technology, is likely to be ineffective.

(c) Support for regulation

Many stakeholders supported stronger regulation regarding the development and use of new technologies, with the explicit goal of protecting human rights.[127] There was particular support for a regulatory focus on AI-informed decision making,[128] and accessible technology for people with disability,[129] which are discussed in detail in later chapters.

Some submissions identified that regulation provides benefits for all sectors. It can promote responsible innovation and provide greater certainty for industry, government and members of the public.[130] The University of Melbourne observed:

> Alongside protection of the public, regulation should facilitate and ensure the appropriate functioning of a competitive marketplace for new technologies. A well-designed regulatory structure could set up a lasting framework for ways in which innovative breakthroughs can be taken to market. Equally, it should be remembered that innovation can be fostered and encouraged within a regulatory regime that encourages best practice, rather than allowing a proverbial 'race to the bottom' in the AI context.[131]

Stakeholders recommended that regulation be guided by international human rights law,[132] with acknowledgement that human rights jurisprudence may develop slowly.[133]

Regulation was said to improve public trust in new technologies, which is currently fragile,[134] especially following controversies in which business leaders have sought to avoid responsibility.[135] For example, effective regulation can build community trust in the handling of personal data and the use of AI.[136] The Australian Information Industry Association stated:

> [Community] trust in AI in the public domain will only be achieved through a strong commitment to transparency of policy, decision making and regulatory development.[137]

The related concept of *social licence* refers to private and public authorities acting responsibly and ethically to maintain the trust of the community.[138] Community confidence and trust might be improved through participatory strategies such as community consultations and consensus building,[139] and effective communication with the public.[140]

On the other hand, the Digital Industry Group, which represents a number of large technology companies including Facebook and Google, submitted that greater government involvement in AI would slow innovation and discourage investment in Australia:[141]

> [T]he time, cost and commercial concerns companies may associate with disclosing automated decision-making to an external organisation may actually serve as a deterrent to due consideration of ethical standards.[142]

(i) International approaches to guide regulation

Some stakeholders recognised that international human rights law should guide the regulation of new technologies.[143] Stakeholders made reference to all the major human rights treaties, as well as the UN Guiding Principles on Business and Human Rights,[144] the Sustainable Development Goals[145] and the Declaration on the Rights of Indigenous Peoples.[146]

Some stakeholders identified specific examples of how human rights mechanisms can be used in a practical way to guide regulation. PwC Indigenous Consulting suggested integrating international human rights standards, such as the Declaration on the Rights of Indigenous Peoples, into national technology and data-related laws to consider their impact on the rights of Aboriginal and Torres Strait Islander Peoples.[147] Similarly, Intopia submitted that the United Nations Committee on the Rights of Persons with Disabilities General Comment No. 2 should guide the regulation of accessible technologies for people with disability.[148]

At a multilateral level, many stakeholders commended the European General Data Protection Regulation (GDPR) as a useful mechanism to consider when formulating Australia's regulatory response to new technologies.[149] The GDPR aims to regulate the use of personal data and is explored in detail in Chapter 6.

(ii) Multi-faceted regulatory approach

Many stakeholders favoured a mix of regulatory responses, from legislation through to self-regulation,[150] with several referring to Braithwaite's regulatory pyramid.[151] This approach allows flexibility and agility in regulatory responses to new technologies.[152] The UN High Commissioner for Human Rights, Michelle Bachelet, has proposed a similar approach—a smart mix of measures to regulate new and emerging technologies.[153]

Some stakeholders focused on the importance of ensuring that the regulatory approach should be founded on enforceable human rights law, with complementary co-regulatory and self-regulatory mechanisms.[154] Suzor et al stated:

> Effective regulation will certainly require the active participation of technology companies, but self-regulatory and co-regulatory approaches will not be sufficient on their own. Self-regulation is often a necessary component of real change, but substantial external pressure is usually required in order for real limits to develop within a company or industry group. Enforceable legal obligations are a core requirement of a human rights regime that is more than aspirational.[155]

Industry guidelines and standards, when used within a co-regulatory framework, may have a number of advantages. These include the speed of establishment and revision, the incorporation of sector-specific knowledge, and encouragement to stakeholders to understand and accept the regulatory process.[156] The advantages of co-regulation have been said to include flexibility, predictability and access to expertise that may be beyond the reach of government agencies.[157]

(iii) Principles-based regulation

Principles-based regulation emphasises rules that set the overall objective that must be achieved, rather than more detailed, prescriptive rules. The Australian Law Reform Commission has described principles-based regulation as:

> provid[ing] an overarching framework that guides and assists regulated entities to develop an appreciation of the core goals of the regulatory scheme. A key advantage of principles-based regulation is its facilitation of regulatory flexibility through the statement of general principles that can be applied to new and changing situations.[158]

In Australia, the *Privacy Act 1988* (Cth) contains principles-based regulation, notably in the form of the Australian Privacy Principles.[159] The GDPR encompasses both principles-based and more prescriptive forms of regulation for the protection of privacy.[160]

Some stakeholders submitted that principles-based regulation is useful here as it allows for technological advancements without needing regular detailed amendment.[161] Further, it can influence 'human rights by design' processes.[162]

Some stakeholders drew attention to limitations of principles-based regulation. For example, principles-based regulation may need to be coupled with prescriptive rules to ensure efficacy.[163] A principles-based approach also can create uncertainty if it relies too heavily on the interpretation of open-ended concepts by regulators.[164]

One type of principles-based regulation has been termed 'anticipatory regulation'. This aims to engage public and diverse stakeholders in experimental, proactive and outcomes-based approaches to regulation.[165]

Some argue that anticipatory regulation can support the emergence of new technologies while also allowing faster responses to ensure citizens are protected and dangers are averted. It is said to provide 'a set of behaviours and tools—essentially a way of working—that is intended to help regulators identify, build and test solutions to emerging challenges'.[166] An industry stakeholder noted the benefit of regulation being 'anticipatory and proactive', to foster multi-disciplinary regulatory teams working on solving problems together.[167]

A regulatory sandbox is an example of anticipatory regulation that has been especially popular in the 'fintech' area. A regulatory sandbox can allow companies to test certain products or services for a period of time under the authority of the regulator, prior to obtaining the usual permit or licence.[168] Regulatory sandboxes are discussed in further detail at Chapter 7.

(d) Challenges of regulation

Some of the challenges associated with regulating the development and use of new and emerging technologies are discussed below.

(i) Regulatory lag

There was widespread concern among stakeholders about the slowness of regulation to respond to rapid technological change—a phenomenon this Discussion Paper refers to as 'regulatory lag'.[169] Rapid adoption of new technologies by government, business and individuals has contributed to delays in implementing adequate rules and oversight mechanisms for its safe development.[170]

Regulatory lag on the part of government has contributed to a drift towards self-regulation in the technology sector, as laws and regulators have not effectively anticipated or responded to new technologies. In particular, stakeholders noted that some existing human rights laws are not being effectively applied in the context of new technologies.[171] This weakens those protections and gives the appearance of gaps in the law.

For example, assume that an AI-informed recruitment tool unlawfully discriminates against an individual because of their age. Several factors could contribute to regulatory lag in this scenario. The individual may be unaware how age discrimination law applies to new technologies; public interest and community lawyers are unlikely to have encountered similar cases, which makes it more difficult to advise the individual how their rights are protected; [172] and new technologies tend to be more opaque in their operation so it can be more difficult to determine whether the law has in fact been breached. Regulators also face these challenges as they attempt to apply and enforce existing laws in the context of new technologies.

(ii) Fostering innovation

Some stakeholders raised concerns that regulation protecting human rights could stifle innovation.[173] It was suggested that poorly designed regulation could inhibit the development of start-ups, concentrating AI research and development within established companies.[174]

Other stakeholders did not see these challenges as insurmountable and suggested that an appropriate balance between legislation and self-regulation can both protect human rights and foster innovation.[175] In this respect, Microsoft observed:

> Governments must also balance support for innovation with the need to ensure consumer safety by holding the makers of AI systems responsible for harm caused by unreasonable practices. Well-tested principles of negligence law are most appropriate for addressing these issues.[176]

(iii) Legal liability and 'Big Tech' companies

Some stakeholders highlighted the challenge of regulating the activities of large, multinational technology companies. A small number of such corporations hold a disproportionately large amount of data (including personal data).[177] The scale of these data holdings is unprecedented and poses challenges, especially to the capacity of a country such as Australia to regulate global operations effectively.[178]

Stakeholders also noted the challenges that can arise in determining the entity responsible for a human rights breach, especially in complex procurement scenarios with numerous actors and opaque algorithms,[179] and when the technology is operating autonomously.[180] Several submissions claimed that proprietary and trade secrets were contributing to this lack of transparency,[181] creating difficulties in examining and explaining decisions informed by such algorithms.[182]

Stakeholders also identified challenges in regulating policy areas where human rights intersect,[183] and in ensuring that regulation responds effectively to technology that can be beneficial or harmful depending on how it is developed and used.[184]

3.3 The Commission's preliminary view

Public trust in many new technologies, including AI, is low.

The majority of respondents to a national survey were uncomfortable with the Australian Government using AI to make automated decisions that affect them,[185] and an international poll indicated that only 39% of respondents trusted their governments' use of personal data.[186]

In Australia, community concern associated with practices such as Centrelink's automated debt recovery program is emblematic of broader concerns about how new technologies are used by the public and private sectors.

Building or re-building this community trust requires confidence that Australia's regulatory framework will protect us from harms associated with new technologies.

Many stakeholders have expressed concern about a power imbalance between the consumer and 'Big Tech' companies. This problem is not unique to new technologies. For example, the recent Royal Commission into the financial services industry stated that an asymmetry of power was one of four factors leading to repeated breaches of consumer rights in delivering financial services:

> Entities set the terms on which they would deal, consumers often had little detailed knowledge or understanding of the transaction and consumers had next to no power to negotiate the terms… There was a marked imbalance of power and knowledge between those providing the product or service and those acquiring it.[187]

A firm foundation for community trust is needed to address the risks associated with emerging technologies, while harnessing the benefits. In the context of AI governance, Eileen Donahoe has said:

> The bottom line is that if we fail to address the range of significant risks of AI for people and society, we will not realize its beneficial potential either. Instead, the backlash from citizens, employees, consumers, and governments will take over. And it will be justified.[188]

The Commission proposes a *National Strategy on New and Emerging Technologies*, which protects and promotes human rights of all people, and especially vulnerable and disadvantaged groups. This is critical to building enduring public trust in those technologies.

More modern and harmonised regulation for personal data has emerged from the European Union in the form of the GDPR,[189] while in the United States there have been recent calls, from both ends of the political spectrum, to use anti-trust (or competition) law to break up Big Tech companies.[190] The debate among technology companies themselves has shifted to focus more on how—not if—their operations should be regulated.[191]

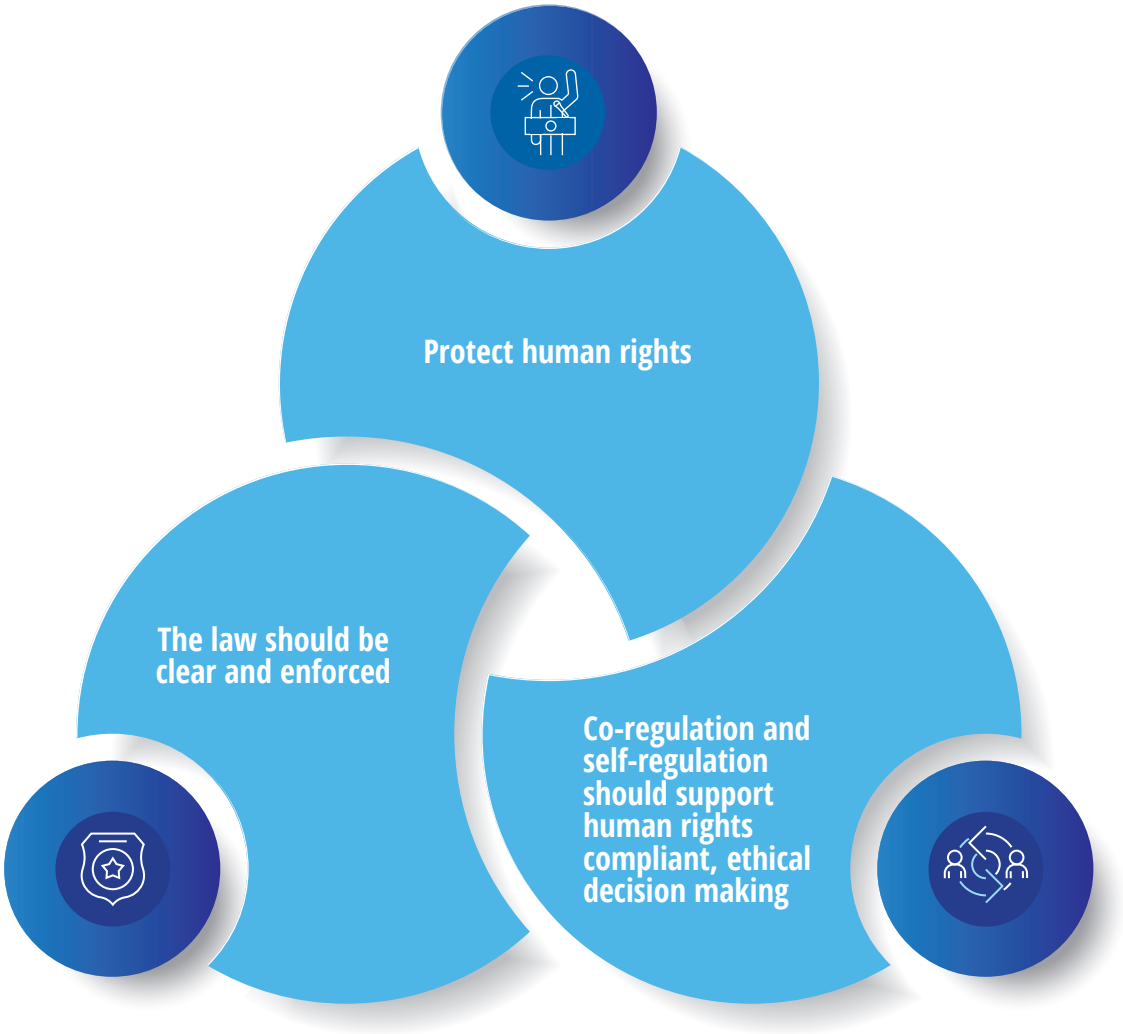
The Commission acknowledges that in respect of new technologies, just as in other fields of endeavour, government and the private sector pursue aims such as economic prosperity, national security and business growth and profits. This is legitimate provided these aims are pursued in ways that respect human rights.[192] The Commission acknowledges that these aims should be given due weight when considering Australia's regulatory approach in this area.

A broad range of human rights are engaged by the design, development and use of new technologies. As outlined in Chapter 2, the international human rights framework should underpin Australia's policy approach to new and emerging technologies.

Effective national regulation should uphold and protect human rights and instil trust in the public about how new technologies are used in Australia. Three key principles should apply:

1. **Regulation should protect human rights.** As a first principle, all regulation should be guided by Australia's obligations under international law to protect human rights.
2. **The law should be clear and enforceable**. Australian law should set clear rules regarding the design, development and use of new technologies, with a view to promoting human rights and other aims consistent with our liberal democracy.

   The broader regulatory framework should contain effective remedies and enforcement mechanisms. Australia's law-makers should fill any gaps necessary to achieve these regulatory aims.
3. **Co-regulation and self-regulation should support human rights compliant, ethical decision making.** The law cannot, and should not, address every social implication of new and emerging technologies. Good co- and self-regulation—through professional codes, design guidelines and impact assessments—can promote the making of sound, human rights compliant decisions by all stakeholders.



This approach involves a range of regulatory mechanisms. What regulation looks like will depend on what is being regulated. Concerns about corporate behaviour, for example, may lead to a focus on accountable governance in private entities. A focus on technology products may lead to solutions such as self-certification or trademarks, or a focus on the design of a product, leading to changes in the design process itself. A focus on how technology products are used by decision makers may lead to rules about procurement.

This regulatory approach should help protect human rights and mitigate risks, as it should include the application of existing laws, as well as law reform where gaps need to be filled. It should also support technological innovation. Innovation can flourish in environments where actors understand how human rights need to be protected and the parameters in which they can operate.[193] Additionally, the public's trust in the design, development and use of new technologies should improve as regulatory parameters are defined and applied.

PROPOSAL 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.**

'The Australian Government should develop a *National Strategy on New and Emerging Technologies*.'

3.4 National regulation and governance

Stakeholders urged reform of Australia's regulatory and governance framework for human rights protection as it applies to new technology, especially in relation to legislative protection, federal regulatory bodies, and means of co- and self-regulation.

(a) Legislation

Many stakeholders identified the most significant gap as being the absence of legislated human rights protection, particularly through a national Human Rights Act or charter.[194]

Given that new technologies, especially AI, rely heavily on personal information, many stakeholders focused on privacy and data protection rights. Several highlighted ways in which the right to privacy receives only limited protection under Australian law.[195] Some stakeholders suggested implementing recommendations arising from several recent legislative reform and policy inquiries,[196] by enacting a statutory tort for breach of privacy.[197]

It was observed that privacy protections often do not apply where an individual has provided consent to the use of their personal data.[198] However, there was concern about over-reliance on this consent model, especially in the context of mass data sharing. The Australian Competition and Consumer Commission (ACCC) recently recommended strengthening consent requirements to require that consent is freely given, specific, unambiguous, and informed.[199]

Similarly, several stakeholders suggested legislation based on the GDPR.[200] Some also urged a regulatory response to:

- link AI policy with data governance[201]
- address discrimination in data sets[202]
- provide better opt-in and opt-out protections[203]
- strengthen security and protection of personal data.[204]

A common theme among stakeholders was that privacy was not the only human right in need of protection.[205] Other rights, such as equality and non-discrimination,[206] and the right to a fair trial,[207] also need regulatory attention. For example, stakeholders pointed to gaps in the law relating to:

- online social media platforms, where there is a rise in technology-based violence, particularly against women.[208] Stakeholders suggested changes in criminal law to help protect against technology facilitated abuse of women and girls, such as the non-consensual sharing of intimate images[209]
- anti-discrimination law, given that it can be difficult to establish discrimination in situations involving use of opaque algorithms or biased datasets[210]
- the related concept of consumer rights,[211] which are not protected in some circumstances where AI is used.[212]

There are few overseas examples of laws that are expressly directed towards protecting human rights in the context of new technologies. The most cited was the GDPR, which applies to all companies processing personal data of individuals residing within the EU, regardless of whether the company has a physical presence in the EU.[213] The GDPR includes protections that do not exist under current Australian law, including the right to access confirmation from companies as to whether personal data is being processed and for what purposes, and the right to be forgotten.[214]

Sometimes legislation seeks to address particular risks associated with a particular technology, such as legislation to address potentially harmful use of facial recognition technology, discussed in greater detail in Chapter 6.

(b) Regulatory bodies

At the federal level, there are two broad categories of regulator with responsibilities relevant to the development and use of new technologies. First, there are bodies that regulate a particular type of technology, such as:

- the Gene Technology Regulator, which administers Australia's national gene technology regulatory scheme to protect human health by assessing and managing risks related to genetically modified organisms[215]
- the Australian Radiation Protection and Nuclear Safety Agency, which regulates the use of radiation by Commonwealth entities[216]
- the Civil Aviation Safety Authority, which regulates Australian aviation safety and operation of Australian aircraft, including technology such as drones.[217]

The second category of regulator is much more common. It does not regulate a type of technology, but regulates particular activities or actors. As new technologies, such as AI, are increasingly used in those activities, or by those actors, this second category of regulator needs to understand how to achieve their aims in this changing context. Examples include the following.

- The ACCC administers the *Competition and Consumer Act 2010* (Cth) to promote competition and ensure consumer protection. It recently recommended that a specialist digital platforms branch be established to build on and develop expertise in digital markets and the use of algorithms.[218]
- The Australian Securities and Investments Commission (ASIC) regulates the incorporation, operations and management of registered corporations and administers the *Corporations Act 2001* (Cth), *Insurance Contracts Act 1984* (Cth) and the *National Consumer Credit Protection Act 2009* (Cth). It hosts several government initiatives relevant to the regulation of new technologies, including the Innovation Hub which provides practical support and informal assistance to financial technology businesses.[219]
- The Office of the Australian Information Commissioner (OAIC) administers the *Privacy Act 1988* (Cth) and has functions related to privacy and freedom of information. It also administers the Notifiable Data Breaches Scheme, which sets out obligations for notifying affected individuals about data breaches likely to result in serious harm.
- The National Data Commissioner (NDC) oversees Australia's data-sharing and release framework for public sector data. It monitors, enforces and reports on the operation of forthcoming data legislation.[220] The National Data Advisory Council advises the NDC on ethical data use, community expectations, technical best practice, and industry and international developments.[221]
- The Australian Communications and Media Authority (ACMA) regulates broadcasting, the internet, radiocommunications and telecommunications. ACMA's Regulatory Futures Section is undertaking a project to explore the regulatory pressures resulting from the communications and media industry's use of AI.
- The Commonwealth Ombudsman investigates the administrative actions of Australian Government bodies and prescribed private sector organisations.[222] The Ombudsman conducted an own-motion investigation into the Online Compliance Intervention program deployed by the Department of Human Services, following a spike in the number of Centrelink debt-related complaints it was receiving.[223]

- The Office of the eSafety Commissioner coordinates and leads online safety efforts to protect and enhance the online safety of all Australians.[224] The *Enhancing Online Safety Act 2015* (Cth) establishes a scheme for the removal of cyberbullying material from social media websites that have partnered with the Office. The Office can also issue notices directly to individuals who have posted cyberbullying material.[225]
- The Australian Prudential Regulation Authority (APRA) regulates the operations of banks, life insurers, building societies, credit unions, friendly societies and superannuation funds.[226] APRA has raised concerns about the risks of the use of AI in the insurance industry.[227]

(c) Co-regulation and self-regulation

Stakeholders highlighted the role that co- and self-regulation can play in this area, such as government guidelines and policies, industry standards, and tools to influence the design of new technologies. Co- and self-regulation can form part of a multi-faceted approach to the regulation of new technologies.

(i) Government guidelines and policies

The Australian Government and regulators develop advice and guidance on the use of technologies by public and private actors in certain contexts. Examples identified by stakeholders include:

- *Automated Assistance in Administrative Decision-Making: Better Practice Guide* to guide government decision-makers about automated decisions[228]
- the *Digital Service Standard*, which provides best practice for the design and delivery of government digital services[229]
- ASIC's *Providing digital financial product advice to retail clients*, which offers guidance regarding automated financial product advice using algorithms and technology and without the direct involvement of a human adviser.[230]

Stakeholders also noted that government procurement policies can support the promotion of human rights, as they provide strong incentives for industry to 'meet, support and develop standards and certification processes'.[231]

(ii) Industry standards and guidelines

A range of standards, certification schemes, guidelines and codes may apply to new technologies. They may be voluntary and self-regulated, or binding if they are included in regulation or made compulsory by a professional industry body.[232] They may also apply to products and services such as the insurance sector, or categories of people such as those accredited or approved to provide insurance products.

Several prominent non-government international entities create standards relevant to new technologies, including the International Standards Organization (ISO) and the Institute of Electrical and Electronics Engineers (IEEE).[233] Standards Australia is the national peak non-government standards organisation which develops and adopts internationally-aligned standards and represents the nation at the ISO.[234]

ISO, IEEE and Australian standards are voluntary, and the respective standard bodies do not enforce standards. However, federal, state and territory parliaments can give standards legislative force.

Several stakeholders submitted that both binding and voluntary standards are useful tools to manage the impacts of new technologies,[235] especially as their development can incorporate co-design to include community members.[236] A common theme identified by stakeholders was the need for consistency in standards.[237] Microsoft noted that

> in a world marked by increasingly complex supply chains, adopting common norms globally, through Standards, can be beneficial for companies and consumers.[238]

Some of these standards and guidelines aim to influence the design process of new technologies, such as the IEEE's standards focused on the social impacts of information and communications technology (ICT) software in the design phase.[239]

Certification and trust marks may be used to incentivise adherence to recognised standards and were identified as effective accountability mechanisms for new technologies.[240] Some stakeholders submitted that voluntary self-certification may be an inadequate form of regulation in isolation,[241] and be more effective when part of a co-regulatory approach.[242]

The UN Global Compact provides guidance, best practices and resources to support businesses fulfil their human rights obligations across all their operations.[243] Stakeholders referred to UN Global Compact guidelines to show how businesses can help protect human rights through following guidelines that are tailored to their particular industry.[244]

Other strategies that promote human rights-compliant design, such as 'human rights by design', impact assessments and capacity building, received positive attention from stakeholders. These strategies are discussed in relation to AI-informed decision making and accessible technology at Parts B and D.

3.5 The Commission's preliminary view

The Commission's proposed *National Strategy on New and Emerging Technologies* should include effective regulation, based on international human rights law, enforceable law, and co-regulation and self-regulation. This approach should include a well-integrated range of the regulatory tools.

Many new technologies are developed and used in ways that engage human rights, and warrant an appropriate regulatory response. In Parts B and D, the Commission outlines detailed regulatory approaches regarding AI-informed decision making and accessible technology for people with disability.

The Commission proposes some key principles to guide law and policy makers, as well as regulators, in the regulation of the development and use of new technologies.

(a) Application and reform of legislation

There is a need to identify gaps in how our law protects human rights in the context of new technologies. When considering the adequacy of our existing law, we need to ask the following questions:

- Is there existing legislation that is relevant to the use of this technology?
- How can existing legislation be better applied and more consistently enforced?
- Does the legislation need to be amended to safeguard human rights?
- Do regulators and other governance bodies need to improve how they support the aims of our regulatory system?

(b) Regulatory and oversight bodies

Regulatory and oversight bodies would benefit from increased understanding of new technologies as they fulfil their functions in an environment that is increasingly defined by the presence of new technologies like AI.

Legal, policy and other experts with senior roles in these entities do not need to be technology experts, but they do need to understand how these technologies operate and specifically how they affect their work.

They also need access to technical expertise to support their regulatory activity.

Equipping these regulatory bodies for this new era should involve the following activities:

- targeted education and training for decision makers and leaders in these bodies
- resourcing these bodies to better understand and explain how existing rules apply in the context of new technologies
- encouraging these bodies to promote human rights-compliant behaviour in the sectors in which they operate.

Expanded regulatory or oversight functions might be required for some regulators. In its *Digital Platforms Inquiry* final report, the ACCC made a number of findings regarding the powers and resourcing of the OAIC, in recognition of the increased volume, significance and complexity of privacy-related complaints that arise in the digital economy.[245]

It also recommended that specialist branches be established within the ACCC to build and develop expertise in digital markets and use of algorithms.[246]

Examples of these regulators clarifying the law and giving guidance on how to use new technologies in a human rights compliant manner include the Ombudsman's Guide to automated decision making,[247] and ASIC's policy development. ASIC proposed a three-pronged approach to algorithmic accountability, requiring:

- a responsible person for an algorithmic system
- that relevant algorithmic outputs be capable of an explanation so that decisions relying on them can be meaningfully explained to customers
- opportunity for redress when mistakes are made due to the dataset or algorithmic design.[248]

The Commission proposes at Chapter 8 a mechanism for leadership on AI in Australia, through the establishment of an AI Safety Commissioner to guide and equip all sectors in their development, use and regulation of new technologies. That Commissioner could be responsible for leading these regulatory and governance bodies in carrying out their functions in conjunction with new technologies.

(c) Co-regulation and self-regulation

The Commission agrees with stakeholder views that each technology and setting requires its own approach to co- and self-regulation. Co-regulation can enhance accountability for new technologies and promote the adoption of a human rights approach to product development. Self-regulation can also play an important role in complementing the law. These co- and self-regulatory measures can be beneficial where there is industry buy-in and involvement in their development, and they can thereby help improve the public's trust in goods and services that use new technologies.

CHAPTER 4
Ethical frameworks

4.1 Introduction

There has been a growing trend to frame many of the social implications of new and emerging technologies, and especially those that involve a risk of harm to humans, as *ethical* issues.

In response, there have been many new initiatives from government, the private sector and civil society, which seek to identify the ethical implications of new technologies, and set out ways to develop and use these technologies ethically. These initiatives include new ethical policies, principles and codes, which apply especially to the use of AI and automated decision making.

This Discussion Paper refers to these initiatives collectively as 'ethical frameworks', and this chapter considers the place of such frameworks in the context of a multi-layered regulatory response to the rise of new technologies.

An ethical framework can mean different things to different people, but almost all are designed to address risks of harm caused by new technologies, via a voluntary system, with no or limited legal enforcement. Some stakeholders expressed support for ethical frameworks in addressing the risks of new and emerging technologies; others doubted the effectiveness of these voluntary measures achieving their aims and especially in protecting human rights.[249]

For the reasons expanded on in this chapter, the Commission cautions against an over-reliance on ethical frameworks to address social harm associated with new technologies, and to fulfil basic human rights. This is for two reasons.

First, there is a lack of consensus, at the level of detail, about the actual ethical rules or principles which should inform these frameworks. Perhaps as a result of this lack of consensus, the ethical principles incorporated in or informing the ethical frameworks that have been put forward to date are generally framed at a high level of abstraction or generality. This does not provide adequately detailed guidance about how new technologies should be developed and used.

Secondly, the ethical frameworks proposed or published to date have generally been voluntary in nature. Given the serious harms that can result from the use of new technologies, voluntary frameworks are not of themselves sufficient to promote and protect human rights in this field.

In Chapter 3, the Commission proposes an overarching regulatory approach for new technologies. This approach effectively integrates laws, regulatory institutions and voluntary systems—such as ethical frameworks—with a particular focus on protecting and promoting human rights. The Commission proposes in this chapter that the Australian Government prioritise ethical frameworks that protect human rights. Such ethical frameworks can play an important role in setting guidance and expectations for industry, and serve a distinct but complementary function as compared with enforceable laws.

4.2 Emergence of ethical frameworks

There is considerable variation in the content and practical utility of the many ethical frameworks that have been recently developed for new and emerging technologies.

(a) What is meant by the term 'ethics'?

Broadly speaking, the field of ethics is concerned with identifying standards, values and responsibilities that allow us to judge whether decisions or actions are appropriate, 'right', or 'good.'[250] In its ethical framework for 'good technology', The Ethics Centre describes what it means for human decision making to be guided by ethics:

> So often we describe the world: what is likely to happen, what might happen or what will happen. Ethics allows us to judge the world—what should happen? Of all the ways you might act, which is the best? Which of all the possibilities should you bring into reality? What ought one to do? That's the question ethics seeks to answer.[251]

The word 'ethics' can take different shades of meaning, depending on the circumstances in which it is used. The *Macquarie Dictionary*, for example, identifies three related, but distinct, senses in which the word is commonly used.

The first is 'a system of moral principles, by which human actions and proposals may be judged to be good or bad or right or wrong'. The second sense is similar, but relates to the particular values held by an individual (rather than those shared generally, or by a community). The third sense is 'the rules of conduct recognised in respect of a particular class of human actions'.

Examples of this final usage are 'medical ethics', 'legal professional ethics', or other professional ethical principles or frameworks. This last sense is still concerned with judging right or appropriate conduct, but in a specific limited context. The values or principles included in systems of professional ethics will frequently include matters that go beyond questions of what is morally right and wrong in the everyday sense, such as the need to protect the reputation or standing of the profession in question.

Frameworks that assist in examining whether current technologies or their uses are 'ethical', or that seek to provide guidance to ensure that future technologies are developed and used 'ethically', are generally using the term 'ethics' in the first of the senses given above. However, some ethical frameworks discussed in this chapter fall within the third of these senses, as they include norms regulating professional conduct which go beyond what might ordinarily be considered to be moral principles. In considering and assessing the utility of 'ethical frameworks' (either in general, or in particular cases), it is important to bear in mind the various senses in which the word 'ethics' may be used and these frameworks' varying aims.

There is no universally accepted ethical system or agreed set of ethical principles. That is reflected in the fact that an examination of current ethical frameworks demonstrates that they take different approaches to regulating new technology.

The most prominent ethical theories include, for example: consequentialism, which emphasises outcomes; the deontological approach, which focuses on doing what is right, regardless of outcome; and the teleological approach, which requires a focus on purpose, such as supporting human flourishing.[252]

While these ethical theories may have significant areas of overlap, applying one approach can lead to a different conclusion about the best course of conduct than applying another.

Much of the current public debate about ethics and new technologies has not been grounded expressly in any particular ethical philosophy. Instead, the term 'ethics' is frequently a catch-all term that refers to a range of frameworks, which are not part of the law, and are generally designed to address the risk of harm. If such approaches have any common or unifying normative core, it seems only to be a very general one: a desire to do good and avoid harm.

The lack of agreement about the content of the ethical rules and principles that should be applied to the development and use of new technologies may partly explain why many recent ethical frameworks set out very general principles. This problem is acknowledged by some working in this field. For instance, the Gradient Institute has written:

> The people ultimately responsible for an AI system must also be responsible for the ethical principles encoded into it. These are the senior executives and leaders in business and government who currently make strategic decisions about their organisation's activities. The ethical principles that they determine will, as with the decision-making they already do, be informed by a mixture of their organisation's objectives, legal obligations, the will of the citizens or shareholders they represent, wider social norms and, perhaps, their own personal beliefs. Gradient Institute has neither the moral nor legal authority to impose its own ethical views on AI systems, nor the domain expertise to balance the complex ethical trade-offs inherent in any particular application.[253]

This lack of agreement does not mean that ethical frameworks cannot play an important part in the development of new technology. However, ethical frameworks cannot provide a comprehensive response to the risks identified in this Discussion Paper.

(b) Proliferation of ethical frameworks

There are many ethical frameworks that have been developed recently in this context across various sectors.[254]

(i) Government and intergovernmental initiatives

Ethical frameworks have been developed in several jurisdictions, primarily to manage the risk and opportunities associated with AI. These initiatives encompass legal, policy and voluntary approaches.

For example, the Australian Government Department of Industry, Innovation and Science published an AI Ethics Framework in November 2019.[255] It contains eight voluntary principles that 'can inform the design, development and implementation of AI systems':[256]

- Human, social and environmental wellbeing: Throughout their lifecycle, AI systems should benefit individuals, society and the environment.
- Human-centred values: Throughout their lifecycle, AI systems should respect human rights, diversity, and the autonomy of individuals.
- Fairness: Throughout their lifecycle, AI systems should be inclusive and accessible, and should not involve or result in unfair discrimination against individuals, communities or groups.
- Privacy protection and security: Throughout their lifecycle, AI systems should respect and uphold privacy rights and data protection, and ensure the security of data.
- Reliability and safety: Throughout their lifecycle, AI systems should reliably operate in accordance with their intended purpose.
- Transparency and explainability: There should be transparency and responsible disclosure to ensure people know when they are being significantly impacted by an AI system, and can find out when an AI system is engaging with them.

- Contestability: When an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or output of the AI system.
- Accountability: Those responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled.[257]

The second principle on human-centred values states:

> AI systems should enable an equitable and democratic society by respecting, protecting and promoting human rights, enabling diversity, respecting human freedom and the autonomy of individuals, and protecting the environment. Human rights risks need to be carefully considered, as AI systems can equally enable and hamper such fundamental rights. It's permissible to interfere with certain human rights where it's reasonable, necessary and proportionate.[258]

Some overseas governments and intergovernmental organisations have also created ethical frameworks for new and emerging technologies. For example, the UK Government's Centre for Data Ethics and Innovation is midway through a two-year public consultation regarding the ethical development and use of data-driven technology.[259] Also, the European Commission's High-Level Expert Group on AI published 'Ethical guidelines for trustworthy AI' in April 2019, following a public consultation.[260]

(ii) Non-government initiatives

Groups of technology firms, individual companies and not-for-profit organisations have adopted ethical frameworks to define what is acceptable conduct in the context of new and emerging technologies, particularly with regard to AI technologies.

- **Private sector initiatives**

Many large technology companies have begun individually and collectively to produce ethical frameworks to guide the development and use of new technologies.[261] These self-regulatory processes are taking various forms and structures.

Salesforce, for example, has established an Office of Ethical and Humane Use of Technology, which will develop and implement an ethical framework for how that company deals with emerging technologies.[262] Facebook has also recently published a draft charter to create an Oversight Board for Content Decisions—essentially to create an internal appeals mechanism in respect of decisions made by Facebook to remove content uploaded by Facebook users. This oversight board will review Facebook's 'most challenging content decisions'.[263]

Similarly, Google has established its AI Principles, which are an ethical charter to guide the responsible development and use of AI in Google's research and products. Among other things, the Google AI Principles aim to promote only socially beneficial AI and avoid creating or reinforcing unfair bias.[264] Google announced an Advanced Technology Advisory Council to support the implementation of its AI Principles.[265] However, widespread concerns were expressed about the membership of the Advisory Council, including from Google employees and human rights leaders,[266] and Google ended the initiative soon after its announcement.[267]

Other technology companies have established their own ethical frameworks to guide the development of their products. Microsoft, for example, has published the 'Microsoft AI Principles' intended to ensure the company designs trustworthy AI products that 'reflect ethical principles that are deeply rooted in important and timeless values', such as fairness, accountability, transparency and inclusiveness.[268]

Large technology companies have also begun to fund research into 'ethical technology'. In early 2019, for example, Facebook partnered with the Technical University of Munich to support the creation of an independent AI ethics research centre. The objective of the centre is to 'advance the growing field of ethical research on new technology' and, specifically, to explore 'fundamental issues affecting the use and impact of AI'.[269]

Other technology companies have established codes of conduct to guide their employees at an individual level. PayPal, for example, has a Code of Business Conduct & Ethics that states that, when faced with ethical dilemmas, PayPal employees must query whether the outcome is honest and fair, and consistent with the law.[270] Similarly, Akamai Technologies, a smaller content delivery network and cloud service provider, has published a Code of Ethics which states that Akamai staff must pursue innovation to continually improve customer value, encourage employee innovation, initiative and appropriate risk taking, communicate openly and honestly and demonstrate professionalism.[271]

- **Professional codes of conduct and standards**

A number of professional and representative bodies, which exercise self-regulatory functions in prominent fields for the development of new technologies, are starting to incorporate ethics into their accreditation or professional development requirements.

The IEEE, for example, has a Code of Ethics incorporating eight core principles. These require software engineers, among other things, to act in a manner that is in the best interests of the client and consistent with the public interest; maintain integrity and independence in their professional judgment; and advance the integrity and reputation of the profession.[272]

The Association of Computing Machinery's (ACM) Code of Ethics and Professional Conduct outlines general ethical principles for computing professionals. This Code stipulates that all computing professionals must use their skills for the benefit of society and human wellbeing[273] and must avoid harm, especially when there are negative consequences that are 'significant and unjust'.[274]

Accenture has provided ethical guidelines to businesses working with and developing AI. Accenture has stated that codes should focus on making it clear where liability lies when systems make mistakes, and that general principles should guide accountability.[275] Accenture also suggests that ethical codes should avoid bias, and that core values such as equality and anti-discrimination must be promoted.[276]

Similarly, the Code of Ethics of the Association for Information Science and Technology urges its members to be ever aware of the 'social, economic, cultural and political impacts of their actions or inaction'.[277]

- **Civil society initiatives**

Civil society organisations have also produced ethical frameworks to guide the development and use of new technologies. The Future of Life Institute, for example, has developed the Asilomar AI Principles, which identify relevant 'ethics and values' such as safety, transparency and human values, the latter requiring AI to be 'compatible with ideals of human dignity, rights, freedoms, and cultural diversity'.[278] Similarly, the Toronto Declaration, adopted by a number of international civil society organisations in 2018, outlines how human rights and ethics should inform the development and use of AI.[279]

4.3 Impact of ethical frameworks

Understanding the proper role of ethical frameworks, within a broader regulatory system, must start with an assessment of the strengths and limitations of an ethics-based approach.

(a) Limitations of ethical approaches

(i) Lack of substantive normative content

Many recent ethical frameworks contain high-level value statements, which are not precisely defined. For example, the principle 'do no harm' is common in such frameworks, but its precise meaning in the context of developing and using new technologies is not widely understood or agreed.

Former US Ambassador Eileen Donahoe, now of Stanford University, has expressed concern about the effectiveness of many ethical frameworks in the context of new technology. She has said that, collectively, they lack 'normative sway' and they fail to grasp the full spectrum of human rights and governance challenges. She states, 'Ethics statements may guide the entities that commit to them, but they do not establish a broad governance framework under which all can operate'.[280]

The Australian Government's AI Ethics Framework, outlined above, is an important, but modest, step that aims to prevent social harm associated with AI. The Framework adopts a 'human, social and environmental wellbeing' principle, where it states that AI should benefit individuals, society and the environment and encourages the use of the UN's Sustainable Development Goals.

The Framework notes that 'ideally, AI systems should be used to benefit all human beings', and that positive and negative impacts on wellbeing should be accounted for. However, 'wellbeing' and 'benefit' are not defined and would likely be subject to numerous competing interpretations. This principle could suggest that a negative impact on an individual or a minority will be acceptable if there are greater benefits for another individual or group.

More work will be needed to apply this principle in practice. Smartphone technologies, for example, support the enhanced independence of people with disability; they can also be used by an abusive and controlling partner in a domestic violence context.[281] It would not be acceptable—on ethical, moral or legal grounds—to justify this harmful use simply because there is also a benign use. Further, many of the positive and negative impacts of AI may be intangible and so, by their nature, difficult to quantify.

(ii) From principle to practice: ethical professional codes

Ethical frameworks that guide the work of professionals, particularly those published by technology companies, frequently outline a commitment to the common good. However, they often provide little practical guidance to those who design new technology products and services, or to those who purchase these products and services.

A professional ethics code primarily aims to guide those who subscribe to the code. A code for electrical engineers, for instance, should be framed in a way that is relevant and readily applicable to those professionals. It should also give practical guidance for electrical engineers in making difficult decisions that involve weighing competing harms, rights or interests.

Academic research on ethical frameworks is beginning to identify the challenges of applying high-level ethical principles in practice. In an empirical study of professional codes relevant to data science and data scientists, Stark and Hoffman observed that professional ethics codes must balance the need for visibility with vagueness or generality. They concluded that

> ethics codes often elide granular attention to professional activities, relying instead on informal everyday rules over which individual practitioners have some (albeit limited) control.[282]

Empirical research is also testing the extent to which professionals are influenced by the codes that apply to them. A recent study found that expressly instructing software engineering graduate students and professional software developers to consider the ACM's Code of Ethics in their decision-making 'had no observed effect'.[283]

A further challenge for technology professionals is the variation between different professional ethics codes. Ethical frameworks are often being developed in isolation from one another, with limited or no interoperability and no formal connection. This will be challenging where professionals may be subject to multiple codes, or where closely aligned professional sectors interact but work within frameworks with different ethical objectives.

(iii) Lack of enforcement and accountability

The ethical frameworks discussed in this chapter generally operate on a voluntary basis. This means that there is frequently no rigorous, independent way of holding an individual or corporation to account in adhering to these principles, and no concrete consequences that flow from a failure to adhere.

This is not inherently problematic. A voluntary commitment to abide by certain ethical principles can influence behaviour. A problem arises, however, if such voluntary commitments occupy the proper place of enforceable legal rules.

The AI Now Institute has observed that, in the rush to adopt ethical frameworks, often responding to a particular controversy, 'we have not seen strong oversight and accountability to backstop these ethical commitments'.[284] Some have gone further. In their empirical review of ethical codes of conduct for computer scientists, Stark and Hoffman concluded:

> If data is indeed the new oil, the IEEE and other data science codes should be at least as explicit as that of petroleum engineers in listing consequences for violations of the code and articulating how those violations can be reported.[285]

Some stakeholders referred to the limitations of voluntary, self-regulatory codes without any accompanying legal framework.[286]

Concern is also emerging among experts in this field that the rapid and widespread adoption of unenforceable, voluntary ethical frameworks by technology companies, particularly with regard to AI, is distracting attention from the need for binding forms of governmental regulation.[287] A member of the European Commission Independent High–Level Expert Group on AI has commented that some technology companies fund research in the area of ethical use of AI to create the 'illusion of action', and with a view to stalling substantive regulation of AI.[288] He went on to say:

> This phenomenon is an example of 'ethics washing'. Industry organises and cultivates ethical debates to buy time to distract the public and to prevent or at least delay effective regulation and policy-making.[289]

Similar concerns were also raised by stakeholders, who noted that voluntary ethical frameworks within private companies would not be an adequate substitute for enforceable legal rules.[290]

(b) Ethical frameworks and international human rights law

This Project examines the challenges and opportunities raised by new and emerging technologies by reference to international human rights law. The same frame of reference has been adopted by a growing number of influential experts, organisations and governmental bodies.[291]

As discussed earlier, international human rights law provides agreed norms, as well as an extensive body of jurisprudence demonstrating how those norms should be applied in concrete situations. By contrast, current ethical frameworks have no equivalent agreed normative content.[292]

To put this another way: a human rights approach rests on 'a level of geopolitical recognition and status under international law that no newly emergent ethical framework can match'.[293] In her recent comment on the need to carefully regulate new technologies, the UN High Commissioner for Human Rights stated there was no need to 'start from scratch', noting the strength of the universal human rights framework to address the challenges posed by new and emerging technologies.[294]

Some ethical frameworks refer to human rights as a source of normative content. The OECD AI Principles, for example, state that 'AI systems should be designed in a way that respects the rule of law, human rights, democratic values and diversity'.[295]

Some stakeholders preferred a human rights approach to these issues, rather than an ethics-based approach.[296] The Castan Centre, for example, noted the limitations of an ethics-only approach, including the 'slippery and ill-defined' terms that often make up ethical frameworks, such as 'good' and 'unfair'.

While acknowledging industry initiatives engaging with these concepts, the Centre also expressed concern that 'these efforts are insufficiently inclusive, participatory and representative', and noted the unique position of human rights institutions to work within an existing structure with a legitimate, normative framework.[297]

4.4 The Commission's preliminary view

(a) Classifying the role of ethical frameworks

As discussed in Chapter 3, the complex challenges presented by new and emerging technologies require a multi-faceted regulatory approach. The Commission proposes an approach that starts with applying our current laws more effectively, and identifying gaps in the law that require reform. Where the law is appropriately silent or does not provide a clear rule to apply, the Commission supports other measures, such as ethical frameworks, to provide guidance that protects human rights.

This approach reinforces the primary role of Australian and international law in setting the core parameters for how new and emerging technologies are developed and used. There is a supportive 'ecosystem' that comes together to apply and uphold the law. That ecosystem consists of courts and regulators, as well as other bodies inside and outside government.

Adopting this approach still leaves considerable scope for all sectors to innovate without being constrained by unwarranted legal regulation. The Commission acknowledges that the law should not seek to control or police every conceivable activity.

Ethical frameworks can be important, but they cannot be a substitute for the law. This is as true amid the rise of new technologies, as it is in any other context. The Commission considers that there is a need to re-articulate the conventional relationship between the law and ethics in regulating behaviour. Well-developed ethical frameworks can play an important role in the shadow of the law—complementing and reinforcing the law, and providing a guide for individuals to make autonomous decisions in developing and using new technologies.

In liberal democracies like Australia, it is axiomatic that enforceable laws are necessary to constrain some activities, and especially to avoid significant social and individual harm. Laws have an important role in crystallising social norms in a diverse range of areas, such as human rights, competition and consumer protection among many others.

As noted above, the law does not, and cannot, guide all human activity. It is common for the law's interstices to be filled by ethical and other principles, which help us to make *good* choices.

Part of the impetus behind an ethics-focused approach in the context of new technologies has been some ambiguity about the law—especially whether and, if so how, existing laws apply in this context. The overarching regulatory system has struggled to keep pace in an era of exceptionally fast technological development.

In this context, ethical frameworks can offer guidance on developing and using new technologies in ways that avoid harm and promote human rights. Ethical frameworks can thereby play a valuable role, supporting appropriate legal regulation.

To this end, the proposed *National Strategy on New and Emerging Technologies,* as set out in Chapter 3, should promote the role of ethical frameworks as co- and self-regulatory mechanisms that support and complement enforceable human rights and other laws.

(b) Existing ethical frameworks

The level of reliance placed on many recent ethical frameworks may be too great, and they seem unlikely to be effective in their aims of ensuring that new technologies are developed and used in ways that promote societal good and avoid harm.

A number of stakeholders emphasised the overlap between ethics and human rights.[298] The Commission endorses calls for a thorough examination of how well ethical principles reflect the requirements of international human rights law, in the context of new and emerging technologies, with a particular focus on 'potential areas of overlap that may lack clarity and/or produce tensions due to differing approaches'.[299]

In this context, there is a need to map and assess existing ethical frameworks, with a view to evaluating their efficacy in promoting and protecting human rights. Analysis along these lines has recently started in some comparable jurisdictions.[300]

(c) Consolidation of ethical frameworks

The proliferation of overlapping ethical frameworks in this area leads to a variety of problems. Multiple ethical frameworks, covering the same sorts of activities, can frustrate attempts to achieve industry-level compliance and streamlined approaches across different settings. For example, it is feasible, perhaps even likely, that software engineers at different companies will approach the same ethical dilemma entirely differently.

In addition to mapping and assessing ethical frameworks, there would be value in consolidating some existing ethical frameworks, perhaps at an industry level.

Individual companies could then apply the applicable framework or frameworks to their own unique circumstances, or where necessary, address gaps in the industry-wide approach. This would help achieve consistent ethical norms across particular industry groupings and provide greater certainty for consumers as well as the market.

The Australian Government could play a role in any consolidation. This could be through the granting of special legal status to some broadly-applicable frameworks that meet particular criteria, such as the incorporation of ethical guidelines into legislative instruments that regulate the development of new technology.

As previously noted, the Department of Industry, Innovation and Science has published an Australian Ethics Framework for Artificial Intelligence. A next stage to this work could focus on the interaction between this document and other ethical frameworks.

(d) Ethical guidance for professional and industry associations

It is essential that professional ethical codes provide sufficient practical guidance to assist those responsible for the design, development, procurement and use of new technologies.

Take a hypothetical scenario, adapted from a well-known real situation.[301] Judges find it difficult to apply objective criteria to determine the appropriate prison sentence for people who have been convicted of particular crimes, especially when considering factors like an offender's chance of re-offending in future. For this reason, imagine that a government department (Department X) engages a technology company (Company Y) to develop an AI-powered tool to assist in that decision-making process.

A range of *legal* requirements would apply to this scenario: privacy law would restrict some aspects of how personal information is used by Department X and Company Y; this decision-making system would need to comply with existing laws that protect equality before the law; and so on.

In addition to these legal requirements, a range of ethical questions also would arise. For example, what safeguards are needed to ensure this information does not further disadvantage groups that have suffered historical prejudice in the justice system? Public servants specialising in justice policy within Department X, data scientists at Company Y, judges, lawyers and others all would be involved in procuring, developing or using this decision-making system. Each may be subject to their own varying professional and ethical obligations.

This hypothetical example illustrates the complexity of how ethical issues arise in applying new and emerging technologies—especially when those technologies are used in sensitive areas of decision making where significant human rights issues are at stake.

It would be helpful to consider how ethical issues relating to new and emerging technologies arise by reference to particular professional or other associations.

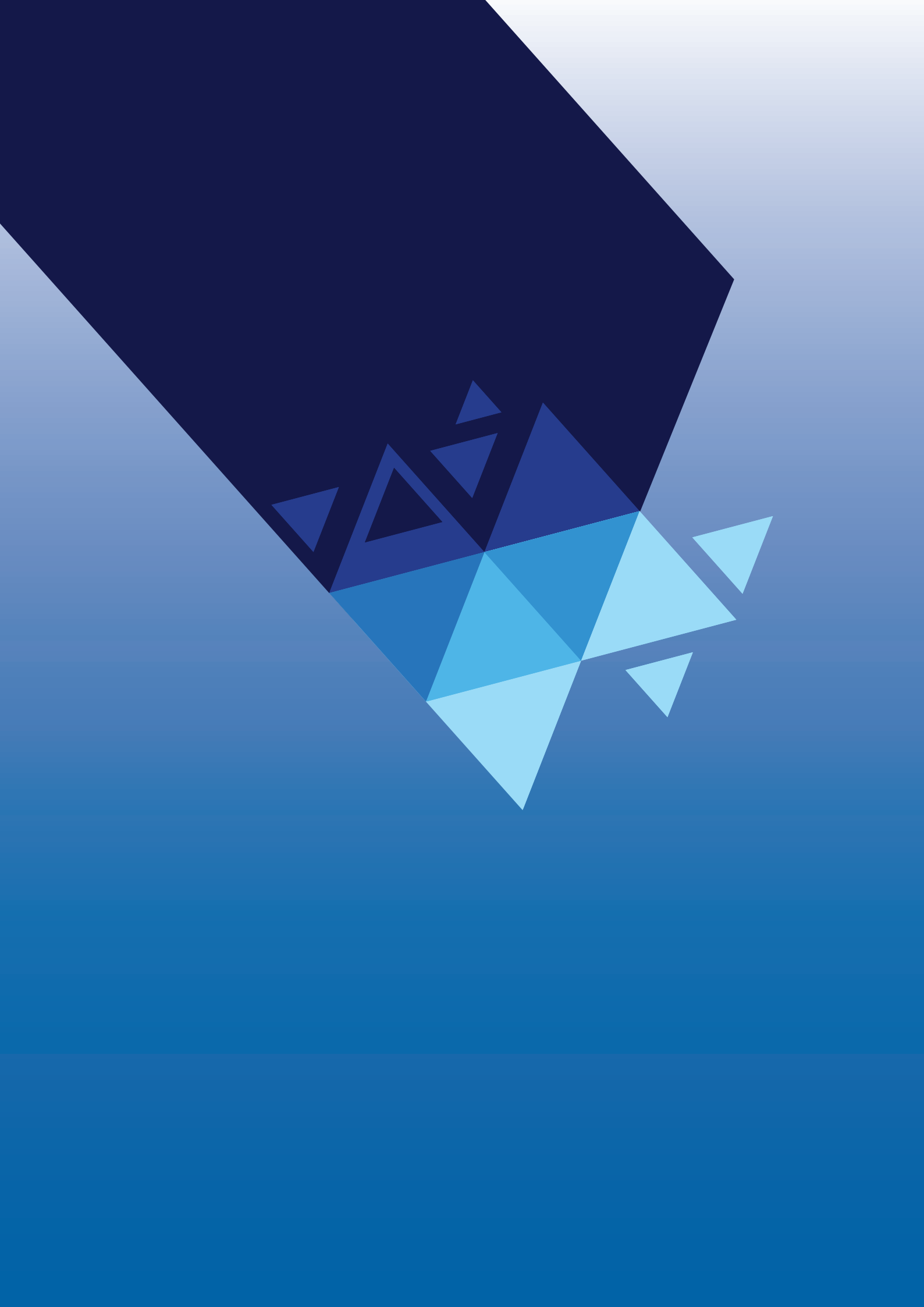
A useful starting point would be to focus on professional and other associations that are most involved in research and development of new technologies associated with AI. This would include, for example, data scientists and certain categories of coding and software engineers. Developing ethical guidance that is at least cognisant of existing ethical and professional obligations, and ideally builds on those obligations, would appear to be a practical way of fostering practices that uphold human rights.

PROPOSAL 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.**

The Australian Law Reform Commission or the Australian Human Rights Commission could be appropriate bodies to undertake an inquiry of this nature. In Chapter 8, the Commission proposes leadership for AI in Australia in the form of an AI Safety Commissioner. In the context of ethical frameworks applicable to AI, that proposed Commissioner also would be well placed to undertake this proposed research and consultation.



PART B: ARTIFICIAL INTELLIGENCE

CHAPTER 5
AI-informed decision making

5.1 Introduction

AI is increasingly used to make a broad range of decisions, with significant implications for our human rights. In the following chapters, the Commission considers the phenomenon of AI-informed decision making and how we can ensure human rights are effectively protected in this context.

Broadly speaking, AI describes 'the range of technologies exhibiting some characteristics of human intelligence'.[302] The term incorporates a cluster of technologies and techniques such as automation, robotics, machine learning and neural network processing. However, AI is not a term of art; it has no precise, universally accepted definition.[303]

As we witness AI disrupting our economic, social and governmental systems, it is difficult to overstate its impact or its potential. Understandably, countries such as Australia are fearful of being left behind in the 'AI arms race'.[304] At the same time, public concern about AI is growing. For instance, over half of people polled recently by Essential Media for the Commission were uncomfortable about government agencies using AI to make automated decisions.[305]

Over the past year, the Commission has asked the community how human rights are engaged by AI, how Australian law should protect human rights in respect of AI-informed decision making, and what non-legislative mechanisms are needed to protect human rights in the context of AI-informed decision making.

Drawing on this consultation, this chapter considers what we mean by the term AI; how AI engages human rights; and guiding principles for regulation in this area. AI-informed decision making is also defined to be a decision with legal, or similarly significant, effect for an individual, where AI materially assisted in the decision-making process.

The Commission proposes three principles to guide how the Australian Government and private sector engage in AI-informed decision making. First, international human rights law must be observed whenever an AI is deployed; secondly, there should be mechanisms in place to minimise harm and safeguard human rights; and, thirdly, AI should be accountable, with affected individuals being able to understand and challenge decisions made using AI they believe to be wrong or unlawful.

These principles are then applied in the following two chapters to AI-informed decision making, with a view to ensuring effective oversight and accountability for decisions made using AI, so that the community can be confident that our human rights are protected in this new domain.

5.2 What is Artificial intelligence?

(a) Defining AI

AI is not a new phenomenon; there are examples of work that could be characterised as AI going back to the 1940s. Russell and Norvig describe the modern development of AI:

> For thousands of years, we have tried to understand how we think; that is, how a mere handful of matter can perceive, understand, predict and manipulate a world far larger and more complicated than itself. The field of artificial intelligence, or AI, goes further still: it attempts not just to understand but also to *build* intelligent entities.[306]

AI should not be considered as one, singular piece of technology. Rather, it is a 'constellation of technologies'.[307] The OECD Group of Experts recently defined an AI system to include a

> machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. It uses machine and/or human-based inputs to perceive real and/or virtual environments; abstract such perceptions into models (in an automated manner, eg with ML or manually); and use model inference to formulate options for information or action. AI systems are designed to operate with varying levels of autonomy.[308]

While AI and algorithms have long histories, their capabilities have advanced at an unprecedented speed in recent years. There are several reasons for this, not least is the increasing availability of large datasets, as well as increased computing power, and new programming and data analysis techniques.[309]

There are generally considered to be two types of AI technologies:

- 'Narrow AI', refers to today's AI systems, which are capable of specific, relatively simple tasks—such as searching the internet or navigating a vehicle. It also encompasses 'algorithmic systems that analyse data and develop solutions in specific domains'.[310]
- 'Artificial General Intelligence' or 'technological singularity' is largely theoretical today. It would involve a form of AI that can accomplish sophisticated cognitive tasks on a breadth and variety similar to humans.[311] It is difficult to determine when, if ever, Artificial General Intelligence will exist, but predictions tend to be between 2030 and 2100.[312]

The focus of the Project, and this Discussion Paper, is on 'Narrow AI', where humans play a central role in designing and using the technology.

The use of the term AI has been subject to critique.[313] For instance, the Allens Hub for Technology, Law and Innovation observed that the term AI 'is not necessarily the most useful lens' given there is wide variation in the implications and risks associated with individual technologies that fall under the general description of AI. A system that uses pre-programmed logic, for example, will raise different issues from machine-learning tools that rely on patterns and trends in historic data.[314]

As 'AI' is a frequently used term, commonly denoting one or a combination of technologies and related products, services and applications, the Commission uses the term in this Discussion Paper.

(b) 'Big Data' and AI

Like the term AI, 'Big Data' is imprecise, but commonly used. It refers to the mass collection of personal data that differs from traditional data sets due to what some call the 'three Vs'—data that is large in volume, diverse in variety and moving with velocity.[315] It is 'big' in two ways: first, the quantity and variety of data available for processing; and, secondly, in the scale of the analytics that can be applied to that data. Both aspects depend on massive and widely available computational infrastructure.[316] Combined with new sources of data, the cost of collection, storage and processing in this context is also declining, resulting in 'a world of near ubiquitous data collection'.[317]

Not all data or information is the same. Different types of data are treated differently by Australian law. This Discussion Paper focuses especially on:

- *Personal information*, including sensitive personal information such as an individual's health information, credit information, an employee record or information related to a tax file number, or their criminal record. Personal information is regulated by Australian privacy and other laws.
- *Aggregated personal information,* which is a dataset consisting of personal information relating to some, usually many, individuals. However, the various sources of personal information have been combined (or aggregated) in a way that no individual's personal information is identifiable. Where personal information is aggregated in a way that strips the detail linking it to an individual—often referred to as de-identified or anonymised data—the resulting dataset will no longer be considered 'personal information' within the meaning of Australian privacy law. Such aggregated datasets are frequently used in AI applications to draw inferences about groups of people that share particular characteristics.

Of course, some data does not relate to individuals at all. This could be because the information has no significant connection to an individual (eg, the weather on a given day). Such information falls outside the scope of Australian privacy laws.

In its 2019 report, *A Day in the Life of Data*, the Consumer Policy Research Centre concluded that, taken together, the vast amount of personal and other non-sensitive data

> enables firms to develop increasingly detailed profiles of individuals. The processing of this data can then lead to inferred personal information such as: socioeconomic status, sexual orientation, political views, personality, mood, stress levels, health, personal interest, consumer worth or value or relationship status.[318]

Data, and especially personal data, is often described as the 'fuel' for AI. Large data sets are used, for example, to train machine-learning algorithms, and to generate profiles to support a determination made about a particular individual.

While this process can have a direct impact on an individual—the data itself may not be considered 'personal information'—the only type of information protected under Australian privacy law. This regulatory challenge is considered in Chapter 6.

5.3 What is AI-informed decision making?

This Discussion Paper focuses on the use of AI to make decisions in circumstances where an individual's human rights are engaged in a meaningful way.

Decision making can affect a person's human rights because of something about the *decision* itself, or because of the *decision-making process*. Therefore, in defining 'AI-informed decision making' for the purposes of this Discussion Paper, the Commission has considered both of these elements.

As explained in greater detail below, the Discussion Paper defines 'AI-informed decision making' to refer to:

- a decision that has a legal, or similarly significant, effect for an individual, and
- AI materially assisting in the process of making the decision.

(a) What types of 'decision' are within scope?

It is self-evident that some types of decision are likely to have an impact on an affected person's human rights more than others. Deciding whether to grant someone a social security benefit raises more acute human rights issues than deciding whether a person is eligible for a loyalty card.

It is natural that closer scrutiny, and tighter rules, should apply to decisions that are more likely to have a significant impact on individuals' human rights, because the stakes are higher with these areas of decision making.

Policy makers are increasingly grappling with how to apply this longstanding principle as AI is used more in decision-making processes.

For example, the GDPR's rules on automated decision-making apply to a decision that has 'legal effects' on an individual or 'similarly significantly affects' an individual.[319] The UK Information Commissioner's Office states that a decision under the GDPR produces 'legal effects' if it affects an individual's legal status or legal rights, such as the ability to access a social security benefit. A decision that has a 'similarly significant effect' has an equivalent impact on an individual's circumstances, behaviour or choices, such as the automatic refusal of an online credit application.[320]

It should be acknowledged that the GDPR is not a perfect model. In particular, Article 22 of the GDPR, referred to above, sets certain limits on how wholly automated decisions are made, with the aim of protecting personal information. This is a narrower focus than the one adopted by the Commission. Nevertheless, the GDPR definitional approach, which focuses on the legal or other significant effect of a decision, is instructive here, because it is likely to bring within scope many types of decisions that are likely to engage human rights more broadly.

With this in mind, the Commission has defined 'decision' in the context of 'AI-informed decision making' to be any decision that has a legal effect, or similar significant effect, for an individual. This would include decisions made by government, such as administrative decisions, and decisions made by non-government entities like corporations.

(b) When will a decision be *informed* by AI?

There are many ways in which AI might be used in a decision-making process. In some cases, the use of AI will be central to how the decision is made, and on the ultimate decision. But for others, AI will have only a trivial impact on the decision-making process.

As the Commission is concerned with the former scenario, this Discussion Paper uses the term 'AI-informed decision making' to refer to decision making that is *materially assisted by the use of AI*.[321] This includes decisions made by humans who rely on data points generated by an AI product or tool, as well as fully-automated decisions.

(i) What types of AI are most relevant to decision making?

A range of specific technologies and processes can be used in AI-informed decision making. These technologies and processes can be used individually, or in combination, to make inferences, predictions, recommendations or decisions. For example:

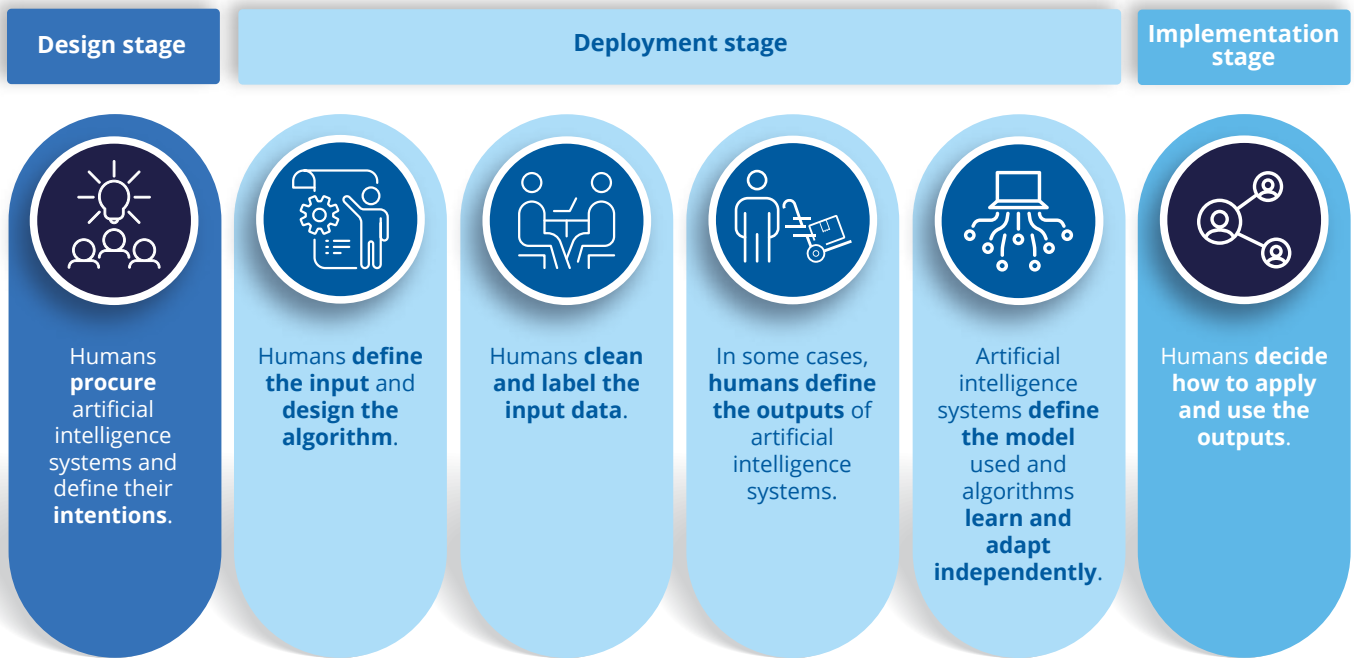
- *algorithms*, *mathematical formulae and computer code* are all 'designed and written by humans, carrying instructions to translate data into conclusions, information or outputs'[322]
- *machine learning* involves a computer program learning to undertake defined tasks using numerous examples in a dataset, detecting patterns in the examples. The computer program then uses those patterns to make inferences, predictions, recommendations or decisions for a new case or situation that was not included in the original data set.[323] This process can be:
  - » unsupervised learning, where the system is fed a dataset that has not been classified or categorised, and it identifies clusters or groupings
  - » supervised learning, where a system is fed a categorised or classified data set that it uses to learn how to complete a task as instructed, or

» reinforcement learning, where a system will take action within a set environment and assess whether that action will achieve its goals, and learn, through this process, which action to take to achieve those goals.[324]

- *automation* in a decision-making process involves a computational system applying algorithms or other rules to particular fact scenarios in an automatic way. A decision-making system can be wholly automated, in which case it produces decisions without human involvement, or the system can produce inferences, predictions or recommendations, which a human will use to make a final decision.

As outlined above, the Commission is focused on forms of decision making that involve at least *some* level of human involvement—in designing and deploying the AI-powered system; in assessing datapoints generated by AI; in overseeing the decision-making system; or in some combination of all of these activities. An example of such an AI-informed decision-making process, relying on machine learning, is outlined in Box 2 below.[325]

Box 2: AI-informed decision-making process[326]



(ii) Has AI materially assisted the decision to be made?

The Commission is focused on decisions where the use of AI is material or significant in the decision-making process. Materiality is a concept well defined in other sectors.

There are two main scenarios here. The first is relatively simple: where all key elements of the decision-making process are automated. The use of AI in such a case is clearly material.

The second scenario arises where AI is used in the decision-making process to generate a data point that bears in a material or significant way on the ultimate decision. In other words, in this scenario, a human decision maker relies on an AI data point that has been generated to make the decision itself or to determine something that was significant in the ultimate decision.

For example, imagine a human decision maker (X) must weigh three factors in making a decision. If X relies on an AI tool in forming a view on one of these three factors, it is likely that AI was material or significant in the decision-making process. Now imagine X types up the decision using a sophisticated word-processing application that was developed using AI. If the function of the word-processing application is simply to record the decision, not to assist in weighing the three factors that bear on the decision, X's use of this word-processing application would not be a material use of AI.

5.4 How human rights are engaged by the use of AI

As noted several times in this Discussion Paper, AI can have a dual impact, promoting human rights in one context, but impinging on human rights in a different context. Drawing especially on the Commission's consultation to date, this section assesses some ways in which the use of AI can engage human rights.

(a) Submissions and consultations

While the rights to privacy and non-discrimination have often dominated public discourse on AI, submissions to the Issues Paper pointed out that many other human rights are also engaged, both positively and negatively, by AI technologies, including other civil and political rights, and economic, social and cultural rights.

Many stakeholders noted AI's potential to transform society for the better, promoting human flourishing and more inclusive communities. Submissions also made clear that there are significant risks associated with AI, and that negative impacts tend to be experienced disproportionately by those already disadvantaged or vulnerable.

(i) Use of AI by government

Government is increasingly using AI to support or make decisions, especially in delivering services.[327] Some commentators predict a revolution in government service delivery that will redefine the citizen's interaction with government.[328]

AI is being used in decision making where the potential impact for individuals is high, including in relation to housing, health, criminal justice and policing. In such contexts, the likelihood of engaging human rights—and, if poorly implemented, having an adverse effect on human rights—is also high. This is particularly evident in the digitisation of essential government services.[329] Typically, guidance on how AI should be used by government does not refer to human rights.[330]

Stakeholders identified examples of how government use of AI can promote human rights:

- using public health data to improve diagnostics, personalise medical treatment and prevent disease[331]
- removing human bias in decision making[332]
- improving access to justice by making it more affordable and accessible, particularly for vulnerable communities that have historically faced significant obstacles in accessing legal remedies.[333]

Stakeholders also noted examples of government using AI in ways that may adversely affect human rights, such as:

- the use of an automated system to promote debt compliance, often referred to as 'Robodebt', described in more detail below[334]
- the use of facial recognition technology, in policing for example,[335] or as a requirement to access government services.[336]

These uses of AI engage, among others, the right to privacy, the right to equality and non-discrimination, and the right to equality before the law.

AI is increasingly being considered in government service delivery, including overseas as a tool to detect children at risk of maltreatment in order to target protection interventions.[337] This can engage the right to privacy, the right to work, the right to a family life and a number of children's rights, among others.

Case study:

AI used in social services

A number of submissions referred to Centrelink's use of an automated debt recovery system as an example of AI-informed decision making that engages a number of human rights.[338]

The system, which some have called 'Robodebt', used an algorithm to identify discrepancies between an individual's declared income to the Australian Taxation Office, and the individual's income reported to Centrelink. An algorithm was used to compare the two debts; where a discrepancy was identified, this was treated as evidence of undeclared or under-reported income, and a notice of debt was automatically generated and sent to the individual.[339]

It subsequently became apparent that this process was resulting in the generation of some inaccurate debt notices,[340] which had a particular impact on a number of recipients who were already disadvantaged or vulnerable.[341] The Senate Standing Committee on Community Affairs received evidence of 'many personal accounts of the stress and distress' the system caused to Centrelink recipients.[342]

The Committee concluded 'the system was so flawed it was set up to fail'.[343] Many of the problematic aspects of the debt recovery program related to how the system was rolled out, including the lack of information, and difficulty accessing information about how to challenge or seek review of a debt nominated in a debt recovery letter.[344] On 19 November 2019, the Minister for Government Services, the Hon Stuart Robert MP, announced the cessation of wholly automated debt discrepancy notices, and a review of all debts identified through the use of the algorithm.[345]

The use of AI in the delivery of government services like social welfare engages a range of human rights. The right to social security, for example, is protected in international human rights law.[346] The Committee on Economic, Social and Cultural Rights (CESCR) has concluded this right imposes an obligation on governments to ensure that eligibility criteria for social security benefits are 'reasonable, proportionate and transparent'.[347] Further, any 'withdrawal, reduction or suspension' of social security benefits should be circumscribed and 'based on grounds that are reasonable, subject to due process, and provided for in national law'.[348]

Case study:

Use of AI technologies in criminal justice

AI is increasingly being used in the criminal justice system, including through predictive policing and risk assessment tools, and facial recognition technology. These tools are being used to predict or determine who, when and where crimes are, or will be, committed, and to assist in deciding sentencing and bail. Several submissions emphasised how such AI-powered tools engage human rights.[349]

Some stakeholders identified the adverse impact on human rights by the NSW Police Suspect Target Management Plan (STMP).[350] The STMP is a risk assessment tool that identifies individuals at high, medium or low risk of offending. It is understood that young people identified using the STMP are then made the subject of more frequent pre-emptive policing strategies, such as stop and search. Evidence has shown that the program has a disproportionate impact on Aboriginal and Torres Strait Islander peoples,[351] something that a number of stakeholders also noted with concern.[352] While the STMP has been described as an algorithm,[353] little is known about its use. Giving evidence to a NSW parliamentary committee in 2016, the NSW Police Commissioner stated that the STMP 'is based on a predictive style method of policing' that involves a risk assessment template that helps us identify those potential recidivist offenders and it puts in place some strategies in terms of trying to disrupt their activity to minimise opportunities for them to commit crimes.[354]

Similar risk assessment tools powered by AI are increasingly being used by police forces in comparable jurisdictions.[355] The impact of this kind of risk-based model to support predictive policing software overseas is beginning to be evaluated.[356] The lack of transparency regarding how such algorithms operate has been challenged in the courts. For instance, the New York Police Department's refusal to provide records regarding the use and testing of predictive policing technologies was challenged by the Brennan Centre for Justice in 2017.[357]

The use of AI in the criminal justice system engages a number of human rights, including the right to equality or non-discrimination, the right to equality before the law, the right to life, personal security and liberty, the right to privacy, the right to a fair public hearing and the right to procedural fairness and due process, including the presumption of innocence.

(ii) Use of AI by corporations and other non-government bodies

Where companies and other bodies outside government develop and use products and services that rely on AI, this also can have positive or negative human rights effects. Stakeholders referred to advances in human rights, often in ground-breaking ways, through commercial applications of AI. For example:

- autonomous vehicles have the potential to increase safety by removing human driving error,[358] and enabling people who otherwise cannot drive because of their age or disability
- some AI-powered recruitment tools are said to remove human bias and better match applicants to job criteria[359]
- AI can be used to improve our response to natural and humanitarian disasters.[360]

Examples of negative impacts on human rights include:

- the development of 'deepfake' videos, news and audio leading to social and political manipulation[361]
- content moderation by, and of, social media platforms, which can control the flow of information and potentially undermine freedom of expression by removing online content[362]
- exploitative and discriminatory marketing practices, causing harm to a number of different groups, such as children[363]
- intrusive surveillance practices in the workplace[364]

- predictions of creditworthiness in banking[365] and predictive analysis in insurance,[366] both of which may potentially have an impact on human rights if, for example, the decision-making system is affected by algorithmic bias.

Case study:

Using AI to target or exclude particular groups

The use of AI to exclude or target a group sharing a characteristic, such as age, gender, or racial or ethnic background, engages a range of human rights.

This can be deliberate. There have been examples of AI being used in advertising on social media platforms, which allows organisations to target advertisements for housing, employment and credit by reference to race, gender, age or postcode.[367]

Even where this targeting is not deliberate, AI can operate inadvertently to exclude particular groups, often by reference to historical data that is itself skewed. One high-profile example involved an AI-powered recruitment tool that 'learned' to put forward predominantly male candidates as potential employees.[368] Both of these examples can result in people being denied an opportunity due to a particular characteristic, such as their gender or race. This engages rights such as the right to non-discrimination[369] and the right to work.[370]

AI also has been used to analyse individuals' activity in social media and elsewhere to develop personal profiles of those individuals. Such profiles allow inferences to be drawn regarding an individual's preferences for products and services, as well as the individual's political and other views. This information can then be used by companies and others in how they communicate with those individuals. While there are innocuous uses for this application of AI, there are also ways in which it can exploit and manipulate individuals.[371] A high-profile example of the associated risks is the Cambridge Analytica controversy.[372] Stakeholders emphasised how this use of AI engages a range of human rights including privacy, freedom of expression and association and the right to non-discrimination.[373]

Case study:

Use of data in health

AI is increasingly used in the health sector. There is demonstrated potential for AI to improve health care, including through more accurate and speedy diagnosis, treatment and management of diseases, and in planning and resource allocation in the health sector.[374] Some stakeholders urged better regulatory oversight to promote ethical and accountable clinical practice in areas where AI research has been focused, such as radiology.[375]

There is also potential for human rights to be breached in this context. For example, many applications of AI will make use of health data in ways that may be unknown at the point of collection. Several stakeholders drew attention to how health data, including genetic data, can be used to discriminate unlawfully.[376] Submissions also raised concerns about the My Health Record system and lack of safeguards against inappropriate access to sensitive personal information.[377]

5.5 AI and regulation

The Australian Government must ensure that Australian law protects and promotes human rights, including by requiring corporations and others to adhere to human rights.[378] This obligation applies to the use of AI, just as in any other area. The Issues Paper asked what principles should guide any regulatory response to the rise of AI.

(a) Submissions and consultations

(i) Support for regulation despite the challenges

There was strong support among stakeholders for an appropriate regulatory response to AI.[379] One submission stated that the need for a proactive regulatory response to AI technologies is an 'inescapable challenge for contemporary societies'.[380]

Many stakeholders emphasised that 'hard' regulation—that is, laws accompanied by effective enforcement mechanisms—should play the primary role in ensuring AI is used in a responsible or accountable way. This could be supported by 'soft law', including co- and self-regulatory reforms and guidance.[381]

As outlined in Chapter 6, stakeholders referred to various laws, guidelines and regulatory oversight bodies that might be relevant to AI, as well as identifying significant gaps in the regulatory framework, or identifying particular problems that currently have no regulatory solution.

Generally, there was concern that the current approach to regulation made 'avoidable harms' more likely,[382] and that the information asymmetries between companies and public regulators regarding AI undermine much-needed collaboration to prevent harm.[383] It was noted that AI created for a specific, benign purpose can nevertheless be put to an alternate, destructive or problematic use.[384] Other submissions simply urged a precautionary approach to deal with the problems posed by AI; that is, action should be taken where it has been identified that human use of AI might cause harm.[385]

Several submissions identified that meaningful stakeholder participation was important for the regulation and design of technology.[386] Some emphasised the importance of genuine participation of stakeholders from vulnerable groups who can be most affected and disadvantaged by AI technologies.[387]

(ii) Addressing the challenges of regulation in this area

Stakeholders identified difficulties in regulating uses of AI. Those difficulties include: the varied nature of the technologies that fall within the scope of AI; the speed of technological change; and the fact that AI is being used in many different contexts.[388]

A further difficulty is in selecting the most appropriate 'regulatory object'. That is, while there have been calls for better regulation of AI *as technologies*, in many instances that is likely to be impossible or ineffective.

For example, researchers at the University of Melbourne cautioned against developing laws that are AI-specific, arguing that law or regulation should apply without reference to AI technologies per se:

> There is no reason why a law or regulation should apply to an AI algorithm but not to other algorithms that are tasked with the same decisions. While there may be additional challenges of opacity applying to modern machine learning techniques, the regulatory response should be independent of algorithmic implementation details.[389]

This point was also made in a recent challenge in the United Kingdom to the use of facial recognition technology by South Wales Police. The court noted:

> The fact that a technology is new does not mean that it is outside the scope of existing regulation, or that it is always necessary to create a bespoke legal framework for it.[390]

Some stakeholders suggested that given the challenges of regulating a specific technology, a preferable approach is to 'focus on the goals that regulatory efforts aim to achieve and confer sets of rights and remedies to that effect'.[391] Another way of putting this would be that it may be more appropriate to regulate the activities in which AI is used or contemplated, in a way that is mindful of the role that AI could play in those areas, as distinct from seeking to regulate AI or its constituent technologies directly.

Some stakeholders were wary about the risk of regulation stifling innovation or weakening the digital economy.[392] Stakeholder objections ranged across the spectrum, from objection to *any* form of regulation in this area, to more specific concerns that regulation that may not properly account for the undoubted challenges in this area.[393]

Stakeholders argued that responsible innovation in this area has a range of advantages including the protection of human rights[394] and economic benefits.[395] Some submissions emphasised the need to balance the benefits of innovation with respect for human rights.[396]

This could include encouraging adaptive regulation and governance models grounded in protecting human rights that do not disproportionately inhibit technological innovation.[397] Some pointed to the prospect of technological innovation being driven, rather than hindered, by positive protection of human rights, such new forms of medical care.[398]

(b) A human rights approach

Stakeholders generally supported the idea that human rights should be central in how we consider AI, and how we regulate in this area.[399] Some argued that a human rights approach can guard against the risks posed by AI, such as the 'potential to lose sight of the fundamental dignity of human beings' when people are treated as 'data points rather than individuals'.[400]

Submissions also recognised the practical support the human rights framework provides to determine whether limitations on human rights by technology applications are appropriate and acceptable.[401] Some submissions recognised the importance of regulation being 'adaptive' or 'anticipatory', but still thought that it should 'remain grounded in protecting human rights'.[402]

(c) Guiding principles for regulation

As outlined in Chapter 3, some submissions made the case for 'principles-based regulation', with the aim of producing rules that are sufficiently flexible to meet the demands of new and emerging technologies.[403] Some suggested such principles should emphasise particular approaches, such as privacy by design,[404] universal design,[405] or 'human rights by design'.[406]

Some stakeholders argued that regulation in relation to AI specifically should be guided by principles such as:

- **transparency** regarding the use of AI to make or support a decision about an individual,[407] including notification of that use to the individual concerned[408]
- **trust**, including in relation to how personal data being collected may be used in future in AI-informed decision making[409]
- **fairness**, including by avoiding bias and discrimination in the development and use of AI technologies,[410] and for private entities to be proactive in their monitoring and evaluation of algorithms[411]
- **mitigation of risk**, including through the design and management of AI technologies,[412] such as identifying potential bias in a data source[413]
- **responsibility**, including by ensuring AI technologies are developed and used for social good.[414]

There have been many recent attempts to articulate ethical frameworks and other such high-level guidance on the design, development and use of AI, as discussed in Chapter 4.

The OECD AI Guidelines, for example, identify 'five complementary values-based principles for the responsible stewardship of trustworthy AI', including that AI should be:

- beneficial to people and the planet
- designed in a way that respects the rule of law, human rights, democratic values and diversity
- transparent, ensuring people understand AI decision making and can challenge a decision
- continually assessed and managed.

Accountability for compliance with these principles rests with those 'organisations and individuals developing, deploying or operating AI systems'.[415] Subsequently, the G20 Trade And Digital Economy Ministers adopted a 'human-centred approach to AI, guided by the G20 AI Principles', which were themselves drawn from the OECD.[416]

Similarly, in 2019 the Council of Europe's Commissioner for Human Rights developed a ten-point plan for human rights compliant design and use of AI, declaring that 'ensuring that human rights are strengthened and not undermined by artificial intelligence is one of the key factors that will define the world we live in'.[417]

The ten-point plan provides guidance for the mitigation and prevention of the negative impacts of AI on human rights. The plan requires, for example:

- legal frameworks to be established incorporating human rights impact assessments, independent and effective oversight, and access to effective remedies
- ensuring the development and use of AI is transparent and non-discriminatory and in compliance with state obligations to protect human rights
- public consultation on AI systems and the promotion of AI literacy across all levels of government and among the general public.[418]

In July 2019, the Berkman Klein Center for Internet and Society at Harvard University published a preliminary report mapping 32 'principles documents' for AI published by governments, companies advocacy groups and multi-stakeholder initiatives.[419] The Center's definition of 'principles' included 'normative and general declarations about how AI *ought* to be developed, deployed, and its usage regulated'.[420] There were 47 principles in total, categorised into eight themes that include human rights.

5.6 The Commission's preliminary view

(a) AI-informed decision making

It is clear from the Commission's consultation, and research more broadly, that the impact of AI on human rights is already profound, and it is growing. The Commission has identified immediate areas of concern, possible solutions, gaps in the regulatory framework and has noted some persuasive calls for innovative solutions to complex problems.

AI can be applied to almost every conceivable area of human activity. This is not an abstract observation: there seems no limit to the appetite to develop, or at least consider, AI applications for use in the real world. Consequently, any attempt to articulate an all-encompassing response to AI seems bound to fail. It is highly likely that such a response would be unsuitable, at best, for the myriad and diverse contexts in which AI may be used.

As a result, the Commission has a narrower, though still ambitious, focus in respect of AI.

This Discussion Paper considers the use of AI to make decisions that have legal or similar effects on people. AI, when properly deployed, can improve some forms of decision making, by making it faster, more data-driven and efficient.

However, there are also significant risks to human rights. Those risks vary according to the design of the relevant decision-making system; and the role, if any, of a human in the decision-making process.

The context of the decision is also important. Where AI is used in decisions that have a significant impact on humans—such as in sentencing criminal offenders, in determining whether an individual is a refugee or in home loan decisions—the stakes are high and the consequences of error can be grave for anyone affected.

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?

For good reason, AI is at the centre of the Fourth Industrial Revolution. The social, economic, political and other impacts of AI are unprecedented and growing. This applies also to the use of AI in decision making that affects individuals' human rights.

The Commission considers that Australia needs to re-think its regulatory approach to AI, especially as AI is used in decision making. A central objective for regulation in this area should be the promotion and protection of human rights. It is legitimate for government regulatory policy to pursue other objectives at the same time, including to seize economic opportunities presented by AI. However, human rights must be at the centre.

So too with AI. Seeking to regulate AI *as a cluster of technologies* almost certainly would be ineffectual. However, we can regulate how AI is developed and especially used in particular contexts. Equally importantly, we can take steps to ensure that the existing laws that protect human rights in Australia are applied effectively wherever AI is used.

This reflects the approach taken in some other jurisdictions. In the UK, for example, the House of Lords Select Committee on Artificial Intelligence concluded in 2018 that, at present, blanket AI-specific regulation would be inappropriate. Instead, it recommended that existing sector-specific regulators and the UK Centre for Data Ethics and Innovation work to identify the gaps in existing regulation.[421] The UK Government has set up a cross-government Ministerial Working Group on Future Regulation to consider some of these issues.[422]

(b) Framework for regulatory reform

This Discussion Paper primarily applies international human rights law; it also refers to authoritative guidance for private sector entities, including the UN's Guiding Principles for Business and Human Rights.

The international human rights law framework is a body of substantive norms that are almost universally accepted as an authoritative legal standard.[423] Human rights law therefore provides an excellent framework for a regulatory response to AI. This position reflects a growing trend internationally among experts seeking to grapple with the challenges presented by AI.[424]

Over the last 70 years, international human rights law has been remarkably adaptable to a diverse range of contexts. It has provided important guidance in protecting humans from a wide range of harms that have changed over time.

Many stakeholders supported the idea that the primary aim of any regulatory response should be to apply international human rights law to AI-informed decision making.

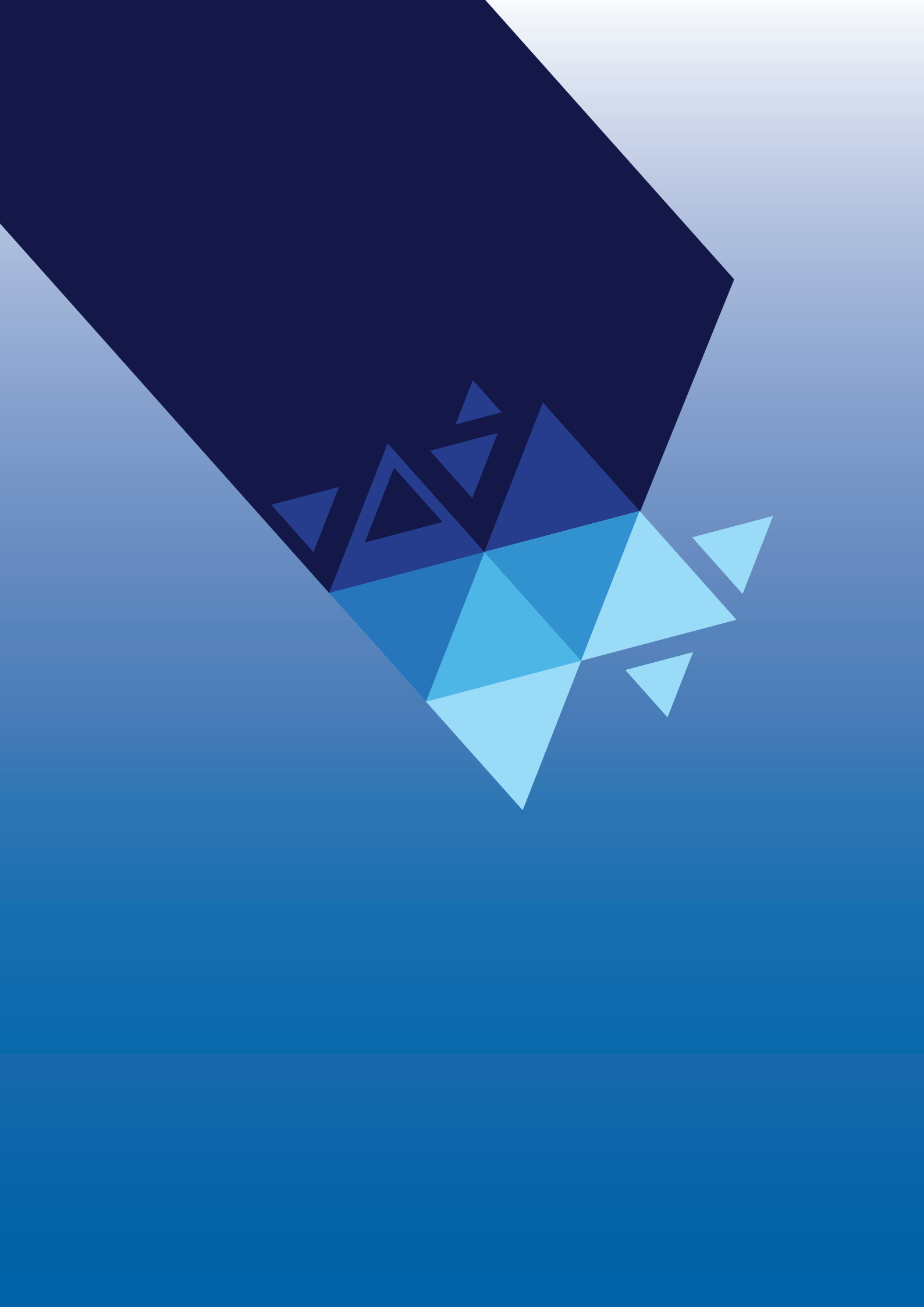
While there might be some gaps in that existing body of international human rights law which need to be filled, it is not apparent that an entirely new human rights framework needs to be developed to address the rise of AI.[425]

The Commission cautions against starting a new debate about ideas such as 'fairness' in AI-informed decision-making systems in a way that pays insufficient regard to the contribution of international human rights law in this area.[426]

International human rights law also requires Member States to put in place a framework to provide an effective remedy where there has been a human rights violation. Effective remedies include judicial and administrative remedies, such as ordering compensation or an apology, as well as preventive measures that may include changes to law, policy and practice.[427] Effective remedies for human rights breaches fall under the accountability principle, which is central to a human rights approach.[428]

In the following chapters, the Commission seeks to apply three key principles to how the Australian Government and private sector should design, develop and use AI in decision making that affects people's human rights:

- **International human rights should be observed.** The Australian Government should comply with human rights in its own use of AI, and it should also ensure that human rights protections are enforced for all entities that use AI.
- **AI should be used in ways that minimise harm.** There need to be appropriate and effective testing of AI-informed decision-making systems, before they are used in ways that could harm individuals, and ongoing monitoring of those systems when they are in operation.
- **AI should be accountable in how it is used.** Individuals affected by AI-informed decisions should be able to understand the basis of the decision and be able to challenge decisions that they believe to be wrong or unlawful. Accountable AI-informed decision making is discussed in detail in Chapter 6.



CHAPTER 6

Accountable AI-informed decision making

6.1 Introduction

Accountability is critical to the protection of human rights in all forms of decision making. This chapter considers how to ensure accountability for AI-informed decision making.

As previously observed, AI can improve the way decisions are made by enabling powerful insights to be drawn from large datasets. These insights can make for more accurate, or better informed, decisions. Automation can increase efficiency—in decision making, as with many other activities. It is little wonder, therefore, that AI-informed decision making, including through automation, has been embraced by the public and private sectors across a range of domains.

However, the use of AI can also lead to decisions that breach human rights and cause other harms. Accountability is fundamental to ensuring that those risks are addressed.

Too often, there is an illusion that the use of AI is unregulated. The Commission considers that the first priority, in ensuring that AI-informed decision making is accountable, is to apply existing law rigorously, including laws that protect human rights. To this end, the High Court of England and Wales stated in 2019:

> The fact that a technology is new does not mean that it is outside the scope of existing regulation, or that it is always necessary to create a bespoke legal framework for it.[429]

While the effective application of existing law will go some way to achieve accountability in this area, there are also some gaps in the law. AI-informed decision making raises some novel issues that go to the heart of accountability. The Commission has reached a provisional conclusion that in order to be accountable, AI-informed decision making must be:

- *lawful*, complying with existing laws and having legal authority where necessary
- *transparent*, encompassing the notion that affected individuals are notified of AI being a material factor in a decision engaging their human rights, as well as transparency regarding government use of AI
- *explainable*, requiring a meaningful explanation for an AI-informed decision
- *used responsibly* and with clear parameters for liability
- subject to appropriate *human oversight and intervention*.

In this context, this chapter includes proposals for targeted reform to ensure that human rights are adequately protected, where AI is used in decision making in areas that carry a significant risk of harm.

6.2 The elements of accountability

Under international human rights law, knowing a decision has been reached, and how that decision has been made, is fundamental to an individual being able to challenge a breach of human rights, and seek an effective remedy.[430] These elements of accountability are also central to the functioning of any liberal democracy.

A human rights approach to accountability requires effective monitoring of compliance with human rights standards and goals, as well as effective remedies for human rights breaches. For accountability to be effective, there must be appropriate laws, policies, institutions, administrative procedures and mechanisms of redress.[431] Accountability, especially in the context of human rights, includes both a *corrective* function, facilitating a remedy for when someone has been wronged, as well as a *preventive* function, identifying which aspects of a policy or system are working and what needs adjustment.[432]

Applying these accountability principles to AI-informed decision making is the subject of debate among experts and others. A common theme in the Commission's consultation has been that regulation generally should combat 'black box' or opaque decision making using AI, and thereby promote accountability.[433]

As explained in this chapter, the Commission has reached a provisional conclusion that for AI-informed decision making to be accountable, it must be:

- *lawful*, complying with existing laws and having legal authority where necessary
- *transparent*, encompassing the notion that affected individuals are notified of AI being a material factor in a decision engaging their human rights, as well as transparency regarding government use of AI
- *explainable*, requiring a meaningful explanation for an AI-informed decision
- *used responsibly* and with clear parameters for liability
- subject to appropriate *human oversight and intervention*.

Each of these elements is addressed in turn.

6.3 Lawfulness

Any decision that affects a person's rights or interests must comply with the law. This applies both to decisions made by government and those made by non-government entities, such as corporations. What it means to make a lawful decision—in other words, the specific legal requirements that must be followed—varies depending on the decision being made and who is responsible for the decision.

Some common legal requirements apply to almost all decision making. For example, it is unlawful for anyone to make a decision that discriminates on the basis of an individual's race, sex, age or other protected attribute. There are few exceptions to this rule.

In addition to these common requirements, some types of decision making carry their own specific legal requirements. For example, a specific body of law regulates medical decision making, with the aim of avoiding harm to patients.

Moreover, the law differentiates between government and non-government decisions.[434] On the whole, government decision making is more highly regulated. Laws dictate what decisions may be made by government, how they may be made, with opportunities for those affected by government decisions to seek review.

For example, the *Administrative Decisions (Judicial Review) Act 1977* (Cth) (ADJR Act) regulates many areas of Australian Government decision making. A decision covered by the ADJR Act must, for example: comply with natural justice (or procedural fairness); follow the decision-making procedure set out in law; and not be unreasonable.[435] Where a decision is unlawful, a court exercising judicial review can require the law to be followed.

As a general rule, the legal requirements applicable to a decision-making process apply regardless of the way in which the decision is made.

For example, imagine a decision maker (DM) must offer a hearing to someone (X) before deciding whether to give them a social security benefit. If DM decides to do their own investigation into X's situation, or if DM asks an assistant to do that investigation, or if DM does no investigation at all, the requirement to offer X a hearing will continue to apply. Similarly, if DM decides to rely on AI to help make DM's decision, DM will still be required to offer X a hearing.

However, as the Hon Justice Melissa Perry has observed, the increasing prevalence of automated government decision making can threaten the administrative law values underpinning Australia's democratic society governed by the rule of law.[436] Justice Perry states:

> It is not difficult to envisage that the efficiencies which automated systems can achieve, and the increasing demand for such efficiencies, may overwhelm an appreciation of the value of achieving substantive justice for the individual. In turn this may have the consequence that rules-based laws and regulations are too readily substituted for discretions in order to facilitate the making of automated decisions in place of decisions by humans.[437]

(a) Applying existing law to AI-informed decision making

While AI-informed decision making is subject to the usual legal requirements that apply to decision making that does not involve AI, some laws are particularly pertinent. In particular, the Commission's consultation highlighted the importance of anti-discrimination law (see below)[438] and privacy law (see below).[439]

Stakeholders also identified existing laws that are especially important in specific decision-making contexts, such as take-down provisions for social media platforms;[440] intellectual property and copyright law;[441] social security law;[442] and consumer law implications of data use and sharing.[443] Some stakeholders suggested potential new legal issues, such as the application of legal personhood to various forms of AI.[444]

AI-informed decision making creates challenges for existing legal frameworks. Where there are gaps, law reform may be necessary to ensure AI-informed decision making remains accountable.

Stakeholders pointed to several gaps in the legal framework for AI-informed decision making, including the lack of federal protection for human rights in the form of a Human Rights Act or charter.[445]

Some stakeholders also suggested other reform, including in federal anti-discrimination and privacy law. For example, the evidentiary onus often rests on an individual claiming discrimination, or seeking to challenge an automated decision, with some suggesting this onus be shifted to those deploying an AI-informed decision-making system, particularly where the decision-making process is opaque.[446] Several submissions drew on examples from other jurisdictions, such as Europe's GDPR, which might improve the regulation of AI-informed decision making in Australia.[447]

In addition, some stakeholders suggested that specific laws might be needed to prohibit altogether, or to closely regulate, the use of AI for particular types of AI-informed decision making—especially refugee status determinations,[448] the use of autonomous weapons[449] and facial recognition technology in specific contexts, such as policing.[450]

The GDPR is a rare example of a law that has been developed specifically with AI-informed decision making in mind. Another example, from the United States, is the proposed 'Algorithmic Accountability Act', tabled in the US Congress in April 2019.[451] The Bill proposes that some organisations conduct impact assessments in relation to data protection and algorithmic decision making under the authority of the US Federal Trade Commission.[452]

(i) Anti-discrimination law and AI-informed decision making

It is unlawful to discriminate on the basis of 'protected attributes', which are characteristics such as race or ethnic background, gender, age and disability. This human right to equality and non-discrimination is contained in all the major international human rights treaties, and in domestic Australian law.[453]

Bias or prejudice can cause unlawful discrimination. The use of AI can assist in identifying and addressing bias or prejudice that can be present in human decision making.[454] Conversely, AI can also perpetuate or entrench such problems. A steady flow of examples is emerging of AI being used to make apparently discriminatory decisions regarding sentencing, advertising, recruitment, healthcare, policing and elsewhere.

AI-informed decision making can be:

- *directly* discriminatory, where someone is treated less favourably because of a protected attribute. The deliberate exclusion of certain individuals from seeing a job advertisement because of their age or gender, for example, is likely to be discriminatory.[455]
- *indirectly* discriminatory, such as when an unreasonable rule or policy is applied that is the same for everyone, but has an unfair effect on people who share a particular protected attribute. One way this can happen is where information, such as the postcode where an individual lives, can be a proxy, or indicator, for a protected attribute, such as the individual's ethnic origin.[456] If a decision is made by reference to that proxy, and the decision unfairly disadvantages members of that ethnic or other group, it could lead to indirect discrimination.

Discriminatory outcomes from AI-informed decision making may be difficult to detect or predict. In a 2018 empirical study, for example, academic researchers found that an advertisement for science, technology, engineering and maths (STEM) jobs, which was intended to be gender-neutral, ended up being shown to far more men than women.

While the advertiser did not intend to exclude women from seeing the advertisement, the advertiser did require that the advertisement be shown in a cost-effective way. The researchers found that this requirement of the advertiser led to women being excluded from seeing the advertisement: the algorithm was applied in a way that caused women to be a more expensive demographic to advertise to.[457]

More research is needed about how AI-informed decision making can lead to unlawful discrimination. Stakeholders have questioned whether current law will be effective in detecting or preventing discrimination.[458]

There are practical challenges to applying current laws. It will be difficult, if not impossible, to establish discrimination where the decision-making system is opaque, or identifying whether the combination of potentially thousands of variables have been used by an algorithm in a way that treats an individual differently on the basis of a protected attribute. Experts have begun to develop technical solutions to combat the potential for discrimination in AI-informed decision making, including 'pre-processing methods', referring to sanitisation of training data to remove potential bias; 'in-processing' techniques, involving modifying the learning algorithm; and 'post-processing' methods, involving auditing algorithmic outcomes to identify and resolve discrimination patterns.[459]

(ii) Privacy law and AI-informed decision making

Much public concern regarding AI-informed decision making has focused on the right to privacy.[460]

Stakeholders referred to privacy risks posed by activities associated with AI-informed decision making. Some of these relate to the large amounts of information stored by those responsible for AI systems; some relate to the operation or use of the systems themselves. Examples include:

- the intrusion of surveillance technology, which may be used to inform policing and other decisions
- the fallibility of anonymising personal information

- the over-reliance on individual consent, including where consent is not informed or freely given, to justify activity that would otherwise breach an individual's right to privacy
- the risks posed by the commercial exploitation of personal data
- the widespread collection and use of personal data by large private technology companies.[461]

Seemingly innocuous personal data can be used, especially in an AI-powered system, to gain insights about an individual, including on sensitive matters. For example, in 2018 a US court accepted that smart meter data, which records when and how energy is used in a home, may allow 'intimate personal details' of an individual's life to be established, supporting inferences 'about a person's lifestyle including their occupation, religion, health and financial circumstances'.[462]

In Australia, in the absence of a general privacy protection, the right to privacy has limited protection in law. Australian law prohibits the misuse of 'personal information' about an identified individual, including sensitive information (such as a person's health information, their racial or ethnic orientation, sexual orientation or criminal record) and credit information.[463]

The Australian Privacy Principles, created under the *Privacy Act 1988* (Cth) and administered by the OAIC, guide organisations about how they should collect, store, manage and use personal information, including in the context of data and data analytics. Principle 10, for example, requires entities to 'ensure that the personal information that the entity collects is accurate, up-to-date and complete'.[464]

Australian privacy law protects information privacy only. It permits de-identified or anonymised personal information to be processed for the primary purpose for which it was collected, based on the consent of the individuals concerned.

AI challenges this model in a number of ways. For example, AI is increasingly capable of disaggregating a dataset made up of a conglomeration of anonymised personal information to reveal the personal information of specific, identifiable people.

The technology is developing quickly with ever-increasing new uses of personal information being developed, many of which could not have been envisaged, let alone specifically consented to, at the point of collection. Certain revealing information, including metadata, has been held by Australian courts not to fall within the parameters of the *Privacy Act 1988* (Cth).[465]

In 2014, the Australian Law Reform Commission (ALRC) examined how new and serious invasions of privacy have arisen in the digital era—without legal protection.[466] The ALRC recommended a statutory cause of action for serious invasion of privacy, especially given the increased 'ease and frequency' of invasions of personal privacy that may occur with new technologies.[467] This was supported by stakeholders to this Project.[468]

More recently, the Australian Competition and Consumer Commission supported the ALRC's recommendation in order to 'increase the accountability of businesses for their data practices and give consumers greater control over their personal information'.[469] Submissions to the current consultation also identified the importance of integrating privacy and data governance policy, and the need for specific privacy legislation to mirror Europe's GDPR.[470]

(b) Use of AI and automation in government decision making

Automated decision making by government is already expressly provided for under some Australian laws. Under s 495A of the *Migration Act 1958* (Cth) (the Migration Act), for example, the responsible Minister 'may arrange for the use, under the Minister's control, of computer programs' to make a decision, exercise a power or comply with any obligation.

The Minister will be taken to have made a decision, exercised a power or complied with an obligation where the relevant action 'was made, exercised, complied with, or done … by the operation of a computer program'. Similar provisions exist to 'use a computer program' to support government decision making in a range of areas, including social security, superannuation, health services and child support.[471]

Section 495A was inserted into the *Migration Act 1958* (Cth) in 2001.[472] The Explanatory Memorandum noted that complex decisions requiring discretionary elements, such as visa cancellations, would continue to be made by the Minister or a delegate, noting that these types of discretionary decisions 'do not lend themselves to automated assessment'.[473] Guidance issued by the Administrative Review Council in 2004 similarly warned that 'expert systems should not automate the exercise of discretion'.[474]

The progress in AI-informed decision making since the early 2000s could not have been contemplated by lawmakers at that time. The possibility of full automation of AI-informed decision making, for example, is now a realistic prospect. This means that older legislation dealing with this issue should be reviewed. Technological development necessitates a new approach to ensure AI-informed decision making is accountable.[475]

(c) The use of AI to make administrative decisions

Most administrative decisions will be subject to the same legal requirements whether or not they are made using AI. However, there may be some difficult, novel issues that arise for at least some forms of AI-informed decision making.

For example, the Australian Constitution entrenches a right to judicial review for any decision made by 'an officer of the Commonwealth'.[476]

However, if a decision is fully automated, and there is no involvement in the specific decision by a human who meets the description of an officer of the Commonwealth, a question arises whether the decision remains within the constitutional ambit. A court has yet to rule on this issue.

Outside of this constitutional context, Australian courts have begun to consider the scope of reviewable 'decisions' in the context of AI-informed decision making. For example, in the 2018 *Pintarich* case, the Full Court of the Federal Court of Australia considered a dispute over a tax debt. The taxpayer received a number of decisions and correspondence from the Australian Tax Office (ATO), including a letter that a human ATO decision maker used a computer program to generate. The letter included information that was at odds with some other ATO communications, regarding the taxpayer's tax debt. The Court decided, by majority, that the letter was not a reviewable 'decision' under the ADJR Act.[477]

The majority judges found that in order for there to be a decision 'there needs to be both a mental process of reaching a conclusion and an objective manifestation of that conclusion'.[478] In dissent, Kerr J found that applying this computation process still involved a reviewable decision:

> What was once inconceivable, that a complex decision might be made without any requirement of human mental processes is, for better or worse, rapidly becoming unexceptional. …The legal conception of what constitutes a decision cannot be static; it must comprehend that technology has altered how decisions are in fact made and that aspects of, or the entirety of, decision making, can occur independently of human mental input.[479]

The *Pintarich* decision, while significant, will not be the last word from Australian courts about the meaning of an administrative decision in the context of AI-informed decision making.

The legislation under which the ATO makes decisions, and the facts of the case, situate this decision in a very particular context. Caution should be exercised in extrapolating from the Full Federal Court's decision. Nevertheless, this case suggests that the use of AI in administrative decision making can affect its legal status, altering how, and even whether, certain government decision making can be reviewed.

(d) Harmful decision making

Some types of AI-informed decision making raise particular risks of harm. This is putting increasing pressure on legislators here and overseas to pass laws that address those specific types or areas of decision making.

Particular concern has centred on facial recognition technology, the human rights implications of which are explored in detail in Chapter 2. For example, Stadiums Queensland recently started using facial recognition technology. In response, the Queensland Privacy Commissioner pointed out that the risks must be considered when using this technology, particularly given the likelihood of any negative impact being disproportionately felt by ethnic minorities.[480]

Some have suggested a moratorium on the use of facial recognition technology to make at least certain types of decisions.

Case study:

A legal framework for facial recognition technology

Automated facial recognition technology, which relies on AI, is increasingly being deployed by governments and the private sector in Australia and overseas. Facial recognition can be a particularly intrusive form of digital surveillance. Depending on where and how it is used, facial recognition can engage numerous human rights, including the right to privacy, equality before the law, non-discrimination and, in more extreme cases, the right to be free from cruel, inhuman and degrading treatment.[481] At present, facial recognition technology can be prone to error, with those errors disproportionately affecting women and people of colour among others.[482]

To date, where government has used facial recognition technology, it generally has relied on existing legislation. However, as public concern is growing about the risks associated with this technology, some question whether there should be new, more specific legislation regulating government use of this technology.

If passed, the Identity-matching Services Bill 2019 (Cth) (the Bill) would establish a legal framework for the use of facial recognition technology, by Australian governments and others, in a range of policing and other contexts. The Australian Parliamentary Joint Committee on Intelligence and Security (PJCIS) reviewed this Bill, issuing a unanimous report in October 2019.

The PJCIS recommended the Bill be redrafted, to adopt an approach that is 'built around privacy, transparency and subject to robust safeguards'.[483] The PJCIS criticised the Bill's breadth and potential impact on most Australians, noting that it failed to adequately inform people about what it would authorise and what rights and responsibilities citizens would have under it.[484]

In the United Kingdom (UK), an independent report, supported by London's Metropolitan Police Service (MPS), concluded the MPS's trial of live facial recognition technology during police operations would be unlawful under UK human rights law. The report concluded:

> [T]he implicit legal authorisation claimed by the MPS for the use of live facial recognition–coupled with the absence of publicly available, clear, online guidance–is likely inadequate when compared with the 'in accordance with the law' requirement established under human rights law.[485]

This followed a call by the UK Parliament's Science and Technology Committee to pause any rollout of facial recognition technology beyond current pilots until concerns regarding bias and effectiveness 'have been fully resolved'. The decision about wider deployment, the Committee argued, is one for elected Ministers and Parliament, rather than individual police forces.[486]

The High Court of England and Wales recently considered a challenge to the use of facial recognition technology by a UK police force. In a decision that is subject to appeal, the Court accepted there was lawful authority for this use of automated facial recognition technology, but it noted the Government's

> pragmatic recognition that (a) steps could, and perhaps should, be taken further to codify the relevant legal standards; and (b) the future development of [automated facial recognition] technology is likely to require periodic re-evaluation of the sufficiency of the legal regime. We respectfully endorse both sentiments, in particular the latter.[487]

In his 2018 and 2019 reports, the UK Biometrics Commissioner and Forensic Science Regulator called for a moratorium on the use of facial recognition technology 'until a legislative framework has been introduced and guidance on trial protocols, and an oversight and evaluation system, has been established'.[488] A similar call for a moratorium has also emerged from a number of NGOs.[489]

In the United States, the San Francisco local government passed legislation banning the use of facial recognition software by the police and other government departments in May 2019.[490] Two other US local governments have imposed similar bans, while some private companies have refrained from using the technology in products marketed to police forces.[491]

6.4 Transparency

This Discussion Paper uses the term 'transparency' to refer to people being made aware about when AI is used in decision making.

In the context of government's use of AI-informed decision making, this form of transparency applies a principle of open government: that there should be 'publicity about the operation of the state'.[492]

Stakeholders emphasised the importance of transparency to protect human rights in the context of AI-informed decision making.[493] Some adopted a similar definition of transparency to the Commission's.

Others also saw transparency as including the idea that an individual affected by a decision has the right to request a meaningful explanation for an AI-informed decision. That related concept, known as 'explainability' in the context of AI, is dealt with later in this chapter.

It is not clear which government departments are using, or contemplating, decision-making systems that rely on AI.

Some submissions suggested law reform to require that individuals are informed if AI has been used in a decision that affects them.[494] Some stakeholders went further, suggesting there should be transparency regarding any data source that is used to train an AI-informed decision-making system.[495]

Transparency can also bring practical benefits. For example, the Commonwealth Ombudsman submitted that transparency is fundamental to the process 'of continuous improvement that is so important in digital transformation processes', including to ensure trust in digital services.[496]

The Australian Council of Learned Academies (ACOLA) report on AI also notes the importance of transparency:

> Transparency and explainability are important for establishing public trust in emerging technologies. To establish public confidence, it will be necessary to provide the public with an explanation and introduction to AI throughout the initial adoption stage.[497]

6.5 Explainability

Requiring a decision to be transparent and lawful would be hollow if a person affected by an AI-informed decision is unable to discern whether the relevant legal requirements were, in fact, followed. In practice, such an opaque decision would be unaccountable.

Many stakeholders observed that a person affected by an AI-informed decision should be able to understand the basis of that decision—that is, decisions should be explainable.[498]

This general principle is particularly important for government decision making. Individuals can request reasons for most types of administrative decision,[499] and in some circumstances the government must provide reasons when communicating the decision to an affected individual.[500] The Commonwealth Ombudsman can also recommend that a government agency give reason, or better reasons, in relation to an action the agency has taken.[501]

What is required by way of explanation for a government decision will vary, depending on the applicable law and the decision-making context. Generally, a statement of reasons in administrative law should contain the decision, findings on material facts, evidence or other material on which those findings are based, and the reasons for the decision.[502]

(a) Explaining AI-informed decision making

There are growing calls for AI to be 'explainable'. For example, the European Commission's 'Ethics Guidelines for Trustworthy AI' promote 'explainability', 'traceability' and communication to affected users.[503]

'Explainability' in the AI context has been defined in several ways. The European Commission's Independent High-Level Expert Group on AI, for example, defines explainability to include:

> the ability to explain both the technical processes of an AI system and the related human decisions... Technical explainability requires that the decisions made by an AI system can be understood and traced by human beings.[504]

Experts also distinguish between two *types* of explanations that may be relevant to AI-informed decision making, namely:

- *system functionality*, referring to the 'logic, significance, envisaged consequences, and general functionality of an automated decision-making system', and
- *specific decisions*, which refers to 'the rationale, reasons, and individual circumstances of a specified automatic decision, eg the weighting of features'.[505]

Dr Jake Goldenfein has listed factors that could be included in an 'explanation' for an AI-informed decision, including:

> disclosures about the specification and design of an algorithm; the system's explicit purpose; the features and weightings the system uses; the kind of outputs it generates and how they contribute to a decision; what level of human intervention remains or is possible; whether the system has been validated, certified or audited, and in what context; whether the system uses a fairness model and what type of model.[506]

Other discussions of explainable AI have a narrower focus on the elements needed for an AI-informed decision to be challenged or reviewed. The European Commission for the Efficiency of Justice, for example, has stated that where an algorithm is being used by a court in the decision-making process, both parties

> should have access to and be able to challenge the scientific validity of an algorithm, the weighting given to its various elements and any erroneous conclusions it comes to whenever a judge suggests that he/she might use it before making his/her decision.[507]

Professors Citron and Pasquale have focused on developing the concept of 'technological due process' to prevent arbitrary outcomes from predictive algorithms. The authors make a number of recommendations to support 'technological due process', such as requiring 'immutable audit trails' to inform affected individuals; opening up data sets and giving rights of appeal at each stage of data collection and analysis; and opening up black box algorithmic scoring systems to an affected individual or a 'neutral expert' representative in order to challenge 'arbitrariness by algorithm'.[508]

There are both commercial and technical obstacles to explaining AI-informed decisions.

First, revealing this information could also reveal commercially sensitive information. The owner of an AI-informed decision-making system, or a third-party developer, might object to revealing information about the system's operation (including any algorithms that the system uses), because it would reveal proprietary information.

Secondly, there might be a technical reason why the system's use of AI cannot be explained. In its recent horizon scanning report, for example, ACOLA noted that in the case of AI that engages in unsupervised learning, 'it is, in principle, impossible to assess outputs for accuracy or reliability'. ACOLA, accordingly, recommended a regulatory focus on trustworthy and transparent data, rather than focusing on an explanation for how a decision has been reached.[509]

Courts are starting to consider this problem. In the US, for example, teachers successfully challenged the use of an AI-informed decision-making system, purchased from a third party and used by the Houston Independent School District, to terminate public school teachers for ineffective performance.[510] The teachers relied on the US Constitution due process protections against substantively unfair or mistaken deprivations of life, liberty, or property.[511]

A number of technology companies have started to develop and market products that include an 'explainability' component.[512] The 2019 US Government research and development strategic plan includes funding to support the development of explainable AI, or 'XAI', which is trusted, acceptable, and 'guaranteed to act as the user intended'.[513] This builds on research programs already being funded by the US Government.[514]

(b) 'Algorithmic bias' and explainability

Where an AI-informed decision is explained, the decision can be scrutinised for error. In particular, it allows analysis of whether the decision is affected by 'algorithmic bias', a term that is widely used, encompassing statistical bias long familiar to computer scientists, data analysts and statisticians, as well as concepts of fairness, equality and discrimination.

Whether algorithmic bias is harmful, and potentially unlawful, depends on the context. Some experts have highlighted the human rights problems that arise where algorithmic bias leads to 'outcomes which are systemically less favourable to individuals within a particular group and where there is no relevant difference between groups that justifies such harm'.[515] In other words, human rights advocates are principally concerned about algorithmic bias that leads to discrimination which is unlawful under domestic or international law.

Sometimes algorithmic bias can be easy to identify, such as an algorithm designed to exclude a particular group or individual, or give material weight to a protected human attribute such as race, age or gender.

However, algorithmic bias also can be harder to detect or address. Much will depend on the data used to train an AI-informed decision-making system.[516] Some refer to data science's 'garbage in, garbage out' problem, where low quality or flawed data produces low quality results that may be unreliable or discriminatory. Algorithmic bias can arise where the designer of an AI system gives undue weight to a particular data set; or where the system relies on a data set that is incomplete, out of date or incorrect; or there might be selection bias, where the data set is not representative of a population so may ultimately favour one group over another.[517]

There is particular concern where an AI-informed decision-making system is 'trained' on historical data that is affected by prejudice or unlawful discrimination. A typical situation might be where an AI system is developed to make home loan decisions, but is trained on many years of human decisions that were prejudiced against female loan applicants.

In this situation, where training data contains a historical bias, the AI system can replicate or even reinforce this bias in the system's outputs. The historical bias can be hidden in the training data, but any unfair disadvantage experienced by a particular group can be transposed into the AI system. This problem still exists even if there is no longer any underlying prejudice or other improper motivation in the design of the new AI system.[518]

An oft-cited example is a recruitment tool that favoured male over female candidates. The algorithm was trained to identify patterns in the company's resumés over a 10-year period; as most of these were male applicants, the system 'learned' that male applicants were preferable, and made recommendations for the future workforce accordingly.[519]

Similarly, profiling individuals through data mining in order to draw inferences about their behaviour, carries risks of unfair and discriminatory treatment.[520] Another problem associated with machine learning is where a seemingly innocuous factor, such as the post code where an individual lives, becomes linked to a protected attribute, such as race, through unpredicted correlations.[521] This, too, might lead to unlawful discrimination (see, above).

(c) Content of an explanation

When an individual is entitled to an explanation for an administrative decision, the law sets out clear principles about what such an explanation should entail. In administrative law, this right to an explanation is most commonly referred to as 'a right to reasons'.

In some situations, legislation will set out in detail what should be set out in reasons for a particular administrative decision. However, if legislation simply provides a general right to reasons for an administrative decision, s 25D of the *Acts Interpretation Act 1901* (Cth) states that those reasons should include 'findings on material questions of fact and refer to the evidence or other material on which those findings were based'.

At least for government's use of AI-informed decision making, these general legal principles should continue to apply. Such an explanation can enable a meaningful review of the decision, including to determine whether it is lawful.

While an explanation or reasons should generally be expressed in language that an ordinary lay person would understand, there are some situations where a more technical explanation may be necessary. Where AI is used to make a decision, it sometimes may be necessary for a detailed technical explanation about how the AI informed the decision-making process. This could be interpreted by a technical expert.

This suggests also that technical expertise to analyse AI-informed decisions, or look into an opaque AI-informed decision-making process, may be required within regulatory and review bodies. The use of technical expertise is common in judicial and regulatory settings, such as the use of expert witnesses in criminal and civil trials and expert 'assessors' in the Land and Environment Court.[522] Technical experts are also relied on in the medical and health context. The Therapeutic Goods Administration (TGA) in Australia, for example, assesses the risks and benefits of medicines prior to registration, including by seeking advice from eminent experts in the relevant field.[523]

(d) Explainability and the GDPR

No state has yet legislated to create an express right to an explanation of an AI-informed decision. As stakeholders noted, the GDPR has perhaps come closest to recognising such a right, although its precise requirements and whether they amount to a right to an explanation, are disputed.[524]

The GDPR establishes several rights of the data subject (individual) that relate to information about the processing of personal data. These include:

- the right to be informed about certain matters when personal data is collected, such as the period the personal data will be stored
- the right to be informed when personal data that has not been obtained directly from the data subject will be processed
- rights of access, including the right to ask a data controller whether or not their personal data is being processed, and about certain aspects of that processing such as its purpose.[525]

Article 22 of the GDPR provides specific rights for a person who is subjected to automated individual decision-making, including profiling. These rights include:

> the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.[526]

Where an individual has granted explicit consent to the use of their personal data in an automated process, the individual retains 'at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision'.[527] This right has been implemented in some EU domestic laws.[528]

Some experts argue that the GDPR does not support a meaningful right to explanation. Rather, its right of access 'provides a right to explanation of system functionality' or a 'right to be informed', which is restricted by the interests of the data controller.[529]

Others argue that the GDPR supports a right of access to a meaningful explanation of a decision involving an individual's personal data. Andrew Selbst and Julia Powles, for example, argue that, read together, Articles 13—15 of the GDPR provide a right to 'meaningful information about the logic involved' in an automated decision. While the authors do not conclude precisely what that explanation should look like, they state:

> We believe that the right to explanation should be interpreted functionally, flexibly, and should, at a minimum, enable a data subject to exercise his or her rights under the GDPR and human rights law.[530]

Other scholars have focused on a broader notion of transparency, rather than explainability. Processor Margot Kaminski, for example, concludes that the GDPR supports transparency by giving regulators significant information-forcing capabilities to access information about the algorithm; requiring companies to set up internal accountability and disclosure regimes, including by performing data protection impact assessments; and by recommending the use of third-party auditors, given access to all necessary information regarding the internal workings of the machine learning system or algorithm.[531]

6.6 Responsibility and liability

Any system of accountability ascribes responsibility for the consequences of erroneous decision making. While responsibility can be moral or legal, liability is an exclusively legal concept. A person may be morally blameless but nevertheless be legally liable for remedying the consequences of a wrongful decision that causes harm.

This section of the chapter considers questions that arise regarding responsibility, and especially legal liability, for AI-informed decision making.

(a) Liability for AI-informed decision making

Some stakeholders suggested that decision-making systems that use AI should be designed in ways that make clear who is liable in the event of harm.[532]

There is a considerable body of law on determining legal responsibility, or liability, for decisions that affect other people. These existing legal principles are likely to resolve many liability questions in the context of AI-informed decision making.

As a general principle, laws covering areas such as product liability and consumer safety, discrimination, and competition, are technology neutral. If an individual applies for a bank loan, it would be unlawful to discriminate against them based on their race. It should make no difference whether that discrimination was caused by a human bank manager being racially prejudiced, or by the manager applying a company-wide policy that was racially prejudiced, or by the bank making its decision in reliance on an algorithm that produced racially discriminatory results.

In the last of the three scenarios, the bank might have primary liability for the resultant discrimination, but the bank might be able to identify others that share some portion of its liability (such as a third party company it contracted with to develop the problematic algorithm). In any event, as the technology develops, traditional concepts of liability could be increasingly challenged.

Stakeholders highlighted this as an issue that needs resolution.[533] Challenges to identifying liability include the potential removal of humans in fully automated decision-making processes; the opaqueness of AI decision-making systems; and the importation of AI-powered software from overseas.

In its current, most common forms, AI tends to be used to identify correlation, rather than causation. Correlation can be a powerful indicator of causation, but not always. If a person uses AI to identify a correlation, then engages in a rigorous process to test whether this correlation is indeed indicative of causation, logically this must strengthen the reliability of their process of reasoning. It would make it less likely that the person will make an error, and it could reduce their legal liability if an error does occur.

This scenario is relatively common in medical research. For instance, if a medical researcher notices a correlation between eating a type of berry and a lower risk of cancer, the researcher might hypothesise that eating the berry reduces a person's cancer risk. However, the researcher would know that the correlation could indicate a causative relationship, or it could be random chance and there is in fact no cancer prevention benefit from eating the berry. Typically, the researcher would test the berry-cancer hypothesis by looking for a causal link, as well as other independent, points of correlation that prove, or disprove, the hypothesis.

In practice, this might involve conducting controlled trials to exclude other possible explanations for the correlation; using statistical tools to assess whether the observed correlation might be a matter of chance; and attempting to understand the pathway by which the berry might have the hypothesised effect. In other words, the original correlation that the researcher observed is useful provided it is treated as a correlation. A problem would arise if the researcher failed to go further in testing this correlation.

This logic should apply also to the use of AI. Where an AI system draws a correlation, this might need to be carefully tested before it is relied on—especially in circumstances where it is being used to make decisions that significantly affect people's rights and interests.

Applying this logic to a different scenario, imagine a bank uses an AI system to trawl through its historical data set of loan decisions, and this suggests that electricians repay their loans at above-average rates. Without more, this might found a *hypothesis* that electricians are particularly reliable at repaying their loans, but more research and analysis probably will be needed before this hypothesis could be treated as proven.

If the bank chooses to rely on the original correlation, without further testing, it could cause the bank to treat some customers unfairly. In any event, if the bank diligently tested the correlation-based hypothesis before relying on it, it would reduce its risk of error and its legal liability.

(b) Responsible use of AI in decision-making

Responsibility extends beyond legal liability, to include also ethics, fairness and good governance. Ensuring AI-informed decisions are made responsibly underpins social trust, and the concept of responsibility has featured in industry and government frameworks and guidance. The Canadian Government's Directive on Automated Decision-Making, for example, guides decisions made by federal government departments to be 'data-driven' and 'responsible'.[534]

Where AI-informed decision making is lawful, but generates harm, its use may nevertheless not be responsible. Some commentators have suggested that where an AI-informed decision-making process might cause individual harm, such as using health data to justify increases in health insurance premiums for certain categories of people, it should not be used even if it is lawful under federal discrimination law.[535] Another example is the impact of AI-informed decision making on social equality, which may raise questions about the responsibility of using AI in certain contexts.

The Supreme Court of Canada, for example, concluded in 2018 that the Correctional Service of Canada was using a risk assessment tool that disadvantaged Aboriginal Canadians.[536] Finding that there was known cultural bias in the tools being used to assess risk within the Canadian prison system, the Court concluded:

> [T]he clear danger posed by the CSC's continued use of assessment tools that may overestimate the risk posed by Indigenous inmates is that it could unjustifiably contribute to disparities in correctional outcomes in areas in which Indigenous offenders are already disadvantaged.[537]

Box 3:

Respecting, promoting and fulfilling the human right to equality

The use of AI can make it harder to promote equality. Machine learning, for example, can be used to create profiles of people in order to predict behaviour, and make decisions accordingly. Such profiles may result in unfairness that is not prohibited by current anti-discrimination law.

Australian law does not, for example, prohibit discrimination on the basis of socio-economic status, and yet it is becoming clear that people of a lower socio-economic status often suffer disproportionately negative effects from AI-informed decision making, particularly when used by governments.[538]

Professor Virginia Eubanks calls this 'a kind of collective red-flagging, a feedback loop of injustice', where marginalised groups face high levels of data collection through accessing social services, living in highly policed neighbourhoods and accessing public health systems. This leads to higher scrutiny through more intense surveillance and even more data collection.[539]

Different social groups tend to be differently affected by AI-informed decision making. Stakeholders noted the importance of acknowledging the 'digital divide', where those without basic connectivity risk being excluded from accessing government services.[540] Some groups may be disadvantaged by the technology itself. Aboriginal and Torres Strait Islander peoples, for example, may be less likely to feel comfortable answering direct questions often used in algorithmic decision-making tools.[541]

Whether governments' use of AI entrenches or reduces social inequality will depend on how AI-informed decision-making systems are designed and deployed. This includes considering suitability of using AI for particular populations, determinations or service provision. One reform option is the introduction of a positive duty to promote equality, whereby government must take measures to reduce discrimination in society and to contribute to the greater realisation of equality on a continual basis.[542]

6.7 Human intervention in AI-informed decision making

Stakeholders emphasised the importance of the role played by humans in overseeing, monitoring and intervening in AI-informed decision making.[543]

The European Commission's 'Ethics Guidelines for Trustworthy AI' refer to various levels of human involvement in AI-informed decision making, including:

- 'human-in-the-loop', referring to the 'capability for human intervention in every decision cycle of the system (although the Guidelines also acknowledge that 'in many cases [this] is neither possible nor desirable')
- 'human-on-the-loop', referring to human intervention during the design phase and monitoring of the system in operation
- 'human-in-command', referring to the 'capability to oversee the overall activity of the AI system (including its broader economic, societal, legal and ethical impact) and the ability to decide when and how to use the system in any particular situation' including deciding not to use AI, establishing levels of human discretion or giving human decision makers the ability to override a decision.[544]

Some commentators have suggested human oversight alone will be insufficient to minimise the adverse impact of AI, particularly if the particular human, who is 'in the loop', has insufficient technical knowledge to understand, or communicate, the explanation.[545] Other experts have suggested that neutral experts could be made available to affected individuals to overcome the challenge posed by lack of technical knowledge.[546]

6.8 The Commission's preliminary view

The Commission observes that the use of AI in decision making can increase efficiency, enable more data-driven decisions and minimise some types of human bias, but it can also lead to opaque decisions, introduce new forms of bias (or replicate old ones), and undermine human rights.

The Commission considers that *accountability* is central to harnessing these benefits and addressing these risks. In the remainder of this chapter, the Commission proposes a number of ways of ensuring that AI-informed decision making is accountable.

(a) Accountable AI-informed decision making

AI-informed decision making must be accountable. The need for accountability derives from international and domestic human rights law, as well as principles such as the principle of legality and the rule of law, which are fundamental to Australia's democracy. Leading technology and law scholars have identified the problem we face:

> As technology develops, and machine learning becomes more sophisticated, forms of automation used by government may increasingly become intelligible only to those with the highest level of technical expertise. The result may be government decision-making operating according to systems that are so complex that they are beyond the understanding of those affected by the decisions. This raises further questions about the capacity of voters in democratic systems to evaluate and so hold to account their governments, including in respect of compliance with rule of law values.[547]

If Australia is to benefit from AI, while safeguarding human rights, our law must promote accountable AI-informed decision making. This is required for Australia to comply with its obligations to protect, fulfil and respect human rights under international human rights law.

Those people who are negatively affected by an AI-informed decision may also face further adverse consequences if there is a failure of accountability. To date, the known harms associated with AI-informed decision making are being disproportionately experienced by those who are already vulnerable and marginalised.

The Commission's preliminary view is that accountability for AI-informed decision making will be improved by more rigorously applying existing laws, as well as some targeted reform. The Discussion Paper proposes five critical areas of focus:

- ensuring that AI-informed decision making complies with all relevant legal requirements, and is expressly authorised by law when carried out by government
- promoting transparency, so that individuals are informed where AI has been a material factor in a decision that affects them
- ensuring that AI-informed decisions are explainable, in the sense that a reasonable and meaningful explanation can be generated, that is communicated to any affected individual on demand
- clarifying who is legally liable for using an AI-informed decision-making system, with strong incentives for acting responsibly in developing and using such systems
- identifying appropriate mechanisms for human oversight and intervention.

Each of these areas of focus is considered in detail below.

(b) Lawfulness

It is vitally important that AI-informed decision making complies with the law, and especially laws designed to protect and promote human rights.

(i) Upholding the rule of law in government decision making

The former Chief Justice of Australia, the Hon Robert French AC QC, has observed that there are 'four basic requirements for just decision making in a society governed by the rule of law—lawfulness, fairness, rationality and intelligibility'.[548] These requirements apply to all government decision making, including where AI is used.

The Commission's research and consultation have highlighted challenges in ensuring that government's use of AI-informed decision-making complies with existing laws, and fundamental democratic principles such as the rule of law.

There are certainly limits to when and how AI can be safely used in this context. The Commission considers there is a need to assess more deeply whether, and how, AI can and should be used to make complex evaluative judgments that are critical to many forms of government decision making.

Take, for instance, the obligation to act *fairly*. There are multiple factors that must be weighed in determining whether something is fair. In 2018, Professor Arvind Nayanan analysed 21 distinct definitions of 'fairness', concluding that they are not only mathematical questions, but shaped by values, philosophies and politics.

To illustrate this point, in 2018, Professor Arvind Nayanan analysed 21 distinct definitions of 'fairness', concluding that they are not only mathematical questions, but shaped by values, philosophies and politics.[549]

These issues merit close attention in two ways.

First, there is a need to monitor, on an ongoing basis, the development and use of AI-informed decision making, particularly when there is a risk that existing laws might be breached or circumvented. This could be undertaken by a new or existing body, as discussed in Part C.

Secondly, challenging questions of law have already emerged that should be considered by an appropriate law reform body in a dedicated inquiry. In particular, questions have arisen in the context of AI-informed decision making as they relate to the principle of legality and the rule of law, as well as the application of anti-discrimination legislation. Considering these issues in a dedicated inquiry is particularly important given the absence of a federal charter of human rights, which would provide more comprehensive legal protection for human rights.[550]

‘Review how the government uses AI to make decisions.’

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.**

(ii) Legislation to regulate government use of AI in decision making

The Commission proposes that where the government deploys an AI-informed decision-making system, this should be expressly provided for in law. In other words, adopting the definition in Chapter 6, where AI materially assists the government’s decision-making process, this should be regulated by express legislation. This is particularly important where decision making has a significant effect on individuals’ human rights.

This approach would ensure there is legislative oversight over the cost-benefit analysis for each area of government decision making in which AI is proposed to be used. This would assist in bringing to the surface the potential for harm, and especially encroachment on human rights, and create an impetus for addressing such risks. It would also enable a careful consideration of how accountability will be ensured in each specific context in which AI is proposed to be used to assist or facilitate government decision making.

Where Parliament decides to permit this use of AI, it would then set legal rules regarding how AI-informed decision making is deployed in each specific context. The Commission's proposal to this effect is set out later in this chapter at Proposal 6 below.

(iii) Protection of privacy and equality

AI-informed decision making clearly poses particular risks to the right to privacy. This is partly reflected in growing community concern about privacy. In the OAIC's 2017 community attitudes to privacy survey, 69% of Australians reported being more concerned about online privacy than they were five years earlier, and 58% decided not to deal with an organisation because of privacy concerns.[551]

Three Australian law reform bodies—the Australian Law Reform Commission (ALRC) and its counterpart bodies in Victoria and New South Wales—have now recommended enacting a statutory cause of action for serious invasion of privacy.[552] In 2014, the ALRC recommended such a law to apply in two contexts: an intrusion upon seclusion, such as the physical intrusion into a person's private space; and misuse of private information, 'such as by collecting or disclosing private information about the plaintiff'.[553]

By extending the protection of Australian law beyond 'information privacy', such reform could address some (though not all) of the concerns about how personal information can be misused in the context of AI-informed decision making. As such, the Commission urges that this ALRC recommendation, which has received support from the ACCC and others, be implemented.

PROPOSAL 4:

The Australian Government should introduce a statutory cause of action for serious invasion of privacy.

As set out in detail above, there are also important questions about the effectiveness of existing anti-discrimination law in protecting the right to equality and non-discrimination.

'Modernise Australia's privacy and human rights laws.'

More research is needed about how AI-informed decision making can lead to unlawful discrimination. In the next phase of this project, the Commission will be working with a number of partner organisations to understand this problem better, and to identify solutions.

In addition, as part of its *National conversation on human rights*, the Commission is investigating changes that may be needed to current Australian anti-discrimination law, including to respond to the rise of new technologies like AI.[554]

(c) Transparency

(i) Notification of the use of AI in decision making

The Commission considers that an individual should be made aware when they are the subject of an AI-informed decision-making process—whether that decision was made by government or a non-government entity.

The Commission notes that there was significant demand among stakeholders for reform to achieve this goal. Knowing that the use of AI is material in a decision gives an affected person important information about how that decision was made. As AI can be more reliable at some tasks than others, this knowledge also will be useful in assessing the reliability of the decision in question.

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.

(ii) Transparency by government

The Commission considers there should be better consultation before AI-informed decision-making systems are used by government. The Commission proposes that, where a Government agency intends to deploy an AI-informed decision-making system, this should be subject to public consultation, especially with affected stakeholders.

Informing the Australian public when the government proposes to use AI-informed decision making would assist in assessing its suitability. Community consultation should take place prior to the implementation of the AI-informed decision-making system. The US-based AI Now Institute sees such consultation as important in resolving fundamental questions regarding the use of AI:

> [R]obust and meaningful community engagement is essential before a system is put in place and should be included in the process of establishing a system's goals and purpose.[555]

This would promote public trust in government use of AI, the importance of which was widely recognised in independent reviews of the implementation of Centrelink's debt compliance program (often referred to as 'Robodebt'; see above at 5.4).[556]

'People should be informed when the government uses AI in decisions that affect their human rights.'

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.**

(iii) Review of current government use of AI

As outlined above, several federal laws already provide for the use of a 'computer program' in some types of decision making. It is unclear the extent to which these provisions influence government agencies in considering the potential human rights impact when those agencies are contemplating or deploying AI-informed decision-making systems. It is also unclear whether these provisions provide an adequate regulatory framework for many fully-automated decisions.

The Commission considers it would be beneficial to review the scope of these current laws against the backdrop of current technological developments.

That proposed review should form part of a comprehensive review of government use of AI, outlined in Proposal 17 in chapter 7 below. This comprehensive review should also include the criteria identified in Proposal 6 above.

'Government should consult the community before using AI to make decisions'

(d) Explainability

(i) The right to an explanation

The Commission is concerned that AI-informed decision making can be more opaque than conventional decisions. This limits human rights and fundamental democratic principles such as the rule of law.

This is especially important in the context of AI-informed decision making by government. The Hon Robert French AC QC has expressed concern that where government generates automated decision, this can pose problems in complying with the legal requirement of rationality, wherein a decision 'must be supported by reasoning which complies with the logic of the statute'.[557]

The Commission's preliminary view is that, where there is a legal requirement for a decision to be explainable, there is no justification for discarding or reducing this requirement simply because AI has been used in the decision-making process.

It should be noted also that the Australian public is increasingly demanding an explanation for decisions that use AI.

For example, in 2019 polling undertaken by Essential Media for the Commission, 71% of individuals surveyed in Australia thought it was very important that a government agency provides an explanation when automated decisions are reached. Those who said they were uncomfortable with a government's use of AI were more likely to say that it is very important that government is able to provide an explanation of how that decision was made.[558]

The precise nature of any required explanation will vary somewhat by reference to the decision-making context. Generally, however, an individual affected by an AI-informed decision should have the right to an explanation for the AI-informed decision, which is accurate and sufficient to enable the individual to understand and, if necessary, challenge the decision.

In addition to a reasonable explanation that a layperson can understand, AI-informed decision-making systems should also be able to produce, on demand, a more technical explanation of their operation.

This would enable an expert to interrogate the simplified explanation to determine whether it is an accurate description of how the decision was made.

A technical explanation might include some material that is specific to AI-informed decision making, such as: the data set used to train the AI; any limitations of that data set; any profiles created from data mining used in the AI decision-making process; any risk factors and mitigating action taken; any impact assessments, monitoring or evaluation conducted by the decision maker; and key data points taken into account in the decision-making process and the weight attributed to those data points.

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.

(ii) The capacity to generate an explanation

There is considerable debate regarding the extent to which some forms of AI-informed decision making are capable of explanation.

The better view, among experts involved in this area appears to be that it is almost always possible to design an AI-informed decision-making system so that it provides a reasonable (albeit not perfect) explanation of the basis for the decisions or recommendations it generates.

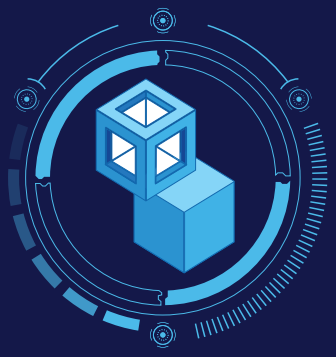
The Commission also notes a trend towards building an explanation function into AI-informed decision-making systems, including on the part of leading software companies.

Nevertheless, building an explanation into an AI-informed decision-making process can be technically difficult and expensive. Overseas jurisdictions, including the United States, have prioritised research and funding grants with the objective of ensuring AI-informed decision making can be explained.

With the recent establishment of the Australian Research Council Centre of Excellence for Automated Decision-Making and Society in October 2019, Australia has a similar opportunity. This Centre aims to formulate world-leading policy and practice and inform public debate, and to create the knowledge and strategies necessary for responsible and ethical automated decision-making.[559]

The Commission considers that where an AI-informed decision-making system has not been designed to enable an affected person to challenge a decision, it should not be used.

The Commission invites feedback on whether this principle should be adopted and, if so, how it should be reflected in Australian law. For example, in the context of AI-informed decision making, a new evidential rule could impose a rebuttable presumption that an AI-informed decision is not lawful, where a meaningful explanation cannot be provided to an individual affected by the AI-informed decision.



‘Government should only use explainable AI to make decisions.’

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.

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?

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.

(e) Responsibility and liability

(i) Apportioning liability

Assigning legal liability is central to ensuring affected individuals are able to access a remedy if they have been wronged.

The Commission considers that many questions regarding liability for errors or other legal problems arising from AI-informed decision making are likely to be resolved simply by applying existing technology-neutral legislation that governs liability in a range of contexts, from contractual disputes to product liability.

However, the Commission acknowledges the persuasive arguments of expert and other stakeholders that assigning liability for an AI-informed decision can present some unique difficulties.

The Commission's preliminary conclusion is that legal liability for any harm that may arise from reliance on an AI-informed decision should be apportioned primarily to the organisation that is responsible for the AI-informed decision. There will be situations where this is inappropriate, so this should be no more than a general rule, or rebuttable presumption, which could be displaced if there are strong legal reasons for doing so.

Legislation that makes this clear, along with guidance about how to apply this legal rule in a range of practical scenarios, may assist in resolving many of the difficulties regarding liability in this context.

Clarifying some of the ambiguity regarding liability in the context of AI-informed decision making could also give impetus to government agencies, companies and others to take a cautious approach when considering and deploying AI-informed decision-making systems. This could increase the effectiveness of some of the tools outlined in chapter 7, including human rights compliant procurement policies and human rights impact assessments, as this could reduce the legal risk associated with such systems causing human rights and other problems.

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 liable for the use of the system.

One problem, which is sometimes said to make accountability for AI-informed decision making more difficult, is that AI can rest on information that is commercially sensitive. It is often said that algorithms generated by companies, and even some generated by or for government, cannot be revealed for reasons connected to intellectual property law.

There are examples, from litigation overseas, where a company has developed an AI-informed decision-making system and then claimed a proprietary interest in the algorithm underpinning that system. The company has then relied on this proprietary interest to refuse to reveal to a claimant, or a court, how the decision was reached.[560]

If such arguments are successful, this can prevent an individual accessing a remedy for harm suffered, as well as obstruct the court's assessment of the lawfulness of the output of an AI-informed decision-making system.

'Whoever deploys an AI-informed decision-making system should be liable for its use.'

While that eventuality would be concerning, Australian courts already have powers to hear matters involving commercially sensitive evidence. Therefore, there is nothing that should preclude a court from receiving and assessing this sensitive evidence (such as an algorithm), with safeguards that prevent its broad publication.

Nevertheless, the Commission invites input regarding whether current Australian law is adequate in enabling courts to assess technical information associated with an AI-informed decision-making system, such as an algorithm, especially in a situation where it is claimed this information cannot be revealed for reasons of commercial sensitivity or confidentiality.

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?

(f) Human intervention

Designing an AI-informed decision-making system such that a human decision maker can intervene, taking control of the decision-making process, in certain situations, has been suggested as a way of combating some risks associated with AI-informed decision making.

Most commonly, it is suggested that a human decision maker might intervene in this way if the system generates a suggested decision that is likely to have a negative effect on an individual (a 'negative decision'). In this scenario, the human decision maker either would consider the relevant evidence to form their own view on the matter before the decision is finalised, or this suggested decision would simply go into a pool of cases to be resolved using a conventional form of decision making that does not materially rely on AI.

The role of human intervention in automated decision making is dealt with in the GDPR. Under Article 22(3), an individual affected by an automated decision, including profiling, has 'the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision'.

Where a significant decision has been made based solely on automated processing, UK legislation now provides that an affected individual can request the data controller to reconsider the decision that has been made, or take a new decision that is not based solely on automated processing.[561]

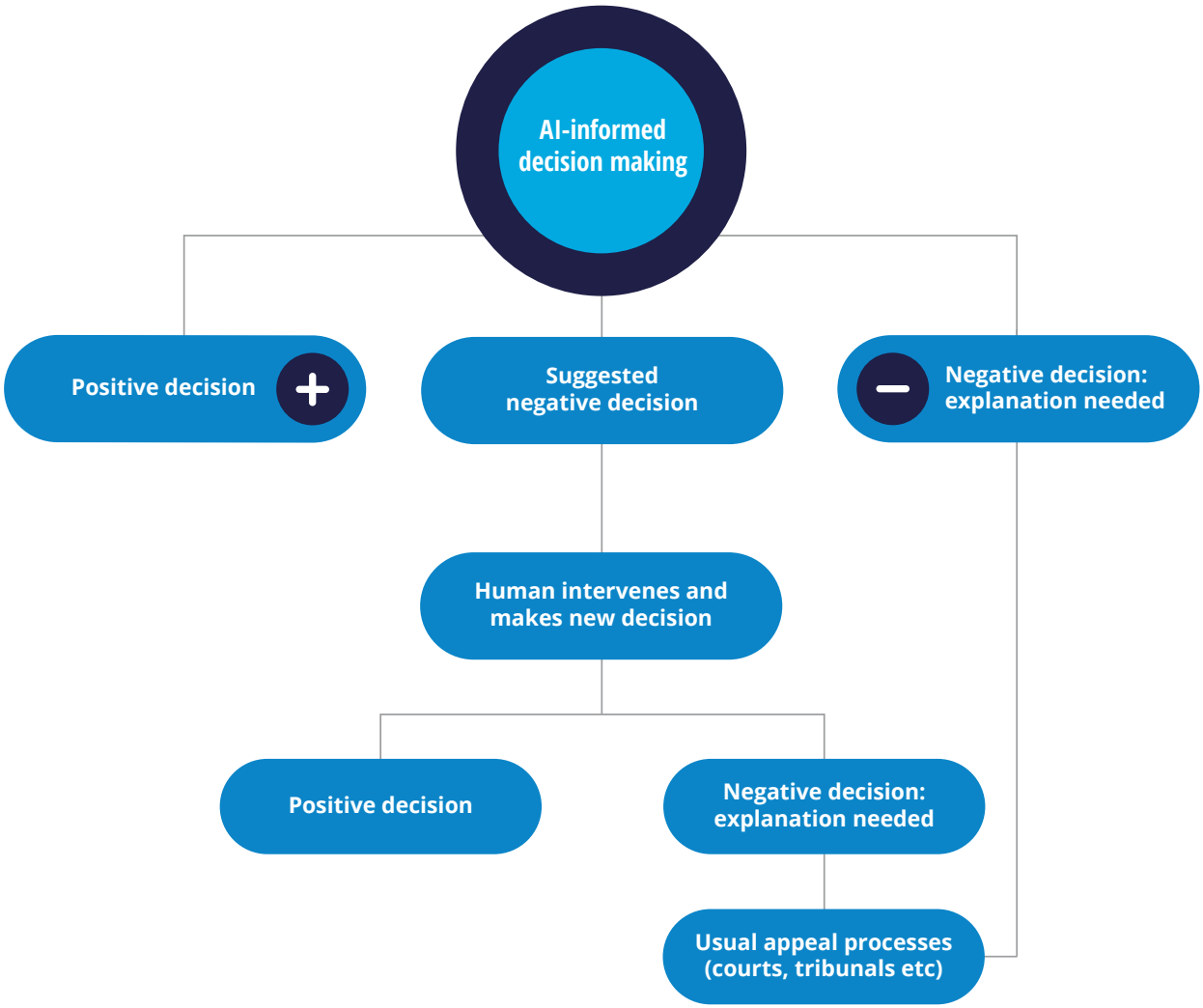
Increasingly, therefore, there is a trend towards permitting or mandating a human to intervene where a person who is the subject of a decision might otherwise receive an automated or AI-informed decision that is a negative decision.

The diagram below summarises a common way in which an AI-informed decision-making system can be designed to include, or exclude, the possibility of a human decision maker intervening in the event that the AI-informed process generates or suggests a negative decision.

The Commission sees value in designing systems that include appropriate 'failsafe' protections against unfair or unreasonable AI-informed decisions, and that make best use of the respective skills of human decision makers and AI in the decision-making process.

However, there is also a high degree of complexity in this area. Simply inserting a human decision maker into this process will not guarantee better decisions.

There is an emerging body of research suggesting that human decision makers can have extreme reactions when presented with datapoints and suggestions generated using AI, and that those reactions can have little to do with a rational understanding of the objective reliability of the technology in achieving its task.[562]



Some of this research points to humans being unreasonably averse or hostile to such AI-powered inferences,[563] and other research suggests unreasonable deference to such material.[564] One example is the 'biometric mirror' project, led by Dr Niels Wouters, which demonstrates how even unreliable AI-powered tools can be seen as useful and trustworthy.[565]

Taking account of such issues, the Commission is interested in stakeholder views about how best to design AI-informed decision-making systems, so that they most effectively make use of human decision makers.

QUESTION D:

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

(g) Ensuring strong safeguards for human rights prior to deployment

(i) Addressing risk of harm

The Commission urges the Australian Government to adopt a precautionary approach to the use of AI-informed decision making, where the identified risk of harm is particularly great.

By way of analogy, the European Court of Human Rights considered the right to privacy in the context of DNA collected from people suspected, but later acquitted, of committing a crime.

The Court cautioned against a rushed, unbalanced approach to deploying new technology without due consideration of the human rights impact. The Court, considering privacy rights in particular, stated that human rights would be

> unacceptably weakened if the use of modern scientific techniques in the criminal justice system were allowed at any cost and without carefully balancing the potential benefits of the extensive use of such techniques against important private-life interests … any state claiming a pioneer role in the development of new technologies bears special responsibility for striking the right balance in this regard.[566]

A balance must be struck in Australia taking up the opportunities offered by AI, but avoiding social and individual harm. Such harm is problematic on its own terms, but it can also reduce community trust in AI more broadly.

As explored in greater detail in Chapter 8, ongoing independent oversight is needed to ensure human rights are protected and promoted in the AI context.

(ii) Addressing known risk

There are also areas of AI-informed decision making where the harm, or likelihood of harm, is already well known. The use of facial recognition technology, the making of government decisions that have a high impact on people's human rights (such as refugee status determinations), and the deployment of autonomous weapons are all areas where there is a compelling case for close government attention, with a view to dedicated regulation.

Increasingly, governments are starting to turn their attention to areas where AI-informed decision making raises serious concern. For example, in November 2019, the Australian Government Department of Human Services made significant changes to its automated debt recovery system (known sometimes as 'Robodebt'), outlined in Chapter 5 above.[567]

As outlined above, the use of facial recognition technology in some areas of AI-informed decision making, is another prime example of where compelling community concerns should be addressed prior to widespread rollout.

These concerns include the privacy impact of intrusive surveillance using facial recognition technology. Implementing the ALRC's 2014 recommendations regarding reform to surveillance laws would go some way to addressing this problem,[568] but more is likely to be needed to protect human rights in this area of activity.

As noted above, there are growing calls for a full or partial moratorium on the use of facial recognition technology in contexts where the risk of harm is particularly acute, and some overseas jurisdictions have done just this.

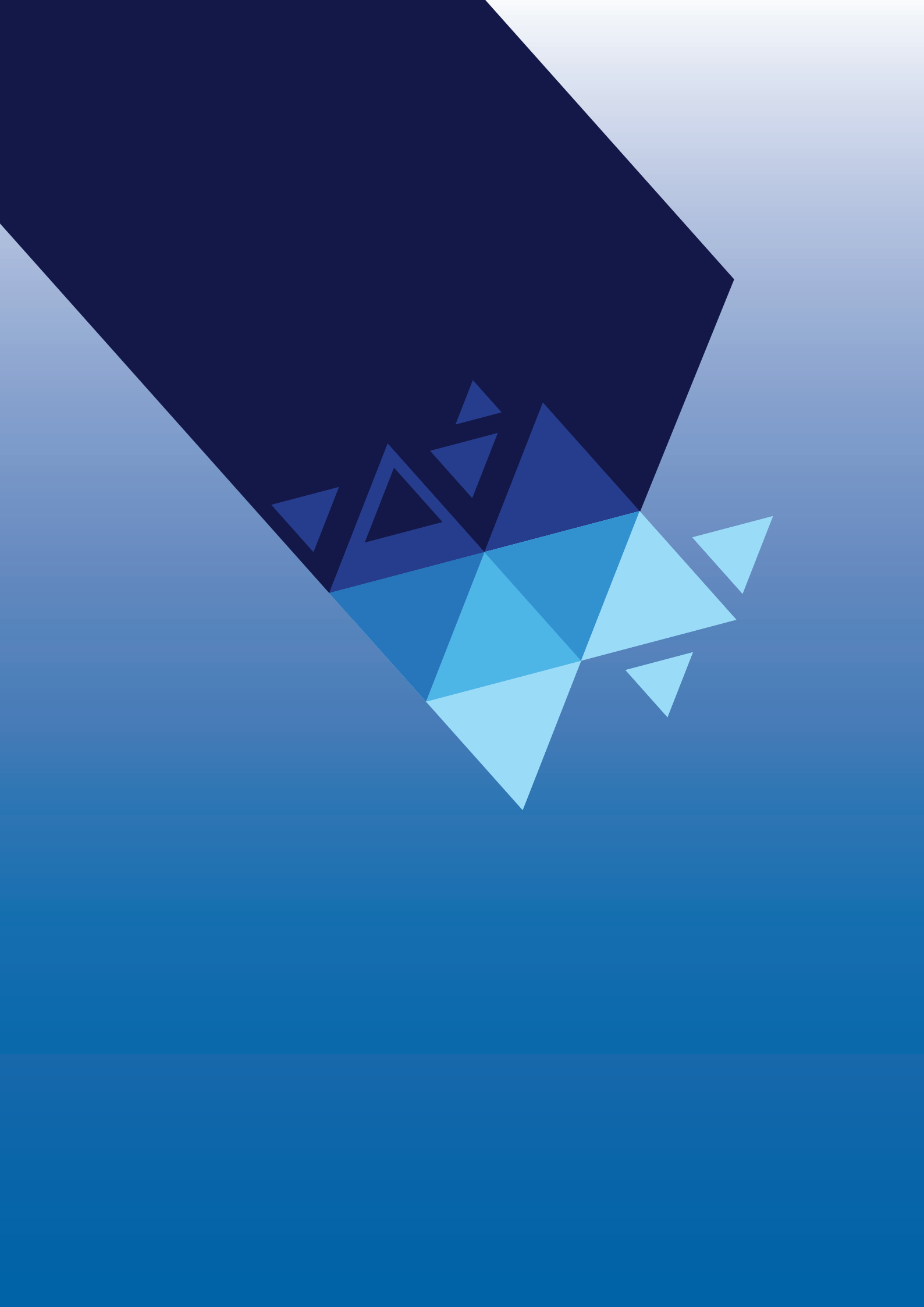
The Commission is concerned about the risks of using facial recognition technology in decision making that has a legal, or similarly significant, effect for individuals. In addition to its privacy impact, the current underlying technology is prone to serious error, and this error is often more likely for people of colour, women and people with disability, among others. The Commission's preliminary view is that legislation is needed for the use of facial recognition technology, and that this should include robust safeguards for human rights. As outlined in Chapter 2, draft legislation has, to date, failed to establish a comprehensive and rigorous framework governing the use of facial recognition technology, with insufficient safeguards for human rights protection.

The Commission proposes a moratorium on certain uses of facial recognition technology, until an appropriate legal framework that protects human rights has been established. This framework should be developed in consultation with experts, public authorities, civil society, and, importantly, the broader public who will be the subjects of this type of surveillance.

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.

'There should be a moratorium on government use of facial recognition technology, until there are robust legal protections for human rights.'



CHAPTER 7

Co- and self-regulatory measures for AI-informed decision making

7.1 Introduction

Acknowledging that legislation is not the only form of regulation, this chapter considers other regulatory measures that can operate in a complementary way to guide AI-informed decision making.

The UN High Commissioner for Human Rights has advocated a 'smart mix' of regulatory measures to protect human rights in the context of new and emerging technologies such as AI.[569]

In this chapter, the Commission makes a number of proposals that would protect human rights at various stages of the AI life cycle. These include proposals that relate to:

- design strategies, standards and impact assessments, which are directed towards the design and testing of AI systems before they are used in the real world
- regulatory sandboxes, which allow for new products or services to be tested in live market conditions but with reduced regulatory or licensing requirements, some exemption from legal liability, and with access to expert advice and feedback
- education and training on human rights compliant technology for a range of groups, including professionals who design, develop and use AI in decisions that significantly affect human rights, as well as the public and particularly affected groups.

This chapter also includes proposals that aim to enhance accountability for government use of AI-informed decision making, including to ensure transparency, such as through better auditing processes; ongoing human oversight and evaluation; and ensuring government procurement rules and practices, as they relate to AI-informed decision making, have stronger safeguards for human rights.

The related question whether an existing or new Australian body should take a leadership role on AI policy is discussed in Part C.

(a) A multi-faceted regulatory approach

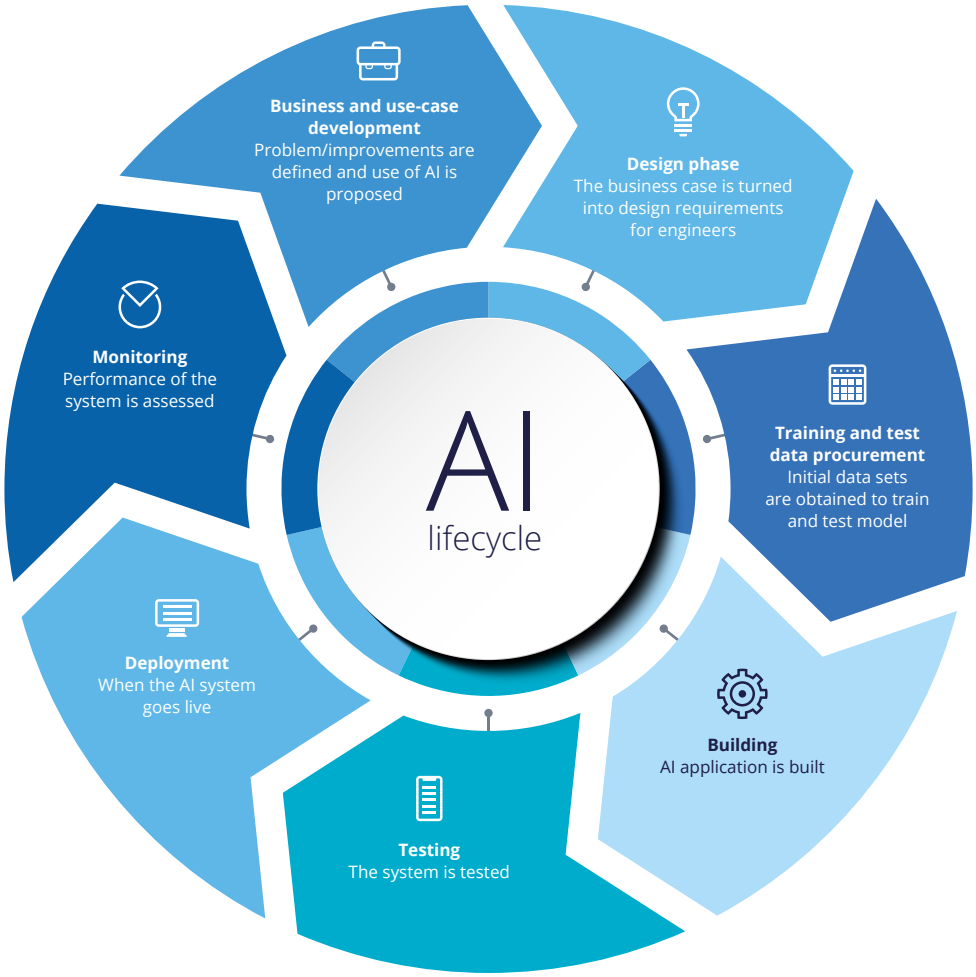
Chapter 6 focuses on the role of the law, and the enforcement of legal rules, to protect human rights in the context of AI-informed decision making. However, the law is not always a sufficient, or the best, way to protect human rights. It is in this context that there is a role for 'soft law', co- and self-regulation and non-legislative measures such as codes of practice, guidelines, directives, industry or sector rules, protocols and other guidance.[570]

These measures can be developed with government ('co-regulation') or by an organisation or body that is independent of government ('self-regulation'). Co-regulation typically involves a legislative framework setting out minimum standards, which is 'supplemented by industry codes or other mechanisms' developed by a non-government entity such as an industry body. These frameworks may then be monitored or validated by a government regulator.[571]

AI is developing quickly, but unpredictably. While the law can be relatively slow to adapt, soft law can adapt more quickly and draw more readily on expertise located outside government. It can also be more easily and speedily changed as technologies develop and the surrounding context changes.

Soft law also offers a range of flexible and adaptable solutions throughout the lifecycle of AI (as depicted in Box 4 below). Decisions made during the earliest stages of AI design, such as selecting a particular dataset to train a machine-learning algorithm, may ultimately lead to a discriminatory outcome when an AI-powered tool is deployed. While the law may offer the best remedy for a discriminatory decision, steps taken early in the AI design process to avoid human rights problems, such as discrimination, can be most valuable of all. Using a mix of measures is more likely to lead to more effective, and comprehensive, human rights protection.

Box 4: Simplified lifecycle of AI[572]



(b) Submissions and consultations

The Issues Paper asked whether non-legislative and other measures are needed to protect human rights in the context of AI-informed decision making and, if so, what those measures should be.

Stakeholders referred to various forms of co- and self- regulation. These include:

- new oversight and monitoring bodies, such as the UK's Centre for Data Ethics and Innovation[573]
- ethical principles and frameworks at the regional organisational level[574]
- design frameworks and principles[575]
- standards for the design of AI products and self-certification[576]
- professional standards and codes of ethics to guide professional behaviour.[577]

(i) Benefits and limitations of co- and self-regulatory responses

There was support among stakeholders for co- and self-regulation to address the human rights and other challenges posed by AI, because they can enable a regulatory response that is agile, flexible and timely.[578]

In addition, some stakeholders linked this with a more co-operative and collaborative approach to protecting human rights in this area. Stakeholders pointed to potential benefits including:

- appropriate ways of obtaining the views and experience of a diverse group of stakeholders, including people who are vulnerable[579]
- the opportunity to demonstrate best practice in technology design and delivery through public/private partnerships or multi-stakeholder initiatives[580]
- sharing of technical and regulatory knowledge.[581]

However, stakeholders emphasised that co- and self-regulation alone would be insufficient to protect human rights.[582] It was suggested that such approaches need to support clear legal rules.[583] On the other hand, some stakeholders argued the law should be the last resort.[584]

(ii) The role of human rights

Some stakeholders noted the central role that human rights could play in guiding co- and self-regulatory approaches to AI. For example, the University of Technology Sydney (UTS) argued:

> [W]hile there is clearly a role for regulatory flexibility, establishing a set of mid-level principles for applying the human rights approach to [AI-informed decision making] can provide a reasonably stable framework (or benchmarks) for evaluating flexible approaches to the regulation of [AI-informed decision making], and then refining those approaches.[585]

As discussed in Chapter 2, there was strong stakeholder support for the UN Guiding Principles on Business and Human Rights as a useful resource to guide self-regulation, particularly as a guide to identifying potential human rights impact throughout the AI life cycle.[586]

7.2 Design, standards and impact assessments

A number of design factors influence how an AI decision-making system engages human rights. These include the decisions it makes, who has designed the system, what datasets are used, how different factors are weighed in the operation of an algorithm, and any process for monitoring and evaluation.

The choices that are made in respect of each of these elements can have a significant impact on individuals' human rights. In turn, co- and self-regulation can help better choices to be made, especially in the design and deployment of AI-informed decision-making systems.

In this section, the Commission considers potential co- and self-regulatory measures, suggested by stakeholders, which could safeguard human rights in the context of AI-informed decision making. Specifically, this section considers the emerging concept of 'human rights by design', the possibility of developing standards for the development of AI-powered products and services, the function of human rights impact assessments, and the idea of certification or a trustmark for AI-powered products.

(a) 'Human rights by design'

Human rights can be advanced through the design and development process. 'Inclusive design', as discussed in Chapter 10, can enhance the accessibility of products and services for people with disability. In a similar way, 'privacy by design' aims to ensure privacy is protected through good development practices. The OAIC described privacy by design as being

> about finding ways to build privacy into projects from the design stage onwards and is a fundamental component of effective data protection. This involves taking a risk management approach to identifying privacy risks and mitigating those risks. In applying this approach, entities take steps at the outset of a project that minimise risks to an individual's privacy, while also optimising the use of data.[587]

Similarly, the Australian eSafety Commissioner has developed 'Safety by Design' principles, guiding companies to use a design process that will assess, review and embed user safety into online services, providing an achievable voluntary standard for private industry.[590]

While there is support for the concept, there is, as yet, no definitive statement about how to implement human rights by design. Some have suggested drawing from related concepts, such as the human rights due diligence approach undertaken by some mining companies, adapting this to the context of designing AI-powered products.[591]

There was considerable interest among stakeholders in how 'human rights by design' might apply in the context of AI-informed decision making. As the Gradient Institute and others have observed, the capacity to use AI to create 'good at massive scale' relies, in part, on design choices. The Institute notes:

> What we have to do is clear: research which design choices for machine learning will lead to more ethical outcomes, apply the research findings to build and spread decision-making systems that are more ethically-aware, and educate individuals and society so they can become active contributors in a world shaped by AI.[592]

Some stakeholders supported elements of human rights by design in this area.

Such an approach could start by rigorously assessing and then addressing the risk of social and individual harm for any proposed AI-informed decision-making system,[593] including the impact on particular rights such as privacy.[594]

Some emphasised the importance of accountability in the operation of such systems, which would require these systems to at least produce explainable or understandable decisions.[595]

Several stakeholders said that a 'human rights by design' approach should also involve considering the composition of design teams: they should be diverse, inclusive and should actively consult with disadvantaged and vulnerable groups.[596]

Some stakeholders observed that 'human rights by design' is an emerging concept, and were interested in knowing more about this and related concepts.[597] Some explicitly supported further research and consideration of this concept.[598]

There was a range of views about how to ensure that 'human rights by design' achieves its aims in relation to AI-informed decision-making systems. Some stakeholders suggested an industry-wide approach,[599] while others argued for design standards to be codified in law.[600]

There are many ways in which human rights by design can be encouraged or enforced. The application of design principles could be entirely voluntary, or there could be a legal requirement to comply with design standards developed by an industry body, for example. However, design standards supported by legislation would be a more robust model for human rights protection.

(b) Standards for products and services that use AI

There is growing interest in the role of standards in the design of AI-informed decision-making systems. Such standards can apply to *categories of people*—for instance, in codes of conduct for professionals who develop such systems. This type of standard is dealt with below, in relation to professional education and training.

As discussed in this section, standards can also apply to *products and services*—for instance, in setting out core design requirements for AI-informed decision-making systems.

Standards Australia, a non-governmental organisation that is also a member of the International Standards Organisation, is currently consulting on the development of standards for AI.[601]

The Australian Government has not given official backing to any specific enforceable standards for AI. It has, however, recently published:

- the AI Ethics Principles, a voluntary and 'aspirational' set of principles that can be used by anyone when designing, developing, integrating or using AI[602]

- a draft voluntary Code of Practice for industry, published for consultation, that provides guidance in designing IoT products.[603]

These initiatives may, in time, result in the development of enforceable standards for products and services that use AI.

There are some Australian standards that relate to technologies more generally. The Digital Transformation Agency (DTA), for example, has developed the 'Digital Services Standard', which recommends a co-design approach to the digital delivery of certain government services.[604]

Standards specific to AI have begun to emerge from professional bodies with international membership. In 2019, for example, the IEEE produced non-binding standards for ethically-aligned design of AI systems.[605] These high-level principles offer specific guidance for standards, certification, regulation or legislation for the design and manufacture for the use of autonomous and intelligent systems.[606]

There are also examples of industry-led self-regulatory approaches to develop standards for AI-related technologies. The Global Network Initiative, for example, is an international membership organisation made up of internet and telecommunications companies, human rights and press freedom groups, investors, and academic institutions. Participants commit to implement the Initiative's Principles on Freedom of Expression and Privacy, which are informed by international human rights law and the UN Guiding Principles on Business and Human Rights.[607]

Some stakeholders supported the idea of non-binding standards to protect human rights, among other goals, in the design of AI-informed decision-making systems.[608]

Any new Australian standards should draw on the work of international standard-setting bodies,[609] particularly given that much of the research and development expertise regarding AI is located overseas.[610]

Good consultation in the development of standards is clearly important, and stakeholders emphasised that the public and private sectors should both be engaged.[611] Effective, in-depth and cross-sectoral consultation could also help to determine whether sector- or industry-specific guidance or standards are needed.[612] For example, there may be a strong case for industry-specific standards for consumer and privacy protection,[613] such as the use of AI in health settings.[614]

Some stakeholders pointed out that non-binding standards can only achieve so much. Where appropriate, standards should be considered along with other processes that aim to encourage particular behaviour, such as rules relating to procurement of AI-powered products and services,[615] or AI impact assessments (discussed below).[616]

(c) Certification and 'trustmark' schemes

Standards and other such initiatives that aim to advance human rights protections, or other similar goals, exist on a regulatory spectrum. At one end, they can be made mandatory by legislation. At the other end of the spectrum, these initiatives can be implemented by industry agreement or made wholly voluntary.

A further option is to tie such measures to certification or 'trustmark' schemes. These schemes are not legally binding, but they can incentivise compliance, including by influencing consumer choice.

For example, the GDPR encourages Member States to establish 'data protection certification mechanisms' and 'data protection seals and marks' to demonstrate compliance with the Regulation.[617]

Some stakeholders expressed interest in the use of trustmarks to certify products and services using AI as ethically or human rights compliant.[618] Several submissions specifically referred to the idea advanced by Australia's Chief Scientist, Dr Alan Finkel, to establish a 'Turing Stamp'—a proposed voluntary certification scheme for 'ethical AI' that would be independently audited.[619]

Work is being undertaken in this area by the OAIC. The OAIC noted in its submission that it is currently considering the role of certification, seals and marks as an accountability mechanism to better protect personal information as defined by Australian privacy law. The OAIC is also monitoring the implementation of trustmark provisions in the GDPR, as noted above.[620]

Other stakeholders opposed the use of certification for AI. For example, some expressed concern about the diminished value of a trustmark in a market with a small number of dominant players and therefore a small number of possible products.[621] It was also observed that AI is not limited to a single industry, making it more challenging to develop an effective trustmark system—in contrast to initiatives such as the 'Fairtrade' trustmark, which applies only to agricultural products and a small number of manufactured goods.[622]

(d) Impact assessments

Especially in areas that involve significant risk of harm, new products and services must be rigorously assessed and tested *before* they are used by or on humans.[623] At the very least, such testing should identify and address harm before these new products and services are used widely.

A harm prevention principle should apply equally to products and services using AI. However, the technology industry has progressed rapidly, in part by speeding up traditional research and development processes, with a view to producing market-ready products and services more quickly.

One illustration of this phenomenon is the 'minimum viable product' (MVP) approach, which involves swiftly developing technology products that can be made publicly available in beta or trial form, and then making iterative changes by observing how the products are used in the real world. A risk of the MVP approach is that such products can cause harm as soon as they are made publicly available.

Some have responded to this problem by proposing better assessment and testing of the human rights impact of new AI-powered products and services, including decision-making systems. For example:

- **Human rights impact assessments** (HRIAs) consider the likely impact of a new product or service on the human rights of affected people, and can be used to address human rights risks. In Australia, private sector organisations are not required to undertake HRIAs. However, such assessments are required for proposed legislation. The Minister or other parliamentarian responsible for tabling a proposed new law in the Australian Parliament must prepare a statement of human rights compatibility, which considers how the draft law engages human rights.[624]
- **Privacy impact assessments** (PIAs) measure the likely impact of a product or service on the privacy of individuals, and make recommendations for 'managing, minimising or eliminating that impact'.[625] Certain activities that involve the collection, storage or use of personal information may require a PIA under the *Privacy Act 1988* (Cth).[626]
- **AI or algorithmic impact assessments** (AIAs) are specifically targeted towards the use of techniques or technologies associated with AI. For example, Canadian Government departments must undertake an AIA before using an automated decision system.[627]

  An AIA considers the likely impact of the system on the rights, health or well-being of individuals or communities; economic interests; and the ongoing sustainability of an ecosystem. Where a system is likely to have a 'high impact', which is defined as decisions that 'will often lead to impacts that are irreversible, and are perpetual', the decision 'cannot be made without having specific human intervention points during the decision-making process; and the final decision must be made by a human'.[628]

Some HRIAs are undertaken voluntarily by companies. Microsoft, for example, has conducted a human rights impact assessment of its AI products, using the process to help refine the company's 'understanding of AI's potential human rights risks and develop mitigation strategies to augment human-centred AI impacts'.[629]

Several stakeholders referred to these sorts of impact assessments,[630] with some emphasising their utility in promoting transparency and 'pre-emptive troubleshooting'.[631] There was specific interest in applying impact assessments to AI-informed decision-making systems, from the earliest stages of policy development and particularly where there is a perceived or actual risk to human rights.[632]

Principle 18 of the UN Guiding Principles suggests that businesses be proactive in identifying the human rights risks or impacts of their activities and relationships. This activity should take place periodically, at all stages of design, development and implementation.[633]

The AI Now Institute has published an AIA prototype, which builds on HRIA models. It aims to support affected communities, stakeholders and government agencies to assess the likely impacts of decision-making systems that use AI, and 'to determine where—or if—their use is acceptable.'[634]

Such impact assessment tools can help to identify human rights problems and other potential harms at an early stage, before they have a real-world impact.[635] The assessment process can itself build up expertise, or highlight where there is a lack of knowledge and thereby build capacity, particularly in public agencies that are required to be accountable.[636]

7.3 The Commission's preliminary view

(a) The benefits of co-regulatory and self-regulatory approaches

The Commission agrees with stakeholders that co- and self-regulatory approaches are important to protect human rights in the context of AI-informed decision making. The flexibility and adaptability of these forms of regulation may be difficult to achieve solely by legislation.

Currently, the expertise, knowledge and ownership of data lie largely with technology companies. A co-regulatory approach is needed to facilitate collaboration and co-operation between the private sector and government and to ensure AI-informed decision making is human rights compliant.

The Commission recognises the emergence of private sector initiatives. These include, for example, the development of products designed to promote human rights, such as AI-powered recruitment tools seeking to promote equal opportunity in employment, open source auditing tools to detect bias in algorithmic decision-making systems, and human rights impact assessments and annual audits. Some of these private-sector initiatives are also expressly designed to comply with the UN Guiding Principles.

One missing piece is a mechanism to scale up good examples of industry tools or approaches across sectors, nations and globally. At the same time, the Commission is concerned about overreliance on self-regulatory, voluntary measures to address the human rights impact of technologies, including AI. Some centralised co-ordination and leadership will be necessary to apply a smart mix of co- and self-regulatory approaches, along with legislation.

(b) A co-ordinated, cohesive approach

During the consultation, some stakeholders expressed the desire for better guidance in assessing the impact of an AI product or service.

A toolkit is being produced by the Australian Government Department of Industry, Innovation and Science to guide industry on the ethics of AI, and work is being undertaken by Standards Australia to develop standards for the application of AI Australia.[637]

The Commission urges that human rights be embedded in these processes. A co-regulatory approach to design, standards and certification should be adopted more generally in relation to AI-informed decision making.

A co-regulatory approach would allow for industry to be more closely involved, and for cross-sectoral education. But this should not be a substitute for regulation through law when this is needed.

This accords with the conclusion of the IEEE that standards and regulatory bodies are needed to ensure automated and intelligent systems do not infringe on human rights, and the need to translate existing and new obligations 'into informed policy and technical considerations' on a global scale.[638]

In turn, this view reflects the requirement in the UN Guiding Principles that States provide effective guidance on how to consider potential human rights impact throughout the operation of a system.[639]

The Commission also recognises the significant challenges associated with establishing 'human rights by design' principles, standards and certification schemes—whether using a co-regulatory approach or otherwise.

For example, many AI applications used in Australia are developed overseas. It is therefore necessary not only to think about what standards might be developed in Australia, but how any standards or certification would be applied to an AI-powered product purchased from offshore. This challenge may be met in a number of ways. For example, it could be considered as part of an Australian certification scheme or safeguards could be inserted in the procurement process. Training and education about these tools will also be needed.

The Commission has not reached a firm conclusion regarding the potential for design, standards and certification to protect human rights in the context of AI-informed decision making.

More work needs to be done to identify what these measures may look like, how they would work in practice and how these approaches could be scaled up to achieve a regularised or uniform approach. There is also a need for coordination across and within government, non-government organisations and private companies.

The Commission proposes that a multi-disciplinary taskforce be established that includes membership from representative industry bodies, relevant government departments, experts in law and policy and technical experts. Importantly, this taskforce should consider how human rights protection can be embedded across all co- and self-regulatory measures.

This taskforce could be supported by a new AI leadership body, as proposed in Part C, or it may form part of the work already being undertaken by the Australian Government.[640]

PROPOSAL 12:

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

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.

(c) Human rights impact assessments

In comparison to human rights by design, standards and certification, HRIAs are more established. Types of HRIA in Australian law and practice include:

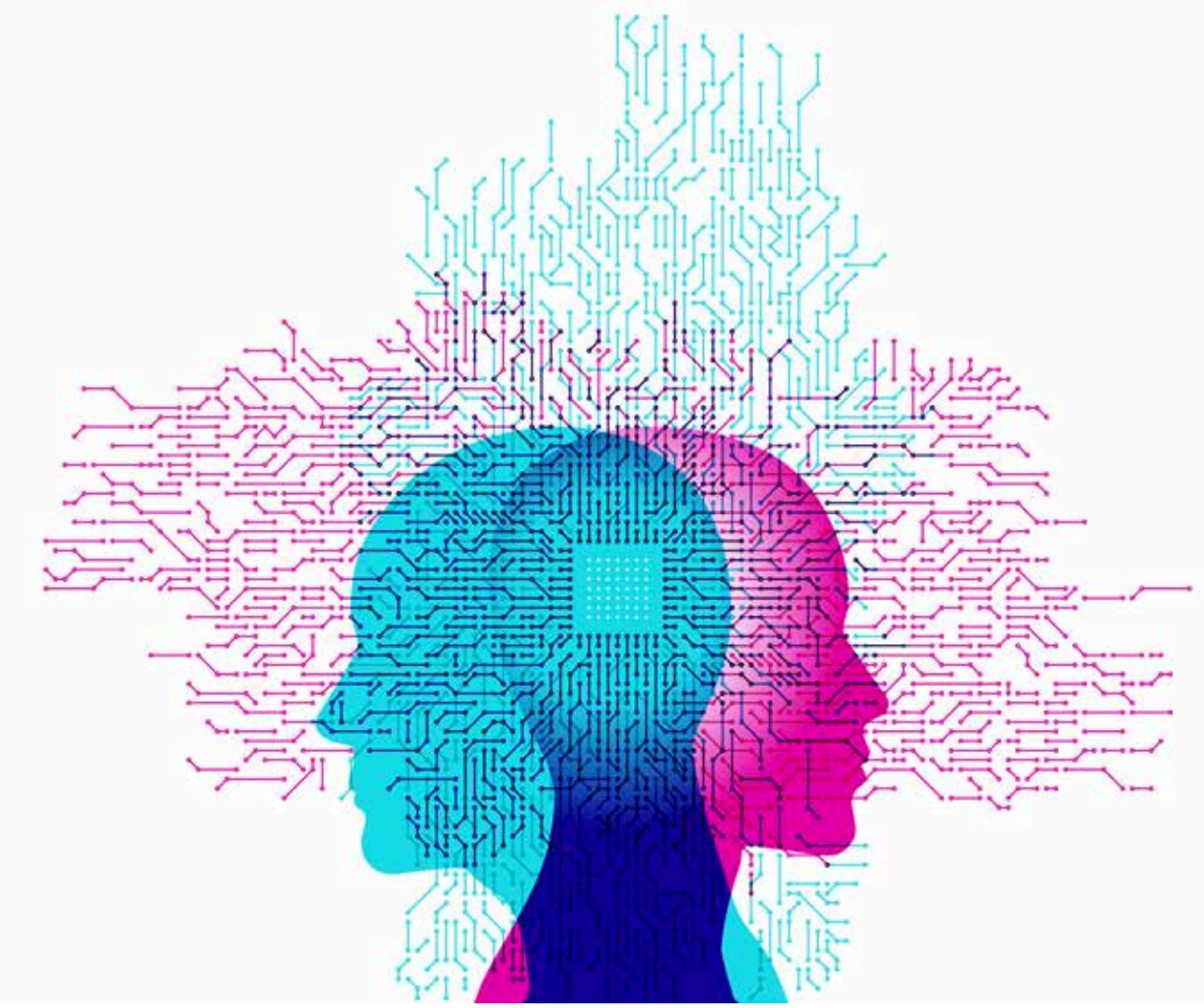
- as part of Australian parliamentary processes[641]
- privacy impact assessments under privacy legislation[642]
- as part of a human rights due diligence and risk assessment conducted under modern slavery legislation.[643]

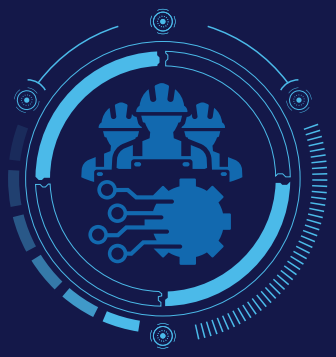
There are also solid examples of HRIAs being undertaken in the context of AI, including algorithmic and automated decision making. These initiatives have been established by private technology companies;[644] governments[645] and non-government organisations,[646] both here and overseas.

The Commission supports the use of HRIAs in the development of AI-informed decision-making systems, because this could assist developers (including companies and governments) to identify the potential harm or benefit of a proposed AI product, and put in place measures to address or mitigate any risk to human rights. A HRIA tool should be developed in partnership with industry and civil society.

Some support for HRIAs in this area was offered during the Australian Government's consultation on ethical AI. Specifically, Data61 proposed that impact assessments, 'which address the potential negative impacts on individuals, communities and groups' be included in a 'toolkit for ethical AI'.[647]

In the Commission's view, such impact assessments should explicitly incorporate consideration of the impact on human rights. This accords with the Government's recently published AI Ethics Principles, the second principle referring to the need for AI systems to respect human rights, diversity and personal autonomy throughout the AI lifecycle, including the careful consideration of risk.[648]





‘A human rights impact assessment is needed before any new AI-informed decision-making system is deployed.’

A number of questions should be considered as part of a process to develop an HRIA tool for AI-informed decision-making systems. These include:

- For what AI-powered systems should an HRIA be used?
- In what, if any, situations should an HRIA be mandatory or encouraged (eg, through a government procurement process, or as part of a certification process)?
- How and when should an HRIA be applied to AI-informed decision making tools developed wholly or partially overseas?

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.

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?**

7.4 Regulatory sandboxes

In a 'regulatory sandbox', new products or services can be tested in live market conditions but with reduced regulatory or licensing requirements and exemption from legal liability, and with access to expert advice and feedback.[649] It allows for dialogue with policy makers, and enables regulators 'to try out new rules and observe their impact on the technology in an environment where wider damage or danger to the public is limited'.[650] The goal

> is to relax or change existing regulation in a controlled and evaluated space to run real-world experiments. These experiences can be collected and inform evidence-based regulatory schemes. Sandboxes are used as an alternative to regulation that is based on speculation about what behaviours could result—and what risks and harms can emerge—from changing technologies or changing policies.[651]

Some see regulatory sandboxes as a way of supporting innovation, while enabling regulators to test new regulatory models. Regulatory sandboxes are being applied to products and services using AI, including in Singapore,[652] India,[653] Finland[654] and in the European Union.[655]

In Australia, regulatory sandboxes have been used in the context of technology that enables or supports banking and financial services (also known as FinTech). For example, ASIC has created a regulatory sandbox that allows certain FinTech products or services to be tested without a licence.[656] ASIC's aim is to foster innovation, while also promoting 'good outcomes for consumers' and recognising that poor outcomes will erode consumer trust and confidence.[657]

Some other jurisdictions have established FinTech sandboxes.[658] For instance, the UK's Financial Conduct Authority established the first FinTech regulatory sandbox in 2016, testing products such as 'robo-advice', and the use of facial recognition technology by a financial adviser to conduct risk profiling.[659] Beyond FinTech, the UK Information Commissioner's Office recently established a regulatory sandbox in the context of data protection.[660]

The claimed benefits of regulatory sandboxes include that they can: reduce regulatory uncertainty for innovators; provide an opportunity to discuss, in confidence and candidly, all potential usages of new technologies; lead to early warning that a feature of a new product may not be acceptable, allowing that feature to be modified.[661]

On the other hand, some criticise the concept because, among other things, regulatory sandboxes: can create a risk or perception that regulators inappropriately support or favour certain tech start-ups; may create limited safeguards against harm to individuals; and can lower barriers to entry.[662] In the specific context of AI, particular difficulties can arise including a tension between the need to be candid, or transparent, with a regulator, and the commercial sensitivity of an opaque algorithm.

Some stakeholders suggested establishing 'regulatory sandboxes' in Australia to test certain products or services that use AI.[663] The Law Council of Australia, for example, recommended

> serious consideration be given to the creation of a 'regulatory sandbox' or sandboxes that expressly address human rights as key criteria for assessment of the given technology. Consideration should also be given to adding human rights and privacy as additional references for existing sandboxes.[664]

7.5 The Commission's preliminary view

All research and development involves some level of prediction regarding how a new product or service will be used and its likely effects. As AI is relatively new, such predictions can be more difficult. That difficulty is exacerbated by a common ethos in the technology industry, which has been described as 'move fast and break things'.[665] That ethos can inhibit testing of AI-powered products and services before they are publicly released.[666]

Inadequate testing can cause harm, including to human rights and public trust. High-profile examples include an AI-powered chatbot that made racist statements,[667] and a facial-recognition application that mislabelled some African-American people as gorillas.[668]

Anticipatory regulation, however, is inherently difficult in the context of rapidly-evolving technology like AI. It can lead to laws that do not sufficiently address human rights risks, or laws drawn so broadly that they undermine innovation.

In this context, the Commission sees benefit in an Australian regulatory sandbox that focuses on assessing the human rights impact of AI-informed decision-making. This could help to encourage such systems to be tested more rigorously than is currently the case, and assist in developing effective regulation in this area.

The experience overseas suggests that regulatory sandboxes can enable multi-disciplinary, multi-stakeholder collaboration and co-operation. ASIC's regulatory sandboxes for FinTech products show how this can work.[669] However, there are no regulatory sandboxes to test the human rights impact of AI-informed decision-making systems.

The OECD AI Principles, to which Australia is a signatory, support the use of controlled testing, with governments called to

> promote a policy environment that supports an agile transition from the research and development stage to the deployment and operation stage for trustworthy AI systems. To this effect, they should consider using experimentation to provide a controlled environment in which AI systems can be tested, and scaled-up, as appropriate.[670]

The Commission is currently partnering with the Gradient Institute, Data61 and the Consumer Policy Research Centre to detect and address algorithmic bias. This work, outlined below, could assist in the development of a new regulatory sandbox.

Box 5:

Algorithmic bias experiment

In partnership with CHOICE, Consumer Policy Research Centre, Data61 and The Gradient Institute, the Commission is undertaking an experiment to better understand the risk of algorithmic bias. The experiment will involve the building of a data-driven algorithm, using a synthetic dataset with relevant input variables such as year of birth and postcodes, and testing the decisions produced by the algorithm for discriminatory impact. The experiment will vary the input variables and explore trade-offs in results, investigate what factors may produce more or less accurate results, or more or less fairness, and analyse these results from a technical, human rights and consumer rights perspective.

The Commission acknowledges that the concept of the regulatory sandbox is in its infancy, and further consultation is needed.

PROPOSAL 15:

The Australian Government should consider establishing a regulatory sandbox to test AI-informed decision-making systems for compliance with 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 eg, 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 eg, transparency and protection of trade secrets?**
- **(h) how should the regulatory sandbox be evaluated?**

7.6 Education and training

Decision-making systems that use AI still involve humans in various ways. Most obviously, humans are involved in designing, deploying, regulating and overseeing such systems. Most commonly a human will make the ultimate decision, but the human will take into account recommendations or information generated using AI (AI datapoints).

However, most Australians do not know what functions can be reliably undertaken using AI, nor are they properly equipped to test the reliability of AI datapoints. As a result, humans can be either too deferential or too dismissive of AI datapoints. In other words, humans can perceive AI datapoints to be more reliable, or less reliable, than is actually the case.

A separate problem is that those people most involved in designing and developing products and services that use AI are often not trained in assessing their likely impact on disadvantaged or vulnerable groups.

Some stakeholders have observed that education and training could play an important role in addressing these and other problems.[671]

(a) Targeted education, training and capacity building

Many stakeholders emphasised the importance of education and training on the human rights implications of AI.[672] This would need to be tailored to different needs and contexts, and might include education for:

- people who design and build decision-making systems that use AI,[673] with a view to building competence regarding the 'intended, known, and *unintended* consequences of innovation'[674]
- public servants procuring and using AI-powered products and services[675]
- officers of regulatory bodies that oversee how AI is used[676]
- the community as a whole, with particular emphasis on those most affected or at risk,[677] such as children and young people.[678]

The needs of these groups will be different, but each will need to be equipped with sufficient knowledge about AI and human rights for their particular purposes.

(i) Professionals who use AI in design and development

Those who design and develop products and services that use AI should be able to assess their likely human rights impacts. Professional bodies that regulate or support groups such as computer scientists, engineers and data scientists, have begun to recognise the need to educate their members on these issues.[679] Guidance and education should be consistent across different but related professional groupings.

Guidance developed by the IEEE, for example, acknowledges that responsible innovation requires designers of autonomous and intelligent systems to 'anticipate, reflect and engage' with users of these systems, in order to design accountable systems and proactively guide 'new technology toward beneficial ends'.[680]

Continuing professional development on these issues may also be needed. This could be incorporated into pre-existing mechanisms, such as training accredited by industry bodies and codes of practice implemented by industry associations and other bodies that administer professional standards.[681]

While there is some guidance regarding responsible development of AI,[682] there are no mandatory professional standards to which professionals designing and developing AI must adhere.[683] Some professional bodies, such as the IEEE, maintain that students should learn what responsible or human rights compliant AI looks like before entering the workplace.[684]

(ii) Government officers and administrative staff

Where government uses AI to make or automate decisions, it is particularly important to have in place a structure that rigorously predicts, evaluates and monitors risks, including risks to human rights.[685] Tools, such as the HRIAs discussed above, can help this analysis. However, to be effective these tools require a level of knowledge among policy makers, public servants and others with responsibility for addressing risks of harm.[686]

Government officials have limited guidance on issues that need to be considered when considering a new AI-informed decision making system.[687] However, the Commission is not aware of any training or knowledge requirements for government employees involved in purchasing or implementing AI-informed decision-making systems. The need to upskill government employees was noted in some submissions,[688] and has been the subject of growing academic and professional commentary.[689]

(iii) Professionals relying on AI

The Australian Council of Learned Academies observed that education and capacity building are needed not only for those working in the technology industry, but also those who rely on AI datapoints to make decisions.[690] It is vital that those human decision makers have the knowledge and skills to evaluate AI datapoints in the contexts in which they work.[691]

For example, it is increasingly common for AI datapoints to be used for decision making in the criminal justice sector, such as to assist in making bail or sentencing decisions. Judicial officers and others need to understand how to assess this sort of information, so that they can determine its reliability and probative value. However, some stakeholders have observed that the legal profession is generally not trained in these skills.[692]

The need for education and training could extend also to company directors who require knowledge of AI and human rights in order to identify and manage risks in respect of any AI-informed decision-making system being deployed by their organisations.[693]

(iv) Education of the public and affected groups

There is limited community understanding about AI, how it is used, and its impact on individuals. Recent polling undertaken for the World Economic Forum indicated significant sections of the community do not know how or why AI is used, but felt uncomfortable with its use in areas such as government service delivery.[694] The IEEE has noted that, given the power of AI tools, 'there is a need for a new kind of education for citizens to be sensitised' to the associated risk.[695]

Some stakeholders urged government to consider education campaigns to build consumer AI literacy.[696] Public education about AI and human rights could assist people to understand, question and challenge decisions made by, or using, AI. The University of Technology Sydney (UTS) framed education as empowerment. Its submission observed that AI is often considered 'something that others design, or is too big to change, leaving us … as passive recipients'.[697]

As AI becomes increasingly central to so many aspects of life, an understanding of AI and its social implications will be vital in navigating the world. This suggests there would be benefit in school education incorporating some of the foundational principles relating to AI and human rights. Bodies such as the Australian Council of Learned Academies and the Australian Government Department of Education have expressed support for this idea.[698]

The need for more comprehensive digital literacy education in Australian schools is an issue that has frequently been raised by the National Children's Commissioner—particularly in relation to issues such as children's exposure to online pornography and cyberbullying.[699]

Experts have started to suggest how such education could be delivered. Professor Lyria Bennett Moses, for example, has argued for a shift from the focus on traditional science, technology, engineering and maths (known collectively as STEM) subjects 'to truly interdisciplinary learning models'.[700]

Education has also been identified as necessary to create a diverse workforce to design and develop AI, including by supporting the role of students from culturally and linguistically diverse backgrounds, who can be disproportionately impacted by AI-informed decision making.[701]

7.7 The Commission's preliminary view

As noted above, recent polling suggests the Australian community is generally unaware of how AI-informed decision making affects them and what the implications may be for their human rights. In addition, there is little training for those who design and develop AI-informed decision-making systems regarding the potential social or human rights impact.

Finally, there is minimal education about the technologies involved for those who procure, use and oversee AI-informed decision-making systems and tools.

This can contribute to a failure to accurately perceive or effectively address risks of harm.

Addressing these issues will require a range of actions. One key requirement is to build our nation's collective capacity to understand the implications of AI-informed decision making. To this end, the Commission proposes the development of a comprehensive plan for Australia regarding education and training on AI and human rights. We refer to this as the 'AI Education Plan'.

As with any field of knowledge, there are different educational or knowledge needs for different parts of the community. Not everyone requires the same level of knowledge about how AI operates, or its human rights implications. Accordingly, the education component of the proposed *National Strategy on New and Emerging Technologies* should be carefully targeted by reference to the particular needs of different groups. Broadly speaking, this might be divided into the following sorts of categories.

Group	Objective and outcome	Leadership to develop educational modules
General public	All individuals will need a basic understanding of AI and how it is being used. Training should enable the community to make informed choices about how they interact with organisations that use AI in decision making, and the accountability mechanisms available to ensure that decisions that affect them are lawful and respect their human rights.	Government Schools Civil society Industry bodies with a public education function
Decision makers who rely on AI datapoints	Decision makers who use AI datapoints to make decisions need sufficient knowledge of the reliability of various types of AI datapoints to make good decisions. This is likely to include an expanding group of decision makers, such as judicial officers, insurance brokers and police officers.	Employers of decision makers Professional regulatory and organisational bodies that are involved in professional development
Professionals requiring highly-specialised knowledge	Designers and developers of AI-informed decision-making systems need to be able predict the likely impact of their work, including how to incorporate human rights in the design process. The accreditation or certification of professionals who design AI should be considered. Regulators, policy makers and those who commission AI-informed decision-making systems need to understand how AI operates and its likely impact on human rights in the context in which systems will be deployed.	Government University curriculum Professional regulatory and organisational bodies Technology companies

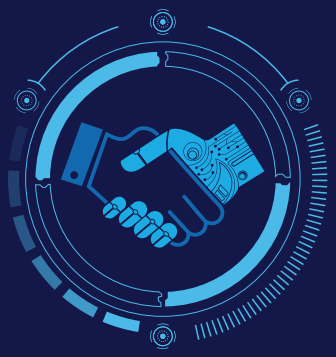
PROPOSAL 16:

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 datapoints and professionals designing and developing AI-informed decision-making systems.

7.8 Oversight of government use of AI

As discussed in Chapter 6, government use of AI to inform or make decisions should be transparent and subject to oversight. Administrative law provides one form of accountability for some government decision making—whether or not it is made using AI. However, accountability is primarily directed towards identifying and remedying errors made in respect of individual decisions.

There is a separate need for systemic oversight of government's use of AI-informed decision-making systems. Australia's federal, state and territory governments are already among the most significant actual and likely developers and purchasers of such systems. It is important to ensure that these systems respect and protect human rights.



‘Develop a national AI education plan focused on AI and human rights.’

(a) Auditing government use of AI

Given the growth of AI in a range of contexts, it will be important to monitor government’s use of AI-informed decision-making systems.

Currently, there is no systematic auditing process, or public register, regarding the use, or planned use, of AI in government decision making. However, the Auditor-General and the Australian National Audit Office (ANAO) do undertake ‘performance audits’, which involve ‘the independent and objective assessment of all or part of an entity’s operations and administrative support systems’, by reference to cost, efficiency, effectiveness and legislative and policy compliance.[702]

Performance audits consider some technological issues. For example, the ANAO has conducted an audit assessing the effectiveness of the rollout of My Health Record by the Australian Digital Health Agency.[703] It is conceivable, therefore, that the use of decision-making systems that rely on AI could be brought within the scope of such an audit process.

In 2004, the Administrative Review Council recommended automated decision making by government be audited, to support ‘transparency, fairness and efficiency’.[704] In a similar vein, the Commonwealth Ombudsman suggested government agencies establish a panel to oversee major digitisation projects, which should include the ANAO.[705]

There are increasing calls for audits of government and private sector use of AI overseas. In 2018, the US-based AI Now Institute urged governments to ‘oversee, audit, and monitor’ the use of AI in different sectors as a matter of priority. Some private technology companies incorporate audits—with publicly available results—as part of ‘internal accountability structures that go beyond ethics guidelines’.[706] The IEEE has suggested that AI systems themselves ‘generate audit trails recording the facts and law supporting decisions and they should be amenable to third-party verification’.[707]

There are, however, numerous challenges to auditing in this area. For example, the prevalence of so-called 'black box' or opaque AI can make auditing difficult, if not impossible.[708] Capacity building might also be needed for bodies that would conduct this new auditing function and in order to establish industry standards for auditing the use of AI.[709]

(i) Submissions and consultations

Some stakeholders suggested auditing the use of AI by government and others,[710] including by a new regulatory body or authority.[711] Some stakeholders referred to examples from other jurisdictions,[712] and other sectors,[713] which could be used to guide the development of AI auditing.

In addition to enhancing transparency regarding the government's use of AI in decision making,[714] some stakeholders suggested that regular auditing could help measure the impact, including the social impact, of AI.[715] Auditing tools, including tools that specifically target algorithmic bias, are being developed by technology companies.[716] These tools could also help to identify any discriminatory impact of particular AI-informed decision-making systems.[717] However, one stakeholder cautioned that auditing alone, without a clear regulatory framework, would be insufficient to protect human rights.[718]

(ii) Oversight and evaluation: establishing a 'human in the loop'

An ongoing audit requirement for use of AI by government would need to be supported by some form of human oversight and evaluation.

Some stakeholders pointed to the need for effective evaluation of the use of AI by government and others, particularly to assess the impact on vulnerable groups.[719]

Where AI is being used to make, or support, a decision that has an impact on an individual's human rights, there is a persuasive argument that human oversight is necessary. This is often referred to as a 'human in the loop', or 'data humanism'. Human oversight and evaluation of an AI decision-making system could take a number of forms and take place at various points in the implementation and operation of the system. Examples include:

- periodic review of a sample of outputs to evaluate accuracy and quality, and ensure decisions being made are not discriminatory or inappropriate
- retraining where alerted that a decision-making system's outputs are unpredictable or incorrect
- ensuring a human is able to manually override a decision-making system in certain circumstances.[720]

Not every AI system necessarily requires human oversight and evaluation. The Canadian Government, for example, in its recent directive on the use of AI by government, established a risk assessment to identify where human oversight is required. The directive was established in recognition of the increasing use of AI to make, or assist in making, administrative decisions and the Canadian Government's commitment to do so 'in a manner that is compatible with core administrative law principles such as transparency, accountability, legality and procedural fairness'. The requirements of the directive are outlined in further detail in Box 6 below.

Box 6:

Canadian Government Directive on Automated Decision-Making[721]

The Canadian Government has issued a directive to ensure automated decision systems are deployed with minimal risk and are used to achieve 'more efficient, accurate, consistent, and interpretable decisions made pursuant to Canadian law'. It will apply to all government procurement from 1 April 2020.

The **expected results** of the directive are: federal government decisions will be data-driven, responsible and procedurally fair; algorithmic impact on administrative decisions will be assessed and any negative outcomes reduced; data on the use of AI by government will be made available to the public, where appropriate.

The **mechanisms** used to achieve these expected results include:

- a public algorithmic impact assessment
- transparency, guaranteed by providing notice that AI is being used, providing a 'meaningful explanation' to affected individuals of how and why the decision was made
- a government right to access and test the AI system for a range of purposes including an audit, investigation or judicial proceeding
- testing and monitoring of outcomes
- validation of the data set to ensure it is relevant, accurate and up-to-date
- employee training in the design, function and implementation of the AI decision making system
- human intervention where appropriate according to the risk level and potential impact on an individual.

(b) AI procurement by government

Procurement in Australia is governed by legislation, rules, guidance and frameworks.[722] The 'core rule' underpinning the Australian Government's procurement system is achieving 'value for money'.[723] Certain procurement processes by Australian Government departments and agencies must comply with the Commonwealth Procurement Rules made under the *Public Governance, Performance and Accountability Act 2013* (Cth).[724]

Large procurement of information and communications technology is supported by the DTA. The DTA's framework for ICT procurement includes best-practice principles to guide a government agency's sourcing of ICT products and services. These principles include 'encouraging competition', and urge agencies to be 'innovative and iterate often', be 'outcomes focused' and 'minimise cyber-security risks'.[725]

It is increasingly common to use government procurement processes as a lever to influence behaviour to achieve policy outcomes. For example, in the *2030 Plan for Australia to thrive in the global innovation race*, Innovation and Science Australia identified government procurement to be a 'strategic lever' to stimulate and promote innovation.[726]

Similarly, a UK parliamentary committee recommended 'targeted procurement to provide a boost to AI development and deployment' and the use of procurement processes 'to capitalise on AI for the public good'.[727] The World Economic Forum is also partnering with governments, including the United Kingdom, to develop detailed guidance for public sector procurement of AI.[728]

(i) Submissions and consultations

Some stakeholders identified that the procurement of new technologies by the public sector could influence positive business behaviour, demonstrate best practice and provide practical safeguards for human rights.[729]

Stakeholders made a number of suggestions for how this could work in practice, including:

- inserting a 'social clause', such as an assessment of a company's human rights record, in public procurement contracts to protect human rights[730]
- tying government funding to compliance with minimum human rights standards, to encourage the development of ethical AI[731]
- including an auditing requirement for new AI-powered decision-making systems[732]
- requiring decision-making systems that use AI to be tested in a real world application as part of the procurement process[733]
- requiring an HRIA to be undertaken in an early stage of procurement.[734]

The need to upskill public sector workers involved in procurement to deliver these kinds of objectives was also noted.[735]

7.9 The Commission's preliminary view

Government decisions about housing, health services, child care, employment benefits and criminal justice can be hugely significant for affected individuals and communities, and particularly for those who are already disadvantaged or vulnerable.

By using AI to make these kinds of decisions, there is the prospect that governments will be able to make better, more data-driven decisions—with resulting social benefit. However, the use of AI can also lead to decisions that are inconsistent with human rights.

To engender public trust, decisions made using AI should be human rights compliant, explainable, procedurally fair, and subject to effective monitoring and oversight. Safeguards for human rights protection will need to be adopted at the earliest stages of policy development right through to deployment and evaluation.

As outlined in Chapter 6, the Commission proposes three key principles for the development and use of systems for AI-informed decision making.

First, AI use should comply with international human rights law; secondly, AI should be used in a way that minimises harm, including by appropriate and effective testing and monitoring; and, third, AI should be accountable in how it is used.

Applying these principles, the Commission has three preliminary conclusions in this context:

1. Government decision making that uses AI should be transparent, especially where it affects human rights. Transparency could be advanced through better auditing processes that show how the Australian Government is procuring and using AI.
2. There is a strong case for ongoing human oversight and evaluation of this form of government decision making.
3. The rules and practices for government procurement could be used to ensure that government AI-informed decision-making systems include strong human rights safeguards.

To support these preliminary conclusions, the Commission makes the following proposals.

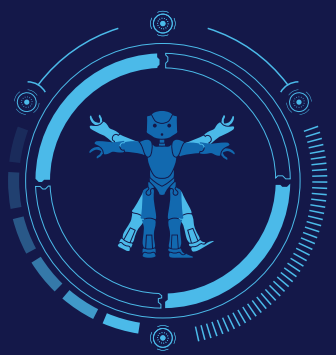
PROPOSAL 17:

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

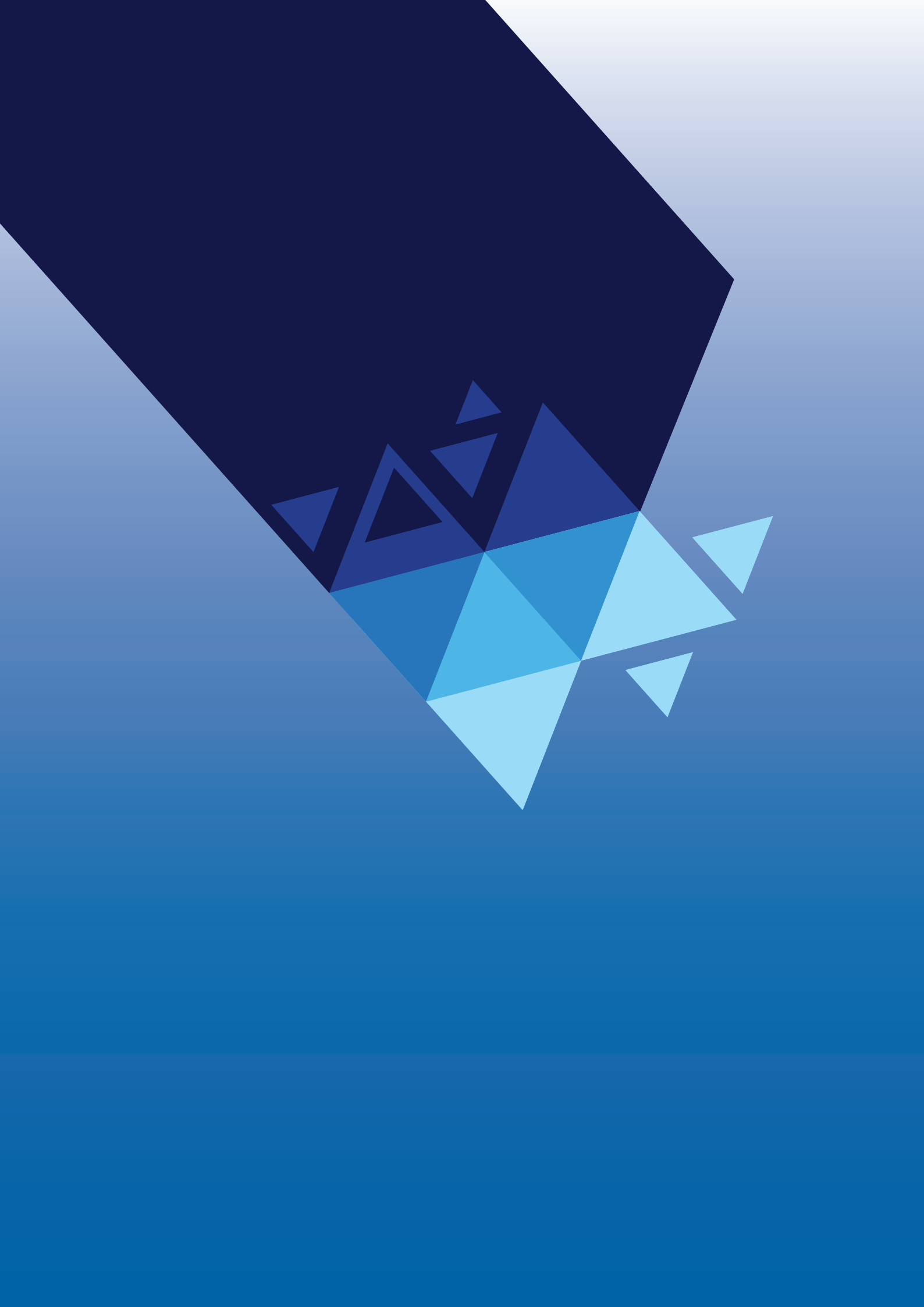
- **(a) identify the use of AI in decision making by the Australian Government**
- **(b) undertake a cost-benefit analysis of the use of AI, with specific reference to the protection of human rights and ensuring accountability**
- **(c) outline the process by which the Australian Government decides to 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.**

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.



'Procurement rules should require that any decision-making system that uses AI include adequate human rights protections.'



PART C: NATIONAL LEADERSHIP ON AI

CHAPTER 8
National leadership on AI

8.1 Introduction

Governance and leadership are critical to Australia's approach to a world that is increasingly driven by AI. Good governance and principled leadership will be central to how we advance the values that are important to us as a country, including the protection and promotion of human rights.

Through its Issues Paper, and a White Paper co-authored with the World Economic Forum, the Commission asked whether Australia needs an organisation to provide leadership on AI governance and, if so, what it should look like.[736] The Commission has consulted extensively on this issue, including by holding a symposium in early March 2019 with 65 leaders from the community, government, industry and academia.

The Commission proposes that a new AI Safety Commissioner role should be established to take a leadership role in AI governance.[737] This proposed independent statutory office, which is also referred to in this chapter as *expert leadership* or an *expert body,* would have a central role in policy development and national strategy relating to AI, and help build the capacity of existing regulators, government and industry bodies to respond to the rise of AI. The AI Safety Commissioner's focus would be on the prevention of individual and community harm, and the promotion and protection of human rights.

8.2 Is there a need for leadership?

(a) A new expert body

Stakeholders from academia, government, industry and the community broadly agreed that Australia needs expert leadership and strong governance for the design, development and use of AI.

But to what end? AI can support economic development, innovation in the development of products and services and many other aims. In pursuing such legitimate aims, our community also must be protected against harm that might arise through the development and use of AI.

In this context, stakeholders generally supported a governance approach to AI that emphasised human rights and agreed ethical principles.[738] In particular, there was support for an expert body that could provide leadership, by assisting and advising the public and private sectors.[739]

There was a sense of urgency regarding the need to establish an expert body to provide proactive policy and legal guidance on rapidly evolving technologies.[740] Such an expert body could anticipate and plan for future technological developments,[741] enabling Australia 'to define the direction of AI innovation, growth, sustainability, and future social impact'.[742]

One technology company described such a body as:

> acting as an interface between existing regulators, engaging with government, academia and industry, both domestically and internationally, identifying good practices and working to promote and implement them. This would necessarily be a largely educative and policy-based undertaking (as opposed a more traditional regulatory function) but the anchoring of such work in the promotion of human rights provides a solid foundation for what might otherwise be a nebulous mandate.[743]

Two common themes expressed in consultation were that an expert body should:

- focus on better coordination across regulators and not duplicate existing regulatory functions,[744] and
- be an independent government-funded or affiliated body working with industry, academia, regulators and other stakeholders, in particular the public, with special responsibility for vulnerable or at-risk people.[745]

(b) Existing governance arrangements

It has been argued that existing governance arrangements do not, or perhaps cannot, respond to the challenges posed by the fast-paced AI industry, and so new approaches to governance are needed.[746] In Chapters 6 and 7, this Discussion Paper considers possible legislative and other reform. In addition, an expert body could play an important coordination role. It could

> act as an 'issue manager' for one specific, rapidly emerging technology, as an information clearinghouse, an early warning system, an instrument of analysis and monitoring, an international best-practice evaluator, and as an independent and trusted 'go-to' source for ethicists, media, scientists and interested stakeholders. The influence of a [governance coordination committee] in meeting the critical need for a central coordinating entity will depend on its ability to establish itself as an honest broker that is respected by all relevant stakeholders.[747]

Some stakeholders submitted that the current mix of government regulatory bodies, and existing self- or co-regulatory processes established by industry, provide sufficient national guidance and regulation for the design, development and use of AI.[748]

Some of these stakeholders noted that more effective collaboration is needed to integrate the work of these bodies,[749] and that a further analysis of the existing regulatory framework might demonstrate the case for a new kind of organisation to govern the use of AI.[750]

(c) Economic and social benefits of a new expert body

The White Paper asked about the economic and social value of a new body. Stakeholders highlighted that such an entity could help protect human rights as well as promote national innovation and investment in AI. An expert body could advance the regulatory aims of ensuring that AI benefits society in general,[751] society is protected from harms associated with the use of AI,[752] and AI is used to enable human flourishing.[753]

Access Now saw this as an opportunity for Australia to define the direction of AI innovation, sustainability and future social impact:

> The ultimate goal should be AI that benefits humanity by contributing to a more responsible and equitable society. By inviting multi-stakeholder input, considering risks, and setting rights-respecting standards, Australia has the opportunity to be among the norm-setters on the world stage.[754]

Some stakeholders suggested that a new body could strengthen the national innovation economy and support local industry to participate in it.[755] This could help create a level playing field for technology companies through reduced barriers to entering a field at risk of monopolistic behaviours.[756] Greater transparency in the use of AI could assist in policing anti-competitive behaviour, with a view to promoting a fairer market.[757]

A new body could help highlight Australia's AI innovation globally, which could enhance our reputation,[758] based on rights-respecting standards[759] and quality-tested products for international distribution.[760] Some stakeholders noted that developing Australia's reputation for AI that protects human rights and the commercial rights of inventors could help build the nation's competitive edge.[761]

An enhanced international reputation for Australia would support Australia's advocacy for effective protections globally.[762] An AI developer said that a new body could enhance the operation of Australia's regulatory system in this area, which would have additional benefits:

> Australia may not be able to compete with other nations in terms of financial investment, but we can compete in terms of providing market leading AI that is held accountable in terms of ethics and transparency. By holding our own to account we can foster a world class reputation for the ecosystems we develop.[763]

Industry stakeholders also noted that a new expert body could help reduce commercial risk through established accreditation, compliance and best practice schemes, leading to business growth.[764] These in turn would enhance public trust in the use of AI, normalising its use in market settings, thereby increasing its economic potential.[765]

Other potential benefits of accreditation include the earlier identification, and prevention, of AI systems that produce poor social outcomes,[766] especially for vulnerable groups.[767] A new expert body could help foster a better informed public and facilitate public participation in the decision-making and regulatory processes concerning the risks and rewards of AI use in society.[768]

(d) Value to organisations

The White Paper invited stakeholders to comment on the value to their organisation of a new expert body.

Stakeholders in the technology industry saw value in a new body developing guidance on AI, through best practice advice and standards, to support business growth. Increased clarity from authoritative guidelines, standards and certification schemes could provide certainty for the technology industry, aid long-term strategic decision making, reduce legal and commercial risks facing the developing sector, and provide a competitive advantage.[769]

Industry stakeholders noted the value of guidance on avoiding algorithmic bias and minimising harm, thereby strengthening the local industry and encouraging global competition.[770] Portable submitted that a new body could add value by

> [e]nsuring risks are mitigated without impeding technical development and by enhancing public perception of AI technology.
>
> ...
>
> Having clear standards and auditing processes would help us provide public benefits by reducing misperception and fear of new technology. We would be able to provide increased clarity around how we develop AI tools by referencing authoritative guidelines and standards.[771]

Such activity could help build public trust and confidence in AI technology. Industry stakeholders saw value in being able to state publicly that they follow human rights guidelines set by an expert body,[772] resulting in better outcomes for the community through human rights compliant design and use of AI, and reduced misperceptions about these technologies.[773]

Some academic experts noted that a new body could support the work of university ethics committees, and contribute to education curricula and research activities that inform public policy development in this area.[774] Swinburne Social Innovation Research Institute stated:

> The [new body] could codify ethical principles around use of data and AI. At present university ethical committees are each, individually, grappling with these issues (guided by NHMRC and other codes) and there are quite different standards applied
>
> ...
>
> It would be useful to have a codification of principles to help non experts on ethical committees with this type of research.[775]

Academics doing research in the area of AI and health noted the usefulness of having recognised national leadership to guide policy development.[776] A new body could also inform and guide international collaborative research activities.[777]

The Royal Australian and New Zealand College of Radiologists anticipated value to the health sector through leadership and development of ethical approaches to the use of AI in medicine, a robust regulatory framework to protect the public, policy development, complaints resolution, and advising health regulators in relation to AI adoption and use.[778]

(e) Business case for the proposed expert body

The White Paper invited comment on the business case for a new organisation to promote responsible innovation in AI. Stakeholders generally focused on the significance of positive social impact and prevention of harm through the operation of such a new body, rather than economic benefits.[779]

It was also suggested that a business case could focus on the risk matrix of benefits and harms,[780] or the consideration of the risks if Australia does not create an expert advisory body.[781]

The Australian Human Rights Institute submitted that the business case could be supported by the new body taking responsibility for developing a code of conduct and training modules related to responsible innovation, as well as surveying industry about the introduction of AI and machine learning.[782]

Stakeholders observed that it would be difficult to evaluate the success of a new body by reference to traditional measures. Some suggested its success could be assessed, at least in part, by reference to social outcomes associated with the use of AI. It might be possible to draw inferences about its impact by considering the number, type and impact of decisions made using AI,[783] or the number of complaints associated with the use of AI.[784]

(f) International leadership

It has been observed that some countries are better positioned than others to benefit, and limit harm, from disruptive technology. Governments that invest in high-quality education, life-long skills training and upgrading, and a flexible and robust social safety net tend to benefit most in the long term.[785]

Eleonore Pauwels of the UN University Centre for Policy Research argues the convergence of AI with other powerful dual-use technologies is leading to an emerging new geopolitics of inequality, vulnerability and potential human suffering. According to Pauwels, the international community needs a shared responsible innovation and preventative approach to address the resultant large-scale risks.[786]

Against this backdrop, numerous strategies are being developed to address AI and other disruptive technologies. Those strategies vary in their focus; in their geographical reach (international, national and local); and in their origin (public, private and cross-sectoral). Some of these strategies are referred to throughout this Paper, with some main points highlighted below.

A recent paper analysed 18 national and regional AI strategies across the globe at the end of 2018.[787] It found eight major public policy themes across the strategies:

- scientific research
- AI talent development
- skills and the future of work
- industrialisation of AI technologies
- ethical AI standards
- data and digital infrastructure
- AI in the government
- inclusion and social wellbeing.

Fifteen of the strategies dealt with AI leadership by proposing a council, committee or task force to develop standards or regulations for the ethical use of AI.[788]

There are examples of initiatives and proposals of this kind at supranational and national levels. Pauwels proposed that the UN Global Foresight Observatory for AI Convergence could bring together a UN strategic foresight team with public and private stakeholders. The foresight capacity would assist the UN to guide strategy development and engage stakeholders in understanding the convergence of technologies and to develop responsible innovation and preventative approaches.

The Observatory could partner with private and academic research in order for the UN to understand and anticipate the implications of AI, and work collaboratively to prevent harm.[789] Some private sector actors have expressed interest in collaborating with the UN to 'foster normative guidance and align technological convergence with the public interest'.[790]

The OECD is launching an online AI Policy Observatory in early 2020 for AI information, evidence and policy options to 'nurture and monitor the responsible development of trustworthy AI systems for the benefit of society'.[791] The AI Policy Observatory will be multi-disciplinary, based on evidence and built on global multi-stakeholder partnerships. It aims to facilitate cooperation across OECD countries on policy coherence and help governments develop, implement and improve AI policies.

At a national level, the UK Centre for Data Ethics and Innovation is an advisory body set up by the UK Government and led by an independent board of experts to investigate and advise on how to maximise the benefits of data-enabled technologies, including AI. The UK Centre has a cross-sector remit and focuses on both innovation and ethics. It has an executive team drawn from both civil society and the broader policy-making sphere, academia, the tech industry, and government.[792]

The UK Centre makes recommendations to government, and produces codes of practice and guidance for industry and government. It also provides expert advice to regulators but it does not have national regulatory and governance functions. The UK Government is bound to consider recommendations made by the body in fulfilling its task to maximise the benefits of data and AI for society and the community, and plans to give the Centre independent statutory footing.[793]

8.3 The Commission's preliminary view

The Commission's consultation to date has revealed support across multiple sectors for expert leadership to take a critical role in coordination, capacity building and in developing regulatory and other policy regarding the design, development and use of AI.

Other reform processes have reached a similar conclusion. For example, ACOLA found that an independent body is needed to provide a critical mass of skills and institutional leadership to develop AI, promote engagement with international initiatives, and develop appropriate ethical frameworks.[794]

The body should bring together stakeholders from government, academia and the public and private sectors, and provide an opportunity for Australia 'to compete on the international stage, become international role models and provide trusted environments for AI development'.[795] The establishment of this body could occur as part of developing the proposed *National Strategy on New and Emerging Technologies* (Proposal 1).

The Commission proposes that this be achieved by establishing a new AI Safety Commissioner to provide leadership on AI governance in Australia. It is worth emphasising that this is a preliminary view, and the Commission is open to other ways of achieving this goal.

There are three critical factors that point to a need for leadership in Australia on AI, which the proposed AI Safety Commissioner could address.

1. **Expertise.** An increasing number of organisations are involved in developing and using AI for an increasing range of purposes. There is persuasive evidence that regulators, oversight bodies and those responsible for developing law and policy need support in better understanding how AI is relevant to their areas of work. The proposed AI Safety Commissioner could provide an authoritative source of expertise, advising on how to advance the public interest, and especially to protect human rights, in the vast array of areas affected by AI.
2. **Trust.** The benefits of AI are possible only if the community can trust that the risks and threats associated with AI are identified and adequately addressed. This factor was recognised in the Australian Government's recently released AI Roadmap.[796] By improving the operation of existing governance mechanisms, the proposed AI Safety Commissioner could play a critical role in building an enduring foundation of community trust.
3. **Economic opportunity.** There is an inherent good in enhancing Australia's governance to prevent harm and promote human rights. However, the Commission also finds persuasive the argument that the local innovation economy will be strengthened, and the global competitiveness of Australia's technology industry will be enhanced, if Australia comes to be known for developing and using AI-powered goods and services that protect users' basic human rights. The proposed AI Safety Commissioner could be pivotal in building Australia's reputation in this area.

The Commission acknowledges that the relative strength of stakeholder support for the proposed AI Safety Commissioner may turn on its precise role and functions. The Commission proposes the outline for an AI Safety Commissioner, with the remainder of this chapter focused on this body's aims, powers, functions, structure and operation. The next phase of the Commission's consultation will seek more specific input on these issues.

8.4 What should the AI Safety Commissioner look like?

(a) The basic model and focus

Two basic models emerged in the Commission's consultation.

The first model is essentially a traditional government regulator—to regulate the use of AI in Australia and carry out governance functions such as capacity building and policy development.[797] Stakeholders supporting this model suggested that the expert body could have a role in the direct regulation of the use of AI, as well as advise existing regulators in their oversight of the development and use of AI.

This model assumes a role in enforcement. Some stakeholders suggested it be able to issue penalty notices and apply for orders or injunctions; require changes to AI systems to comply with agreed standards; provide remedies for aggrieved parties; and decide what information should be made publicly available and kept commercially sensitive when legitimate human rights concerns arise with the use of AI.[798]

The second model was for an independent advice and policy centre, coordinating and building capacity among current regulators and others.[799] Under this model, any new regulatory powers relating to the use of AI would be given to existing regulators, not to the proposed new body.[800]

The University of Melbourne submitted that a specialist advisory organisation could coordinate regulators and other key players.[801] It could tap into the work on innovation, ethics and regulation being undertaken by industry, think tanks, universities, international organisations, and civil society groups globally, without duplicating efforts.[802] Other remits were suggested, such as oversight of a code of conduct,[803] or coordination and education of regulators.[804]

Stakeholders clearly envisaged that the expert leadership should focus its work on the development and use of AI. However, some caution is needed here. The Centre for Policy Futures drew attention to the blurring of boundaries across scientific fields, and between digital and non-digital technologies, such as the combination of genetic technologies and synthetic biology.[805]

(b) Core purposes

The most commonly assumed core purposes of the proposed expert body raised by stakeholders are the *protection of human rights and promotion of innovation*, or both. However, a number of stakeholders pointed to an inherent conflict or tension between protecting human rights and promoting innovation.[806] For example:

> There is a big risk in having one body that simultaneously exists to both promote innovation and protect human rights; its mere existence could become a reason for not acting on key issues. The human rights mandate needs to sit separately to the promotion of innovation.[807]

Other stakeholders did not see an irreconcilable conflict between these aims. For instance, a senior government representative suggested it would be normal for a single government body to promote both human rights and innovation.[808]

Several stakeholders suggested that protecting human rights should be an explicit aim of the new body.[809] The Office of the eSafety Commissioner stated that the positive and negative impacts of new technology are not experienced equally by all parts of the Australian community.[810] Therefore, any new body should focus on protecting the rights of groups who can be at particular risk from the use of AI, such as people with disability.[811]

Stakeholders raised other possible core purposes. For example, it was suggested that the proposed expert leadership should promote AI being used for the *public good*. This could include a focus on human flourishing and social well-being, so that AI produces positive social outcomes for individuals and the community at large.[812]

In this regard, some stakeholders referred to the protection of economic, social and cultural rights. The Australian Council of Trade Unions submitted that Australia needs an organisation to oversee the impact that AI and automation will have on our society, including the significant impact on the right to work.[813] The Australian Services Union stated that 'just transition' principles should be at the 'centre of government regulation to hopefully prevent the unintended consequences of AI whilst accentuating its benefits to workers and society'.[814]

Thirdly, there seemed agreement that the proposed expert body should have a core purpose of *unifying the various existing and proposed national AI initiatives*. It could lead in coordinating, standardising and improving self- and co-regulatory initiatives regarding the design, development and use of AI.[815] Those initiatives are discussed in greater detail in Chapter 7. Good examples of such leadership overseas were said to be the UK Centre for Data Ethics and Innovation and the US-based Partnership on AI.[816]

(c) Powers and functions

This section considers the powers and functions that could be given to the proposed expert body.

(i) Supporting co- and self-regulation

Many stakeholders supported a new expert body fostering good-practice regulation that applies to the development and use of AI. The proposed body could be integral in developing 'innovative forms of regulatory responses to complex, rapidly changing technologies',[817] such as the regulatory sandboxes discussed in Chapter 7.

While there was disagreement among stakeholders on whether the proposed expert body should itself be a regulator, there was broader support for this body overseeing and coordinating relevant co- and self-regulatory mechanisms, such as industry standards and self-assessment schemes. It could play a role in developing and administering:

- industry standards for AI[818]
- best practice frameworks, tools and methodologies[819]
- tools to predict and monitor the impact of the use of AI in particular contexts, with a particular emphasis on human rights.[820]

In Chapter 7, the Commission considers the utility of certification schemes for products and services that use AI. It could be that the proposed new expert body comes to have a role in running or overseeing such a scheme.[821]

(ii) Monitoring and investigating the use of AI

Some stakeholders suggested that the proposed expert body be given the power to monitor and investigate the use of AI in specific cases,[822] or systemic or industry-wide issues.[823]

The Consumer Policy Research Centre submitted that such powers should include the ability to request algorithmic design details: the outcome to be predicted; the inputs that are used by the algorithm; and the training procedure used.[824] Inquiries into systemic concerns could include: emerging new-generation AI products and capabilities; effects of AI systems on the broader population; and social inequality and discrimination impacting particular population cohorts.[825]

Some stakeholders suggested the new body could focus on particular sectors. For example, two stakeholders submitted it could cover the use of AI in healthcare. It could prioritise cross-sectoral collaboration with regulators and standard setting bodies, especially healthcare regulators, and provide advice and support for the research, adoption and deployment of AI and ML tools in health.[826]

(iii) Building capacity: education, training and guidance

A common theme across submissions was the importance of the proposed new expert body building the capacity of all sectors in the responsible design, development and use of AI. A number of options have been canvassed.

First, it could provide general guidance materials for the public and private sectors regarding how to comply with human rights and other such requirements in respect of AI.[827]

Secondly, it could offer more specific expert advice. For example, it could advise an organisation that is testing the potential impacts of an AI-powered product, or it could give advice on complying with industry standards.[828]

Thirdly, the expert body could provide education and training for government and private sector bodies involved in designing and using AI. Some stakeholders saw this as contributing to a cultural shift in the technology industry. For example, one said:

> There is an important role for the [organisation] of education and outreach to the technical, scientific and engineering communities in order to shift their worldview from 'Positivism' to 'Participatory' and human flourishing, so that they can understand and incorporate appropriate human values and ethics as 'goodness and quality criteria'.[829]

Training could be targeted to particular bodies or sectors that most need this assistance, such as building knowledge and skills regarding AI among existing government regulators.[830]

Fourthly, the expert body could be involved in public education and awareness raising for the community at large,[831] with particular attention to vulnerable groups.[832] This might be through the body's own activities, or by advising others, including by integrating information about AI in primary, secondary and tertiary education.[833] The University of Technology Sydney submitted:

> Education is important in part to ensure that the public understands the issues, but also equally importantly to create a collective creative environment in which the public can participate in the creation of new visions for Australia's future.
>
> This requires the participatory co-design and co-evaluation of technology whereby an educated population—and not just technology assessors—is able to set goals, set pertinent questions, evaluate answers, and influence the shape of technology.[834]

Fifthly, the expert body could promote collaboration on the responsible design and use of AI.[835] Some stakeholders noted the importance of building public-private collaborations so that government and industry can communicate positively throughout the introduction of any new regulation and help moderate any socio-political and economic tensions.[836] In this respect, one industry stakeholder noted:

> AI is constantly evolving and improving, and there is a role in Public-Private partnerships. By leveraging these partnerships we can share learning between industries concerning the development of AI and further promote diversity and inclusion.[837]

The proposed body also could facilitate international relationships. This could include collaborating with other similar entities for knowledge sharing, coordinating across governments to manage the potential global risks of AI, and collaborating with the international community on the societal implications of AI and emerging technologies.[838]

(iv) Policy development, research and advocacy

Stakeholders emphasised the need for leadership in policy development, research and advocacy to promote a consistent agenda for responsible AI across sectors. The expert body could be well placed to advise government on policy and regulatory issues regarding AI, with a view to promoting human rights without stifling innovation.[839] The Law Council of Australia suggested a governance body could be

> 'translational' to anticipate and articulate issues in a way that empowers policy makers, civil society and regulators to engage with issues that arise from data driven decision-making and technological innovation and do so in a way that is principles based and technologically neutral.[840]

The expert body could be well placed to monitor national and international research to help promote best practice.[841] Its research arm could support industry in the practical testing and examination of technical issues as they arise,[842] promote the training and development of local AI talent,[843] and inform a long-term AI research roadmap for Australia.[844]

(d) Independence and resourcing

Many stakeholders emphasised the expert body should be independent, in the sense of being free from actual or perceived bias, avoiding conflicts of interest, and being separate from both government and industry in order to maintain the confidence and trust of all sectors.[845] The clear message from stakeholders was that public funding should be the primary resourcing method.[846]

Some stakeholders noted the importance of maintaining independence through public funding, to avoid the perception or reality of regulatory capture.[847] LexisNexis submitted:

> The fostering and articulation of social norms and industry best-practice should be undertaken by an organisation which is commercially disinterested but able to act as a forum for businesses, academics and civil society to engage with issues as they arise.[848]

While acknowledging that public funding should be the primary source of income for the new expert body, some stakeholders noted there could be a need for additional funds from other sources. Financial support from industry could augment public funding,[849] though any such arrangement would need to be monitored to prevent any conflict of interest for the body.[850]

Nichols noted that industry and other sectors may wish to financially support the new body to demonstrate their commitment to the pursuit of responsible innovation.[851] Others suggested that the body could receive funding via a taxation mechanism or a fee levied on industry for certification,[852] or the provision of educational services.[853]

(e) Structure and expertise

Stakeholders were invited to comment on the structure of a new expert body and the expertise it would require.

(i) Leading the organisation

Stakeholders generally urged that the new body be led by a commissioner or chief executive, with support from specialist commissioners, possibly seconded from relevant regulatory bodies and expert disciplines.[854] One stakeholder suggested oversight by the Human Rights Commissioner,[855] and another suggested an internal 'chief ethical and humane use officer'.[856] The Centre for Policy Futures submitted:

> We envisage a relatively small non-hierarchical organisation, a Director, with short-term input and leadership from specialist part-time Commissioners relating to specific projects or programs, international expert advisors and a Board reflective of leading government, commercial and civil society stakeholders.[857]

Three stakeholders argued for diversity in the board and governance leadership of a body such as the proposed expert body, drawing on a range of technical and non-technical experience.[858]

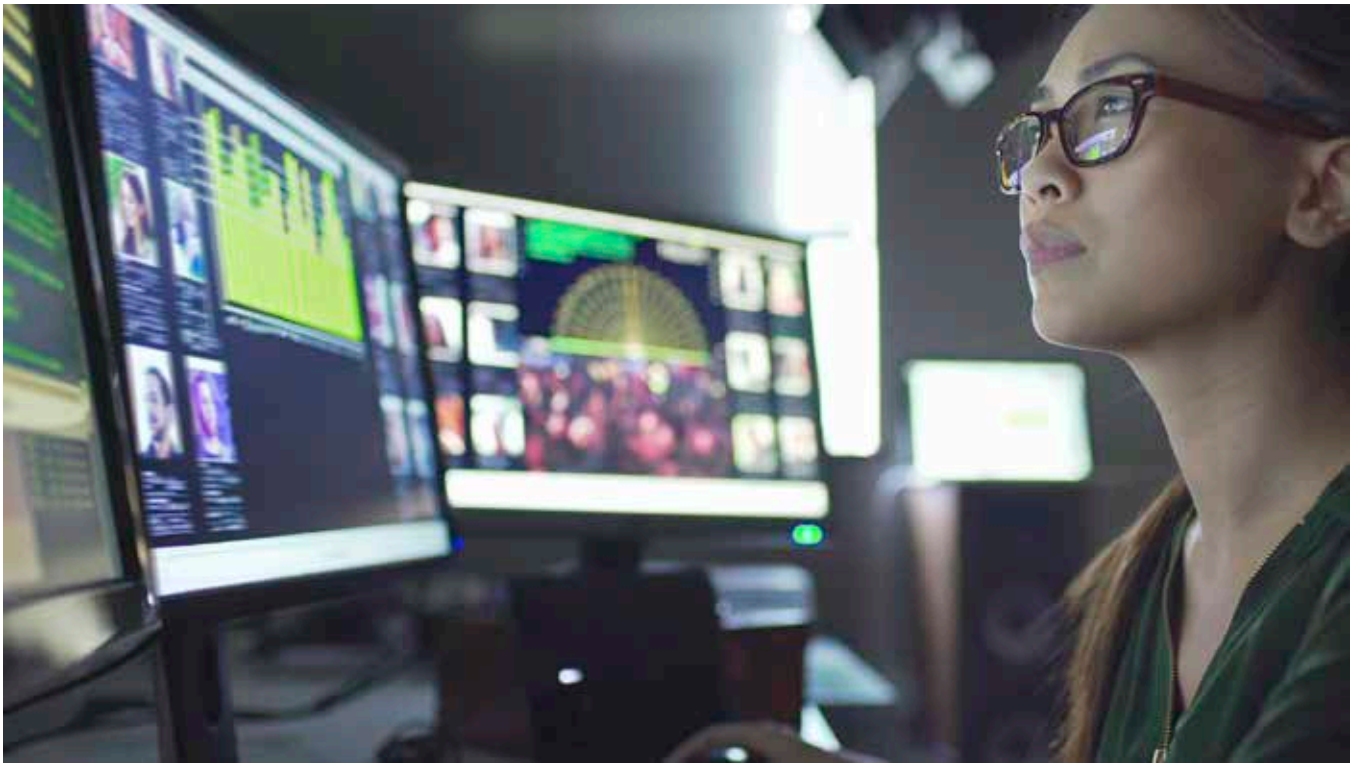
(ii) Expertise and representation

The overwhelming message from stakeholders was that the expert body must be interdisciplinary in its skills and expertise, and it should represent the community it serves.[859] Portable submitted that any governance body

> should foster the creation of multi-disciplinary regulatory teams who offer unique insights from academia, government, nonprofits, and the private sector in order to provide balanced, holistic, and forward-thinking analysis of AI trends and risks. [860]

According to stakeholders, the body should have expertise in the following fields:

- engineering, science and technology[861]
- the law and business[862]
- academia and ethics[863]
- social science and policy[864]
- the capacity for wide-ranging research and effective public communications.[865]



External expertise could be called on when required, including from: government agencies;[866] national human rights institutions;[867] social psychologists to help predict social impacts of AI;[868] and humanitarian and social service organisations which use AI, or are dealing with the consequences of the use of AI, in the community.[869]

There was overwhelming support for civil society representation and input in the new body's operations.[870] This could include advocacy for, and representation of, the public interest generally, as well as particular groups.[871] It could also include workers, representatives and union leaders who understand the impact of AI development on jobs.[872]

Stakeholders noted the importance of the new body establishing appropriate expert and advisory panels and committees. This could draw on industry, community and interest groups.[873] It could help surface interdisciplinary and culturally sensitive issues, foster legitimacy and trust with stakeholders, and ensure that AI is used consistently with community values.[874] The Montreal AI Ethics Institute submitted that the organisation

> should also consist of members from the public-at-large that would like to serve on the regulatory and technical committees within the [organisation]. Leveraging grassroots expertise will not only serve the function of being more inclusive but will also encourage the development of public competence and public engagement will increase the trust and acceptability of solutions coming from the [organisation].[875]

(f) Interaction with other bodies

The White Paper invited comment on how a new expert body should interact with other bodies with similar responsibilities. Stakeholders emphasised the importance of effective cross-sector collaboration and engagement.[876]

It was suggested that a collaborative multi-stakeholder approach was more likely to produce outputs that are fit for purpose and widely adopted by industry,[877] such as co- and self-regulatory approaches.[878]

Submissions focused on the desirability of any new expert body working alongside and coordinating with government, regulatory entities and other relevant bodies, to avoid duplication of activities. The expert body should be an independent source of policy to support or oversee the use of AI by government and the private sector.[879] In addition, it could facilitate international relationships, including with counterpart bodies in other jurisdictions, such as the UK Centre for Data Ethics.

(g) Monitoring and evaluation

Stakeholders suggested that generally accepted monitoring and evaluation benchmarks be adopted to demonstrate the value of the new expert body, including the tabling of annual reports in Parliament, reporting of accreditation activities, and ad hoc reports or studies into emerging technologies.[880]

Key performance indicators could cover the positive and negative impacts on individual and societal wellbeing, as well as the economic development of AI in Australia.[881] Monitoring and evaluation activities could also include metrics that could set the body apart as a leader in innovation.[882]

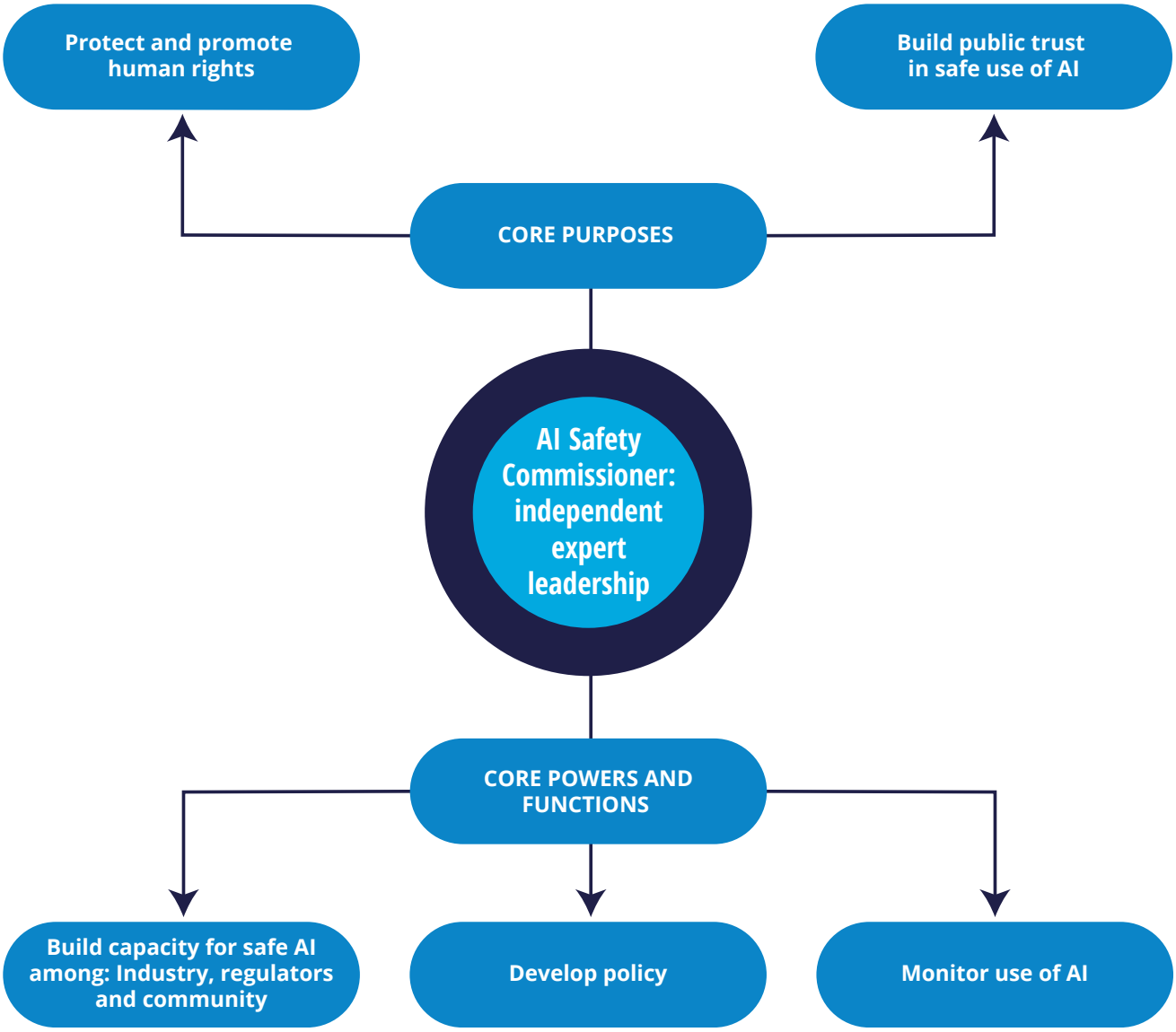
Transparency and accountability were common themes among submissions. Stakeholders submitted that the expert body should report on its activities in a transparent and accountable manner in order to engender trust and confidence in its operations.[883]

8.5 The Commission's preliminary view

(a) The basic model and focus

The Commission proposes the establishment of an AI Safety Commissioner as an independent statutory office, with the role of developing policy and coordinating and building capacity among current regulators and those involved in the oversight, development and use of AI. This is a model of collaborative leadership, in which the AI Safety Commissioner would share its expertise, while also coordinating and fostering cooperation across government and industry.

As an independent statutory office holder, there is some flexibility about where the AI Safety Commissioner is positioned, how it is resourced, and the possible expansion of its functions and powers. The Commission retains an open mind about whether the AI Safety Commissioner should be placed within an established regulatory or government body, or within a new 'Office of the AI Safety Commissioner'. Its proposed functions and powers are explored further below.



The Commission does not support the AI Safety Commissioner, or any other single body, becoming the sole regulator for AI, or its development and use. AI is a cluster of technologies that can be, and already is being, used in widely divergent ways in almost limitless contexts.[884] Some uses of AI engage human rights, while others do not. A focus on regulating AI *as technologies* would not be as effective as regulating AI in its context and use.

Instead, the Commission proposes that the AI Safety Commissioner *support* the existing regulatory structure, and build the capacity of regulators and others involved in protecting the rights of people who may be affected by the use of AI in various different settings.[885] The proposed AI Safety Commissioner could help bring existing regulators up to speed regarding the development and use of AI in their respective spheres of regulatory responsibility, rather than the Commissioner taking on their regulatory role.

(b) Core purposes

The protection of human rights and the promotion of innovation are each legitimate and important, but tension can arise between them.

Given that the Australian Government already promotes innovation in a number of domains, and given there are many organisations with a commercial interest in promoting innovation involving AI, there is significant unmet need in combating harm to humans and especially in protecting human rights.[886] While the Commission acknowledges the legitimacy of promoting the economic and other related benefits of innovation, the proposed AI Safety Commissioner should focus on protecting and promoting human rights.

The title 'AI Safety Commissioner' aims to connote both the protection and promotion of human rights, and the Commission invites feedback on this title.

This conception—the protection and promotion of human rights—of the proposed AI Safety Commissioner's role would not be at odds with the broader innovation agenda, which relies on building community trust in AI. The AI Safety Commissioner would contribute to public trust in AI by working to ensure that AI's risks and threats to the community are effectively identified, understood and addressed.

(c) Powers and functions

Four core powers and functions for the expert body have emerged in the consultation to date: supporting existing regulators; monitoring the use of AI; capacity building; and policy development. The Commission supports the AI Safety Commissioner being focused on those activities. The AI Safety Commissioner would draw on evidence and insights from across regulators, government, industry, academia and the public. The Commissioner would be responsible for developing human rights compliant policy and building the capacity of the various sectors in their understanding and use of AI.

Building capacity and policy development are essential in addressing emerging human rights concerns with AI-informed decision making. The proposed AI Safety Commissioner should focus on equipping the public and private bodies involved in the development and use of AI to ensure that they are acting in ways that comply with human rights and avoid harm.

While the AI Safety Commissioner should not be a regulator itself, it might have a useful role in overseeing or facilitating some of the mechanisms proposed in AI Chapters 6 and 7 in the future. The results arising from some of the reforms proposed in this Discussion Paper may necessitate increased powers and functions. For example, the AI Safety Commissioner could be a candidate to oversee a mooted 'AI trustmark' scheme, take on a role in respect of any new regulatory sandbox that focuses on AI, or be granted a statutory power to investigate AI-informed decisions as a 'technical expert'.[887]

(d) Independence and resourcing

The Commission proposes that the AI Safety Commissioner be a statutory appointment, similar to the eSafety Commissioner.[888] The AI Safety Commissioner should be an independent appointment, with core funding from government to help ensure independence and secure the public's trust. There could be opportunities for other income sources from industry or the community, with necessary protections from conflicts of interest or commercial pressures.

(e) Structure and expertise

The Commission proposes that the AI Safety Commissioner be resourced with interdisciplinary expertise across engineering, science and technology; law and business; civil society; academia and ethics; and social science and policy. The AI Safety Commissioner's office would develop strong professional relationships in government, industry, community and academia, with a ready access to external expertise.

Resourced with this cross-sectorial and multi-disciplinary expertise, the AI Safety Commissioner would be well placed to be at the forefront of technological advances nationally and internationally, and effectively lead and understand the perspectives and driving motivations of each key sector.

(f) Interaction with other bodies

A consistent message throughout this consultation has been the need for a collaborative and unified approach on AI governance. AI initiatives and strategies have emerged from the private and public sectors, at the national and international levels. It is essential that the proposed AI Safety Commissioner build on these existing partnerships and work.

This would necessarily include collaboration with Australian Government bodies that have responsibility for regulating the development or use of AI, industry bodies and civil society organisations, such as the ACCC, as well as bodies outside government such as Standards Australia. It would also include international counterpart bodies and initiatives like the UK Centre for Data Ethics.

Human rights compliant policies and practices are more likely to be adopted by industry, government and regulators when these sectors are included, or at least consulted, in the proposed AI Safety Commissioner's policy development and activities. A collaborative multi-stakeholder approach will improve outcomes for all sectors, as human rights compliant behaviour is more likely to be adopted by industry, government and regulators. It will also help promote policy unity across sectors and minimise the risk of duplication of activities.

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.**

‘Establish the role of AI Safety Commissioner to take a national leadership role in the development and use of AI in Australia.’

PART D: ACCESSIBLE TECHNOLOGY

CHAPTER 9
The right to access technology

9.1 Introduction

This chapter considers how people with disability experience and are affected by new and emerging technologies.

Technology is becoming the main gateway to participation across all elements of individual and community life. With rapid growth in the use of digital and more recent technologies in almost every aspect of life, access to education, government services, employment and other activities increasingly depend on the ability to access and use technology.

Those who experience technology as inaccessible can be excluded from everyday life. All members of the Australian community must have equal opportunity to participate, regardless of their disability, race, religion, gender or other characteristic.

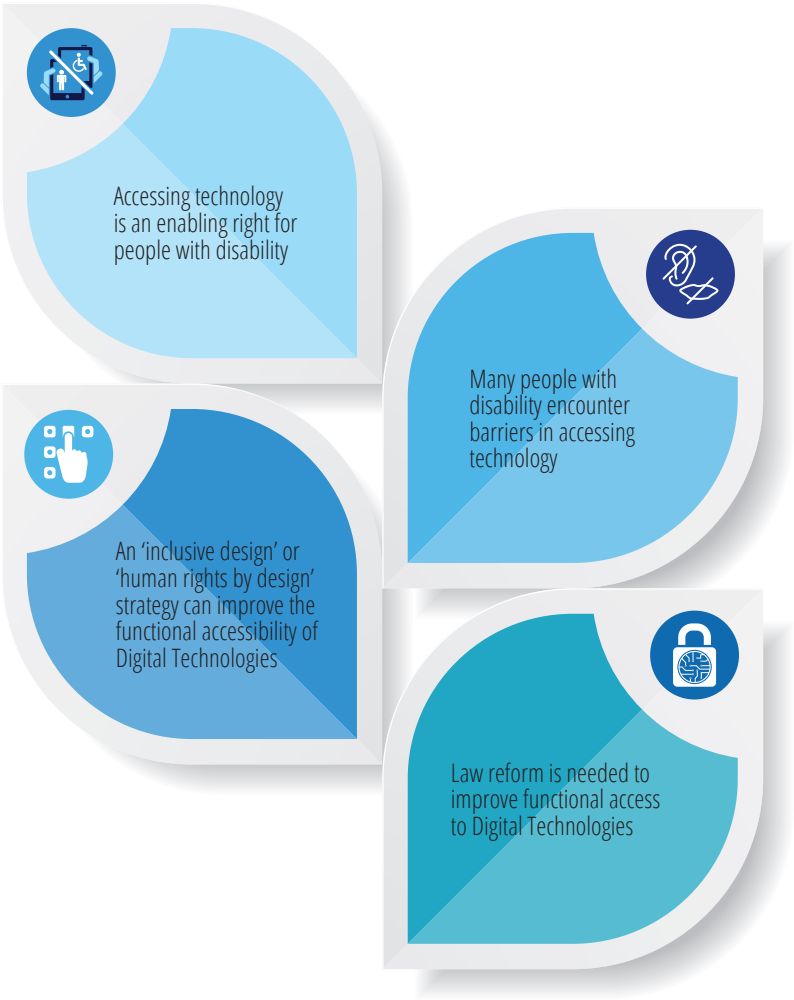
The Commission has consulted the community on how new and emerging technologies engage the human rights of people with disability, including disability advocacy and representative groups, and individuals. This uncovered many examples of opportunities and challenges arising from new technology for people with disability. The aim of Part D is to bring together some common themes that emerged throughout consultation, as well as identify some barriers that many people with disability encounter when accessing technology, noting that people with disability have individual and varied experiences with accessing technology.

New technology can enable the participation of people with disability like never before—from real-time live captioning to smart home assistants. Exclusion from technology, such as cashless payment systems that are inaccessible to people who are blind, may be discriminatory on its face and can also have detrimental effects on an individual's participation, independence and inclusion in all aspects of life.

The Commission particularly heard about the importance of accessibility for those new technologies that are most frequently used in common or important goods, services and public facilities.

Stakeholders focused on new technologies associated with information and communication, connected devices and the Internet of Things (IoT), and virtual and augmented realities. This Discussion Paper refers to these technologies collectively as 'Digital Technologies', and they are the focus of Chapters 9, 10 and 11.

Two types of 'access' are referred to throughout these three chapters—*obtaining* technology, and *using* technology (the latter is known as functional access).[889]



The Commission has formed four key preliminary conclusions:

1. Accessing Digital Technologies is an enabling right for people with disability.
2. Many people with disability encounter barriers in accessing Digital Technologies.
3. An 'inclusive design' or 'human rights by design' strategy can improve the functional accessibility of Digital Technologies.
4. Law reform is needed to improve functional access to Digital Technologies.

(a) International framework

All people have the right to equality and non-discrimination under the *International Covenant on Civil and Political Rights* (ICCPR) and the *International Covenant on Economic, Social and Cultural Rights*.[890] All the rights contained in the 'international bill of rights' apply equally to people with disability as to others. The *Convention on the Rights of Persons with Disabilities* (CRPD), to which Australia is a State Party, aims to 'promote, protect and ensure the full and equal enjoyment of all human rights and fundamental freedoms by all persons with disabilities, and to promote respect for their inherent dignity'.[891]

The following articles are particularly relevant to the accessibility of new and emerging technologies for people with disability.

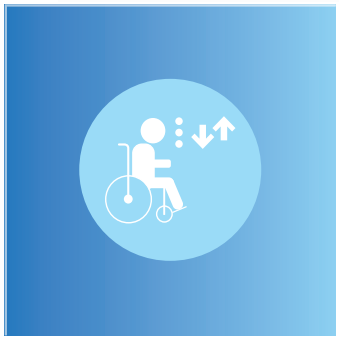
Box 7: Key provisions of CRPD:[892]

Convention principles Article 3

a) Inherent dignity, individual autonomy, freedom of choice and independence;
b) Non-discrimination;
c) Full and effective participation and inclusion in society;
d) Respect for difference and acceptance;
e) Equality of opportunity;
f) Accessibility;
g) Equality between men and women;
h) Respect for evolving capacities for children and their identities.

Accessibility Article 9

People with disability have the right to access all aspects of society on an equal basis with others including the physical environment, transportation, information and communications, and other facilities and services provided to the public.



Living independently and being included in the community Article 19

People with disability have the right to live independently in the community

Freedom of expression and opinion, and access to information Article 21

People with disability have the right to express themselves, including the freedom to give and receive information and ideas through all forms of communication, including through accessible formats and technologies, sign languages, Braille, augmentative and alternative communication, mass media and all other accessible means of communication.

The CRPD also provides for:

- **equality in education**, including taking appropriate measures facilitating and delivering education in the most appropriate modes and means of communication[893]
- **the highest attainable standard of health** without discrimination on the basis of disability[894]
- **equality in employment**, prohibiting discrimination in all forms of employment and promoting equal opportunities for work of equal value[895]
- **equality in political rights**, including by ensuring that voting procedures, facilities and materials are appropriate, accessible and easy to understand[896]
- **equality before the law**, which includes equal legal capacity and support to exercise that capacity[897]
- **the enjoyment of cultural life**, including access to cultural materials in accessible formats and to enjoy access to television programmes, films, theatre and other cultural activities in accessible formats.[898]

The CRPD also requires States Parties to:

- adopt all appropriate legislative, administrative and other measures for the implementation of the rights contained in the CRPD[899]
- collect appropriate information, including statistical and research data, to enable them to formulate and implement policies to give effect to obligations under the CRPD,[900] and
- implement national coordination strategies, such as ensuring there is a framework to promote, protect and monitor the implementation of the CRPD.[901]

(b) Accessibility in the CRPD

As noted above, art 9 of the CRPD sets out the right of people with disability to access all aspects of society equally, including in respect of 'information and communications technologies and systems' (ICT).

How art 9 applies to new and emerging technologies, including ICT and other goods, services and facilities available to the public, is a central focus of Chapters 9, 10 and 11.

States Parties must ensure that art 9 is fulfilled, which includes:

- taking measures to identify and eliminate barriers to accessibility
- implementing minimum accessibility standards
- promoting the design of accessible technologies and systems at an early stage.[902]

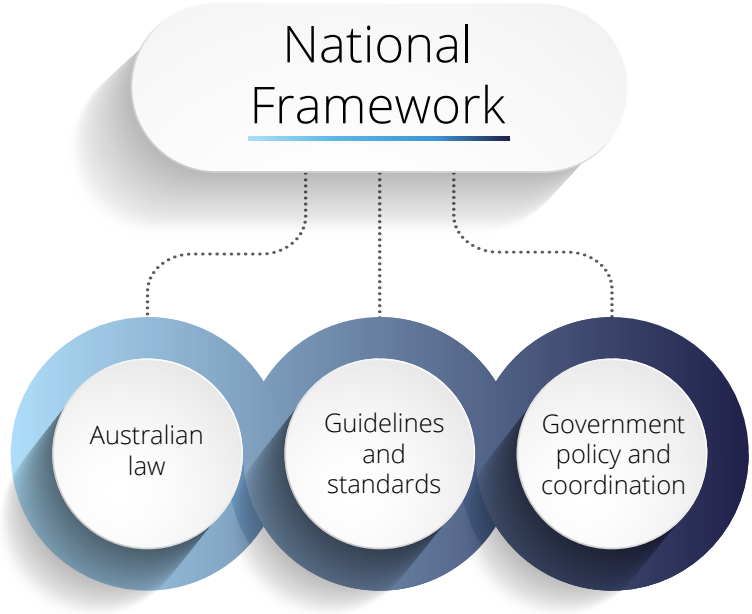
Accessibility is a key underlying principle of the CRPD,[903] because it is 'a vital precondition for the effective and equal enjoyment of civil, political, economic, social and cultural rights by persons with disabilities'.[904]

Accessibility under art 9 is therefore 'a disability-specific reaffirmation of the social aspect of the right of access'. Rights to access are protected also in a number of other human rights treaties. For instance, art 19(b) of the ICCPR protects the right of everyone to freedom of expression, regardless of frontiers and through 'any…media'. That requires that new technologies used for communication can be accessed by all.

The Committee on the Rights of Persons with Disabilities has referred to the right to access public service protected by art 25(c) of the ICCPR, and the right to access places and services protected by art 5(f) of the *International Convention on the Elimination of All Forms of Racial Discrimination*, in support of the view that there is a more general right of access reflected in the international human rights framework.[905]

The CRPD sets out durable principles that continue to apply as technology develops swiftly and often unpredictably.

As noted above, art 9 expressly refers to ICT, which includes any information and communication device or application and its content, and 'access technologies, such as radio, television, satellite, mobile phones, fixed lines, computers, network hardware and software'.[906]



The new technologies canvassed in this Discussion Paper include ICT, as well as other technologies referred to in the Commission's consultation process, such as the Internet of Things (IoT), virtual reality (VR) and augmented reality (AR).

The right of people to access ICT and other technological advances, which are available to the public, will be referred to throughout this chapter as the 'right of people with disability to access technology' or the 'right to access technology'. This CRPD-protected right is also related to the right to benefit from scientific progress.[907]

Australia must take appropriate measures to identify and eliminate obstacles and barriers and promote the design and development of accessible technology to fulfil the right of people with disability to access technology.[908]

(c) National framework

(i) Australian law

There are a number of federal, state and territory laws relevant to the right of people with disability to access technology.

Discrimination on the basis of disability is unlawful under federal, state and territory law.[909]

The *Disability Discrimination Act 1992* (Cth) (DDA) prohibits direct and indirect discrimination on the basis of disability in employment, education, accommodation and in the provision of goods, services and facilities. Direct discrimination involves a person with disability being treated less favourably than a person without that disability in similar circumstances.[910]

Indirect discrimination can arise where a person with disability must comply with a rule that unfairly disadvantages them because of their disability.[911] In some situations, the DDA requires reasonable adjustments be made to accommodate people with disability.[912]

For example, a person who is blind may require a screen reader to do their job. The DDA could require the person's employer to provide a screen reader as a reasonable adjustment, unless the employer shows this would cause unjustifiable hardship on the employer.

A person who has been discriminated against under the DDA can complain to the Commission,[913] which has the power to investigate and conciliate complaints. If a complaint is not successfully conciliated, the person can take their matter to the Federal Court or Federal Circuit Court,[914] which can determine the dispute and order remedies such as damages.

The DDA enables the Commonwealth Attorney-General to make Disability Standards that set out more detailed requirements about accessibility in a range of areas. There are currently Disability Standards in relation to education, buildings and public transport.[915]

Some Australian laws deal more specifically with disability and technology. For example, the *Broadcasting Services Act 1992* (Cth) (BSA) regulates Australia's television broadcasters and authorises the Australian Communications and Media Authority (ACMA) to monitor and regulate the broadcasting, datacasting, internet and related industries.[916] The BSA provides for industry codes and standards, including minimum requirements for broadcasters to caption television programs for people who are deaf or hearing impaired.

The *Telecommunications Act 1997* (Cth) authorises ACMA to make standards to regulate features in telecommunications that may be required by people with disability, including voluntary standards.[917] For example, one standard prescribes requirements for telephone headsets or keypads, and recommends design features that remove barriers to access for people with disability.[918]

(ii) Guidelines and standards

There is a growing body of guidelines and standards that can promote access to technology. Much of this is not legally binding and is known as 'soft law'.

Standards Australia is the nation's peak non-government standards organisation. It develops and adopts internationally-aligned standards in Australia and represents the nation at the International Organisation for Standardisation (ISO).[919]

Australian standards are voluntary and Standards Australia does not enforce or regulate standards. However, federal, state and territory parliaments can incorporate those standards into legislation and some have been incorporated in the Disability Standards referred to above.

Standards Australia adopted European Standard 301 549 in 2016, which deals with accessibility requirements suitable for public procurement of ICT products and services.[920]

This Australian Standard—AS EN 301 549—was adopted by the Australian Government in 2016.[921]

The Web Content Accessibility Guidelines (WCAG) aim to provide a single shared standard for web content accessibility for people with disability. The standard was developed by the World Wide Web Consortium Web Accessibility Initiative (W3C WAI) and the most recent iteration, WCAG 2.1, was released in June 2018.[922]

It provides guidance on the design of website content that is accessible to a wide range of people with disability including: blindness and low vision; deafness and hearing loss; learning disabilities; cognitive limitations; limited movement; speech disabilities; and photosensitivity.[923] The Australian Government encourages agencies to meet WCAG 2.1 in all communications.[924]

Other international guidelines and standards include the following:

- The International Telecommunication Union, a United Nations body, develops technical standards to improve ICT access for underserved communities.
- As noted above, the ISO creates standards that provide requirements, specifications, guidelines or characteristics that help ensure materials, products, processes and services are fit for their purpose.[925] In addition to standards that directly cover accessible products (eg, portable document format (or PDF) specifications that allow for greater accessibility), the ISO produces guides for addressing accessibility in standards.
- The International Electrotechnical Commission produces international standards for 'electrotechnology' products, systems and services.[926]

(iii) Government policy and coordination

The National Disability Strategy 2010–2020 (Disability Strategy) was established through the Council of Australian Governments (COAG). It incorporates principles from the CRPD into national policy and practice regarding disability, with a view to helping Australia meet its obligations under the CRPD.[927]

Consistent with the stated principles underlying the National Disability Insurance Scheme (NDIS),[928] the first 'outcome' of the Disability Strategy is that people with disability live in accessible and well-designed communities with opportunity for full inclusion in social, economic, sporting and cultural life.

Policy directions for this outcome are interrelated and include increased participation in, and accessibility of, communication and information systems for the social, cultural, religious, recreational and sporting life of the community. The strategy has associated implementation plans and disability access and inclusion plans for federal, state and territory governments.

The Digital Transformation Agency (DTA) administers the Digital Service Standard, which ensures Australian Government digital services are simple, clear and fast.[929] Criterion Nine of the Standard provides that services are to be accessible to all users, regardless of their ability and environment.

Government services are required to evidence usability testing of their digital platforms, including users with low-level digital skills, people with disability, and people from diverse cultural and linguistic backgrounds. The Australian Government's design content guide outlines digital accessibility and inclusivity considerations for a range of population groups.[930]

As noted above, the Commonwealth Procurement Rules incorporate AS EN 301 549 which covers accessible ICT. WCAG 2.1 was released after that Standard was adopted, which is encouraged by the Australian Government. This is an example of how the law, standards and government policy on accessible ICT interact.

9.2 Right to access technology as an enabling right

New and emerging technologies can enable the enjoyment of other human rights. Access can enable people with disability to enjoy equal participation and inclusion in political, social, economic and cultural aspects of life in the community.

Conversely, inaccessible technology can restrict human rights, excluding people with disability from community participation and limiting independence.

The Commission's consultation to date has highlighted that inaccessible technology is a pervasive problem, making it more difficult for people with disability to participate and be included in a range of activities, and live independently. When commenting on accessibility, stakeholders tended to focus on ICT, IoT, VR and AR technologies.

This chapter considers:

- the right of people with disability to access Digital Technologies as an enabling right
- the experiences of people with disability with regard to functional accessibility
- the experiences of people with disability obtaining Digital Technologies.

The Commission proposes ways in which the Australian Government and private sector should increase the provision of accessible Digital Technologies below.

(a) Functional accessibility—submissions and consultations

In the Commission's consultation process, the most frequently-cited concern was that denial of the right to access Digital Technologies can limit enjoyment of CRPD rights. Stakeholders also provided examples of how accessible and assistive technology can provide opportunities for people with disability for increased participation and inclusion in political, social, economic and cultural aspects of life.

(i) Inaccessible technology and impact on other human rights

Stakeholders from government, industry and academia all attested to the importance of accessible technology for people with disability enjoying the full range of human rights.[931]

People with disability and community and advocacy groups cited many examples of how an inaccessible technology can impinge on the rights of people with disability.[932]

When a person with disability is not able to access technology, their right to equal access may not be the only human right restricted or violated. The right to access technology on an equal basis with others is particularly important for the realisation of the CRPD's overarching aims, especially participation, inclusion, independence, autonomy and equality of opportunity.

Human rights are indivisible, interdependent and interrelated. The fulfilment of one right often depends, wholly or in part, on the fulfilment of others.[933] This principle is evident in stakeholder feedback on the significance of accessible technology as a precondition for people with disability to enjoy the right to live independently and participate fully in all aspects of life.[934]

For example, Digital Technologies are pervasive and indispensable in workplaces, education, health and as a means of communicating.

As a consequence, the right to access information and communications technologies and systems in art 9 of the CRPD is increasingly critical to people with disability enjoying the right to work (art 27), study (art 24), health (art 25) and to freedom of expression and opinion (art 21), on an equal basis with others. It is in this sense that the right to access technology can be considered an enabling right.

It can be compared with other enabling rights, such as the right to education, which enables the realisation of other economic, social and cultural rights. That is, access to technology helps build skills, capacity and confidence to help people achieve these rights.[935] The UN Committee on the Rights of Persons with Disabilities notes that accessibility is 'a vital precondition for the effective and equal enjoyment of civil, political, economic, social and cultural rights by persons with disabilities'.[936]

In the Commission's consultation process, people with disability reported that accessing technology was critical to their enjoyment of other human rights and having greater independence and participation in society. In this vein, the Australian Communications Consumer Action Network (ACCAN) stated:

> Access to technology can offer expanding opportunities to people with disability. With greater access to accessible technology comes greater inclusion within society and more equal enjoyment of human rights, including for instance more inclusive workplaces, better access to education (including lifelong learning), and greater participation in cultural life, in recreation activities, leisure and sport.[937]

Other stakeholders emphasised that access to technology was critical to enjoying the right to communicate with others, which is itself a precondition to the fulfilment of other rights.[938]

Stakeholders pointed to the impingement on several CRPD-protected rights where a person with disability is unable to access technology. In particular:

- Where the technologies required for work are inaccessible, this compromises the **right to work** (art 27).[939] This problem is exacerbated by the near ubiquity of Digital Technologies in the workplace.[940] For example, more and more job advertisements appear primarily or solely online, which disadvantages those who cannot access such websites.[941] Similarly, it was reported that ICT routinely used in employment settings—such as for content management and internal human resources and finance systems—is not accessible. This creates a barrier for people with disability working in those environments.[942]
- The **right to freedom of expression and opinion** (art 21) can be constrained by difficulties for people with disability accessing ICT hardware,[943] easy-to-read online information[944] and websites.[945]
- The **right to education** (art 24) is impacted when accessible educational materials are not provided.[946] Inaccessible technologies can have a significant effect on the education of children with disability.
- The **right to privacy** (art 22) is affected where technology tools collect personal information from people with disability.[947]

- **Political rights** (art 29) can be compromised when accessible ICT hardware is not provided at voting stations and information portals.[948]
- The **rights to equality before the law** (art 5), **equal legal capacity** (art 12) and **effective access to justice** on an equal basis with others (art 13) can all be negatively affected by the use of predictive data,[949] especially where it leads to algorithmic bias affecting groups such as people with disability.[950] Bias and discrimination in automated decision making are considered in detail at Chapter 6.

(ii) Accessible technologies as enabling tools

Many illustrations of the positive impact of accessible and assistive technologies for people with disability were received by the Commission.

However, it was also clear that the same item of technology often can act as both an enabler and a barrier for people with disability, depending on the context. For example, a smart home assistant may support a person with a physical disability with independence at home but be inaccessible for a person with a cognitive or communication disability and act as a barrier to independence.[951]

An accessible or assistive item of technology can unlock opportunities for the realisation of human rights and support CRPD principles of independence, autonomy, and participation and inclusion in society.

Examples include voice and virtual assistants, text-to-speech and text-to-sign-language applications.[952] Complex technologies—including mouth-controlled joysticks, head-controlled pointing devices and eye-gaze technology—can similarly support the realisation of human rights of people with disability.[953]

National Disability Services summarised the benefits for people with disability:

> Assistive technology is of great benefit to many people with disability. It can boost personal independence; improve quality of life; assist with social inclusion; and reduce the need for personal supports.[954]

Stakeholders noted the opportunities afforded by the Internet to share information and make it more easily available.[955] It can support the right to freedom of expression and opinion, access to information, access to employment opportunities, social participation and services.[956]

Family and friends who care for people with disability likewise benefit from these technologies. As independence increases for the person with disability, carers can pursue employment and education opportunities.[957]

Several stakeholders pointed to advances in health care technology. For instance, machine learning could be used to predict more accurately future health events, outcomes and treatments for people with disability.[958]

The disability sector agreed that technology can improve independence and autonomy,[959] and increase participation in every aspect of life.[960] The Public Interest Advocacy Centre and the Australian Rehabilitation and Assistive Technology Association provided case studies illustrating increased independence and participation through the use of accessible technology.

Case study:

Voting in NSW[961]

Work is underway to have all state and territory electoral commissions adopt a single form of electronic voting based on a telephone keypad. A system has been in operation in NSW since 2011 through the use of iVote. The system has allowed blind and vision-impaired people, as well as other voters with a disability and those living in remote areas, to cast a secret and unassisted vote remotely using an interactive voice recognition-based phone number or an internet-enabled computer. Once lodged, iVotes are printed out in a central location as completed ballot papers and included in the manual count processes.

Electronic assisted voting has greatly improved the franchise of people with disability, with many electors who are blind or have low vision responding positively to the use of electronic voting machines.

Case study:

Automated home[962]

I am a C3 quadriplegic so I automated the entire house. Lights (including dimmer settings), fans, air conditioner, elevator, air conditioning and audio visual equipment.

The AV equipment includes TV, digital video recorder (including navigating menus and recording shows), home network hard drives, Netflix, YouTube etc.

By automating I mean everything is controlled by voice through Google home and currently working on Amazon Alexa... Everything can also be controlled by my smart phone by tapping icons all by scrolling through the menus using a single switch scrolling method. We also tested my home using the Apple watch which worked really well. The differences automation makes to the lived experience is huge. Instead of having to call people I can just ask Google to do whatever I need. This is particularly important in the middle of the night. If I wake up feeling hot I can turn on the bedroom fan or air conditioner I can turn on and off the TV, movies or music if I can't sleep.

(iii) Accessing technology under the NDIS

Some stakeholders raised issues related to the NDIS. This Project is not a detailed review of the NDIS itself, and there are other review and evaluation processes that focus more deeply on the NDIS.[963] Nevertheless, issues raised by stakeholders are relevant to the NDIS and other government service provision.

Some stakeholders emphasised that access, through the NDIS, to new technology products and services could transform positively the lives of people with disability.[964] An example of this transformative effect is that of a young person with a non-verbal disability who was provided an iPad with speech processing through their NDIS package, and is using this technology to run their small business.[965]

The main concern raised by stakeholders is part of a larger debate about whether the NDIS should prioritise the provision of assistive or accessible technologies.

For example, many stakeholders reported that the National Disability Insurance Agency (NDIA) takes an inconsistent approach when determining whether to provide accessible 'mainstream' technology, such as smartphones, under NDIS packages.[966] It was claimed that some NDIS planning regions allow the purchase of accessible technology, such as a smart phone; others a lease; and others do not allow this at all.[967] This is said to have caused confusion and frustration across the disability sector.[968]

A related concern was that a critical factor that often determines whether goods and services that use Digital Technologies are provided under an NDIS plan is how effectively the individual can articulate their goals and needs, or whether a planner is able to identify them.[969]

(iv) Functional accessibility

A problem of 'functional accessibility' arises where an individual cannot use a particular item of Digital Technology (eg, a smartphone) because it does not accommodate the individual's disability.[970] When considering whether a digital good, service or facility is functionally accessible to the public, the question is usually whether the user interface (UI) of the product or service is designed in a way that includes or excludes people with disability.[971] Functional accessibility concerns the design and functionality of the Digital Technology, including software, hardware and any other items that are central to its operation.

Functionally accessible Digital Technology differs from 'assistive technology'. Assistive technology is specifically designed to support a person with a disability to perform a task. An example is a screen reader, which can assist a person who is blind, or who has a vision impairment, to read the content of a website.[972] Assistive technology is generally used only by people with disability.

As different disabilities bring different accessibility requirements,[973] functional accessibility is sometimes described as a match between the user's individual accessibility needs and the provider's UI.[974] Therefore, it needs to be considered at both the hardware and software levels in production and along the consumer chain.

For example, an Android or Apple smartphone may be accessible, but the applications procured through that smartphone may not be accessible. Industry stakeholders observed that the online stores where consumers obtain such applications play an important gatekeeping role regarding the basic requirements for these applications.[975]

Those gatekeepers could specify that application developers adhere to minimum standards on matters such as security and accessibility. Ultimately, however, application developers are primarily responsible for making applications accessible.[976]

When an accessibility problem relates to a Digital Technology that is used in many individual products and services, all of the resultant products and services can be rendered inaccessible.[977]

A good example of this is where a device requires a PIN to be entered by using a touchscreen (also known as 'PIN on Glass' or POG). Unless specially modified, such Digital Technology is generally inaccessible for people who are blind or have a vision impairment, because it is impossible for them to know where to enter the PIN.

(v) Examples of functional accessibility problems

People with disability and community and advocacy groups outlined many examples of products and services that use Digital Technology that is not functionally accessible.[978] Some of these are summarised in Table 2 below.

Context	Digital Technology	Limits on functional accessibility
Domestic and personal	Connected and smart digital devices	Some digital home assistants may be more accessible for people who are blind or have a vision impairment, but those that require voice input are inaccessible for people with complex communications needs.[979]
	Whitegoods and electrical devices	Products with touch screens are sometimes inaccessible for people who are blind or have a vision impairment. Connectedness through IoT may help address this problem—at least for people with disability who are able to connect online.[980]
Employment	ICT	Some content management systems, client referral systems, internal human resources and finance, and business technologies are inaccessible for people with different types of disability.[981]
News and entertainment services	Broadcasting and video content	There is currently no minimum requirement for audio description, for people who are blind or have a vision impairment, in respect of free-to-air television, on-demand and subscription broadcasting and video content services.[982] A high volume of broadcasting and video content available to the Australian public is inaccessible to people who are deaf or have a hearing impairment. There are recommendations to increase captioning quotas.[983] There are minimum captioning quotas under the BSA. However, the quality and frequency are not routinely monitored, with investigations only occurring when a complaint is made to ACMA.[984]
Online information	Web accessibility	Web accessibility has improved through wider use of accessibility guidelines, but there remain many inaccessible web pages and services,[985] especially for people with print disability.[986] Inaccessible tools such as CAPTCHA,[987] which are incorporated into otherwise accessible webpages, can render the whole content inaccessible for the user.[988]

(vi) Functional accessibility: the role of government

Stakeholders representing people with disability advocated a range of measures to combat the prevalence of inaccessible technologies. Some submitted that the Australian Government should increase captioning quotas and introduce minimum quotas for audio description on free-to-air television, on-demand and subscription broadcasting.[989]

ACCAN submitted that captioning should be increased to cover all free-to-air television channels for 24 hours per day, audio description be introduced at a minimum of 14 hours per week across all free-to-air channels with mandatory annual increases, and that all video content distributed in Australia include accessible features.[990]

Some stakeholders recognised the important role of the private sector by focusing on commercial incentives to increase private sector adherence to accessibility requirements, and the adoption of accessible design approaches (discussed below). Some stakeholders proposed that all levels of government be required to procure accessible ICT (including ICT that complies with WCAG 2.1 and Australian Standard EN 301 549).[991]

Two stakeholders pointed to the Voluntary Product Accessibility Template (VPAT) and submitted that Australia could adopt a similar template. VPATs are used by ICT providers to explain to government agencies how their product complies with accessibility standards under US law.[992]

Portable, an Australian digital design company, submitted that government tenders for new technology (or service re-design) should include research phases that bring diverse voices into the room and include as many perspectives as possible.[993]

ANZ suggested that major providers of ICT to government and banking services be required to meet the procurement standard EN 301 549 in their own business.[994] Similarly, the Australian Rehabilitation and Assistive Technology Association submitted that agencies, such as the NDIA, should adopt purchasing and procurement policies that promote flagship examples of excellence in local inclusive design.[995]

ACCAN recommended that Australian businesses make information about the accessibility of their goods and services available and accessible to consumers.[996]

Other incentives to increase private sector commitment to accessibility could include:

- incorporating accessibility requirements into the terms of government grants[997]
- tax concessions and grant incentives to businesses that provide accessible technology[998]
- a trust mark to symbolise a business' adherence to accessible technology[999]
- industry awards, prizes and showcases for best designs and processes.[1000]

(vii) Digitisation replacing accessible services

Stakeholders provided examples of public and private services which have implemented inaccessible technologies that wholly or substantially replace previously accessible services.[1001] The National Association of Community Legal Centres noted that digitisation of essential and government services can create serious problems for people with disability:

> [E]xisting inequalities are often perpetuated through access to technology, and access to technology is increasingly central to engagement with essential services and opportunities.[1002]

Stakeholders provided many examples of how changes associated with digitisation can make some goods and services inaccessible. These include:

- Some machines used for credit card payment in shops require payment via a POG touch screen, and automated self-service and information kiosks can be inaccessible or more difficult to use for people with vision impairment.[1003]
- Automatic airline check-in using passport facial recognition technology to replace human customer support, may present barriers for people with eye or facial differences.[1004]
- Smartphone transport applications that replace the provision of Braille or talking text at transport interchanges are difficult for people who do not have access to mobile phone applications or are not confident digital users.[1005]
- Interactive voice response software, which is often used on service telephone lines, is inaccessible for some people with intellectual disability who require human customer support.[1006]
- Essential government services, such as My Health Record, My Aged Care and NDIS becoming primarily web-based consumer portals may present barriers for people with limited or no Internet coverage, the required software compatibility with devices, and computer literacy.[1007]

Stakeholders supported all levels of government communicating in ways that make goods, services and facilities more accessible. Specific desirable changes included better adherence to WCAG 2.1 and Australian Standard EN 301 549[1008] (noting the progress made by the DTA and NSW State Government in this respect);[1009] providing important information in Easy English;[1010] and providing an option to speak directly with an operator.[1011]

Other stakeholder proposals directed towards government included:

- updating the Australian Government's Digital Service Standard and the Commission's World Wide Web Access DDA Advisory Notes version 4.1 (2014) and keeping these documents up to date as accessibility guidelines and standards are released[1012]
- committing to employ more people with disability across the Australian Public Service and to meet the existing targets, through the provision of flexible arrangements.[1013]

9.3 The Commission's preliminary view—functional accessibility

The right to access technology, and especially several new and emerging technologies, is an enabling right that supports people with disability to realise a range of political, economic, social and cultural rights. This perspective was shared by many stakeholders.

(a) Experience of inaccessible Digital Technology

There is a lack of national data about the accessibility of Digital Technologies. However, the Commission's consultation has provided a clear picture of the challenges and opportunities for people with disability accessing Digital Technologies.

The Commission heard from stakeholders representing thousands of people with disability across the nation. Many of the views expressed are broadly consistent with stakeholder input to a 2017 parliamentary committee looking at inclusive and accessible communities,[1014] and a 2016 DSS report on the experiences of people with disability in Australia.[1015]

In particular, there are significant difficulties associated with functional accessibility, obtaining Digital Technology, and the intersectionality with other socio-economic factors. The Australian Bureau of Statistics has also reported that people with disability experience higher rates of difficulty in accessing service providers compared with people without disability.[1016]

Inaccessible Digital Technologies can have a profoundly negative impact on people with disability. As a State Party to the CRPD, Australia has an obligation to eliminate obstacles and barriers to accessibility.

(b) Government leadership

(i) Government services and publicly-funded services

The CRPD requires States Parties to adopt all appropriate legislative, administrative and other measures for the implementation of the rights contained in the CRPD.[1017] This includes providing accessible public services,[1018] which necessarily entails the procurement of accessible goods, services and facilities.

The UN Committee on the Rights of Persons with Disabilities notes that 'it is unacceptable to use public funds to create or perpetuate the inequality that inevitably results from inaccessible services and facilities'.[1019]

Government leadership is vital in promoting human rights compliant Digital Technology, especially through the provision of public services. Government procurement policies can be a lever for change in large organisations and government itself.[1020]

For example, the Commonwealth Indigenous Procurement Policy incentivises the inclusion of Indigenous enterprises at all levels of the supply chain to 'help maximise the development of the Indigenous business sector'.[1021] Outside government, procurement policies are also gaining traction, as demonstrated by the Business Council of Australia's 'Raising the Bar' initiative which is a $3 billion procurement agreement between some of its member organisations and Indigenous suppliers.[1022]

Some government agencies are taking positive steps to promote accessibility. A future priority for the National Disability Strategy includes promoting universal design principles in procurement, and all governments adopting web accessibility standards.[1023]

The Commission considers that the Australian Government could do more and adopt a more consistent approach. For example, there is currently no whole of government approach to the provision and procurement of accessible goods, services and facilities, with a view to meeting the latest international accessibility standards. A commitment to do so would improve access to government services for people with disability.

The Commission proposes the adoption of government-wide accessibility and accessible procurement standards. This would enhance accessibility for public sector employees and users of public services. It would also model to the private sector the importance of providing accessible Digital Technology and incentivise meeting these requirements for any involvement in public sector procurement or funding.

The DTA strongly encourages the use of WCAG 2.1 across Australian Government services, and the Department of Finance requires procurement of accessible ICT for the Australian Government, according to Australian Standard EN 301 549.[1024]

The Australian Government's National Transition Strategy (NTS) to implement the earlier WCAG 2.0 commenced in 2010 with a strategy for implementation across all levels of government over four years.[1025] There is no recent Government audit or evaluation on the NTS and adoption of WCAG 2.0.[1026] An external assessment suggested that conformity with WCAG 2.0 varied across agencies, and that the NTS was 'successful in the raising of awareness of the issues and requirement of website accessibility, particularly for government agencies'.[1027]

The experience of people with disability accessing public services is likewise varied, depending on the service and the person's disability. The Commission considers that the Australian Government should implement a transition to WCAG 2.1, incorporating lessons from the NTS, in partnership with the community and ICT sector.

There should also be an update in procurement policies to include WCAG 2.1, and to encourage a preference towards providers that conform to those accessibility requirements in their own organisational operations.

These policies would further improve accessibility across major public and private institutions that receive government funding. Government leadership on accessible Digital Technology could also involve the promotion of best practice and showcasing of leading goods and service providers in the field.

PROPOSAL 20:

Federal, state, territory and local governments should commit to using Digital Technology that complies with recognised accessibility standards, currently WCAG 2.1 and Australian Standard EN 301 549, and successor standards. To this end, all Australian governments should:

(a) Adopt an accessible procurement policy, promoting the procurement of goods, services and facilities that use Digital Technology in a way that meets the above accessibility standards. Such a policy would also favour government procurement from entities that implement such accessibility standards in their own activities.

(b) Develop policies that increase the availability of accessible communication services such as Easy English versions and human customer supports.

PROPOSAL 21:

The Australian Government should conduct an inquiry into compliance by industry with accessibility standards such as WCAG 2.1 and Australian Standard EN 301 549. Incentives for compliance with standards could include changes relating to taxation, grants and procurement, research and design, and the promotion of good practices by industry.

(ii) Broadcasting services

The Commission supports calls for broadcasting and online content to be made more accessible for people with disability.

As discussed above, this consultation process has revealed a significant barrier for people who are blind or vision impaired when accessing free-to-air television, on-demand and subscription broadcasting.

People with disability have a right to access broadcasting services, including news and entertainment broadcasting on traditional media like television, and newer forms of content transmission such as online streaming services and social media.[1028] In particular, people who are blind or have a vision impairment have long advocated minimum quotas for audio description on free-to-air television, on-demand and subscription broadcasting.[1029]

Australia is the only English-speaking OECD country without compulsory minimum audio description for free-to-air television.[1030] Audio description is only available to Australian audiences through some DVDs and a few programs on international subscription video-on-demand services, according to a 2019 ACCAN and Curtin University national scoping study.[1031]

Stakeholder feedback is consistent with the findings in the national scoping study on audio description—people who are blind or with vision impairment have very limited access to accessible broadcasting content and video services.

By contrast, the BSA requires national broadcasters to caption all news and current affairs programs and any program screened on primary channels between 6pm and midnight, to provide access to people who are deaf or have a hearing impairment.[1032]

Some groups in the disability community have urged mandatory minimum accessible content for broadcasting services, with progressive increases of this content over time. For example, one suggestion was for a minimum of 14 hours per week (with mandatory annual increases) for audio description,[1033] and 24 hours for captioning.[1034]

'Amend the *Broadcasting Services Act 1992* (Cth).'

The Broadcasting Services Amendment (Audio Description) Bill 2019 (Cth), a private member's bill introduced by Senator Steele-John, proposes a minimum of 14 hours for the first three years of operation, followed by annual increases to 21 and 28 hours per week.[1035] These requirements would apply to national broadcasters, commercial television and subscription services.

The Commission proposes that a priority should be improving the accessibility of content that is broadcast by national broadcasters, commercial television broadcasters and subscription television licensees, all of which are regulated by ACMA.

PROPOSAL 22:

The Australian Government should amend the *Broadcasting Services Act 1992* (Cth) to require national broadcasting services, commercial broadcasting services, and subscription broadcasting services to:

- **(a) audio describe content for a minimum of 14 hours per week for each channel, with annual increases**
- **(b) increase the minimum weekly hours of captioned content on an annual basis.**

The Australian Government should consult with people with disability and their representatives throughout this proposed reform process and in relation to annual increases.

Minimum hours of audio description and increased hours of captioning in broadcasting services could help promote increased accessibility in other forms of content that require increased accessibility such as video, DVDs, films, social media, and online services.

9.4 Obtaining Digital Technologies

Inaccessibility can also arise because of difficulties in obtaining Digital Technologies. This refers primarily to the individual being able to acquire goods and services incorporating particular technologies, not to the physical or intellectual characteristics of their disability.

For example, socio-economic disadvantage, which is more common among people with disability, can make it harder to access technology. A smartphone may have functional accessibility for a person with vision impairment, but a user may not be able to afford it.

The consultation process revealed several barriers for people with disability in their experience of obtaining Digital Technology. These barriers may be grouped into three broad categories: (physical) access, affordability, and digital ability. The Australian Digital Inclusion Index (ADII) seeks to measure progress against these measures.[1036]

The various barriers having an impact on a person's access to Digital Technology are often interrelated. Stakeholders submitted that the 'digital divide',[1037] and other barriers to gaining access to Digital Technology, can further exclude people with disability.[1038]

Digital inclusion is influenced by differences in income, education levels, and the geography of socio-economic disadvantage.[1039] Low internet access is correlated strongly with disability, low family income, and employment and education status.[1040]

Further, the distribution of poverty and inequality in Australia means some people start from a position of disadvantage when it comes to digital inclusion, and there is a real risk that the changing digital environment may exacerbate experiences of poverty and inequality.[1041]

The Australian Red Cross submitted:

> Pervasive service digitisation and dependence on technology in our private and public lives can further disadvantage vulnerable Australians. Improving digital inclusion is critical to ensure that everyone in our community is empowered to participate and contribute. Technology can empower people in so many ways – they can stay connected and involved with their social and community networks, access knowledge and services to help them stay well, or link to learning and job opportunities.[1042]

The National Association of Community Legal Centres provided an example of intersectionality. 22.4% of community legal centre clients have at least one disability. The entire client group faces barriers in accessing technology, including the high cost of accessing the Internet, limited access to training in the use of technology, and limited capacity to access essential services online.[1043]

(a) Physical access

The Commission's consultations revealed that people with disability experience greater difficulty in obtaining access to Digital Technology, compared with other groups. Some stakeholders referred to the ADII, which rates the degree of access of people with disability to the internet and digital platforms as lower than in the general population.[1044] Other stakeholders referred to digital or social exclusion and the relationship between socio-economic factors and access to Digital Technology.[1045]

Several stakeholders raised the concept of 'internet freedom' as a human right,[1046] pointing to unequal access to media, information and communications infrastructure across the population. Case studies on the right to access technology were often interrelated with socio-economic factors and inequitable outcomes for people with disability.[1047]

(b) Affordability

Stakeholders frequently raised the issue of cost for accessible and assistive technology.[1048] Those at risk of being excluded due to the cost of technology include: people with disability and older people who are more likely to be on low incomes;[1049] and people who are unemployed, underemployed or on a pension.[1050]

One participant gave an example of how difficult it was to buy an accessible TV as a blind person in her regional town.[1051] Researching potential TVs was difficult given the digital and hard copy instruction manuals were inaccessible. After finding an accessible TV, the cost was over $1000 more than comparable smart TVs.

Stakeholders submitted that assistive technology and technologies that make conventional products and services more accessible are often costly. Examples include screen readers, voice activation software and eye-gaze software.[1052]

(c) Digital ability

The concept of digital ability covers the individual characteristics, experiences and preferences of the user of the technology, such as their attitudes towards technology and basic skills required to use technology effectively. Stakeholders focused in particular on four barriers to digital ability for people with disability.

The first barrier relates to the skills required to use new technologies. Speech Pathology Australia noted that people with communication disability face a challenge in obtaining the computer literacy required.[1053] Further, there is relatively little easy-to-read online information.[1054] Jargon associated with new technologies—sometimes known as 'tech-speak'—can also make it difficult to participate.[1055]

People with disability often need training to gain the full benefit of new technologies.[1056] Training can also be necessary for support people.[1057] Further, information about accessible technology itself can be inaccessible. Consumers sometimes need to purchase a product before finding out whether the product will be accessible for their needs.[1058]

Secondly, there can be relatively low awareness and understanding of new technologies among people with disability,[1059] and their supporters.[1060] It can be difficult to keep up to date with new technological developments and opportunities for their individual circumstances.[1061]

For example, the Ability Research Centre said:

> There is a significant paucity of knowledge about what assistive technology is available, and how all kinds of technology can be harnessed for the specific purpose of assisting people with disability … In essence, the sector languishes from a lack of clear knowledge about new, often generic technology, how it integrates with existing, often specialised, technology, how options can be meaningfully evaluated, compared and trialled, and where to turn for this information.[1062]

Thirdly, people who have been victims, or are at greater risk, of abuse and exploitation through technology, may be more reluctant to use technology.[1063] Where new technology products and services are implemented in a context where the trust of people with disability has been negatively affected, this can reduce those people's willingness to use or have confidence in them.[1064]

Fourthly, carers of people with disability are more likely than non-carers to have lower incomes, not otherwise be in the labour market, have lower levels of education and have a disability or illness themselves.[1065] Carers Australia suggests that this means carers are more likely to be disproportionally represented in groups with low digital inclusion rates. Many carers rely on government and community services and on social security—these are overwhelmingly accessed through digital channels.[1066]

(d) Intersectionality with other population cohorts

As noted above, the factors contributing to barriers to accessibility are complex and interrelated. The Law Council of Australia submitted:

> Unequal access to new technologies can exacerbate inequalities, especially where access is affected by factors such as socio-economic status, geographical location and cultural or linguistic diversity.[1067]

Stakeholders reported that digital exclusion can compound when disability is linked with some other vulnerable groups, including the following:

- Older people may experience issues with technology and systems such as e-government services. Many older people report being uncomfortable with online systems or finding it difficult to participate because of disabilities such as low eyesight or having a shaky hand which makes mouse use difficult.[1068]
- Aboriginal and Torres Strait Islander people with disability can face acute inequality across all relevant support services,[62] including transport, employment and education.[1069]
- People with disability who are experiencing health conditions and whose health could benefit from technological innovations, such as self-monitoring systems for diabetes, often find those systems inaccessible.[1070]
- People in rural, regional and remote areas who have inadequate and unreliable internet coverage[1071] and lower access to education, employment and social amenity, can find these factors exacerbate the digital divide.[1072] People in these areas who live with chronic diseases sometimes rely on tele-health services, which are reimbursed by Medicare on some but not all occasions and require digital connectivity.[1073]

9.5 The Commission's preliminary view—obtaining Digital Technologies

(a) Private sector leadership

Some designers and developers are pioneering accessible Digital Technologies, while others do not appear to prioritise accessibility. The role of 'human rights by design' is considered in the next section; however, there is scope for the private sector to improve the accessibility of products, goods and services at the point of sale and use by the consumer.

Some of the barriers to access raised by stakeholders may be addressed quite simply by Digital Technology providers. For example, information and instructions directed towards meeting the needs of people with disability would support informed consumer choice.

Goods, products and services should come with accessible instructions and dedicated information about accessible features. For example, a smart TV's accompanying product instructions should be in an accessible format and include information about its accessibility features and functionality. Businesses that adopt such measures could benefit from increased market share and goodwill among the disability community.[1074]

As noted above, Standards Australia works with industry and the community to develop specifications, procedures and guidelines to ensure consistent and reliable products, services and systems.[1075] The Digital Technology industry should work with Standards Australia and the disability community to develop standards, technical specifications or other guidance material regarding accessible instructions for goods and services.

PROPOSAL 23:

Standards Australia should develop an Australian Standard or Technical Specification that covers the provision of accessible information, instructional and training materials to accompany consumer goods, in consultation with people with disability and other interested parties.

The barrier of affordability was a common theme in consultations. Affordability was raised as a concern in relation to a range of technology from the cost of Internet, smartphones and accessible whitegoods through to assistive technology such as screen readers. Consultation showed that people with disability face challenges including: fewer consumer options when accessibility is needed;

the need for specialised assistive technology; and the potential for higher up-front connection fees for telecommunications services if people require specialist equipment or software for a vision or hearing impairment.[1076]

These results confirm what other Australian research has shown: people with disability experience more barriers to digital inclusion than the general population.

In an attempt to improve digital inclusion for people with disability, ACCAN advocated that the National Broadband Network (NBN) provide a wholesale broadband concession for low-income households.[1077] Such a scheme was supported by national NGOs, including the Australian Council of Social Service and the Australian Digital Inclusion Alliance.

ACCAN suggested that the NBN offer low-income households a 50 mbps unlimited broadband service at $20 per month, which would result in eligible households paying approximately $30 per month through a service provider—about half the current average cost.[1078]

ACCAN stated that preliminary estimates indicate that the provision of this concession to the 2 million Australian households on the lowest incomes could be budget neutral when offset against outcomes of increased take-up of broadband services including an increase in average incomes and creation of new businesses.[1079]

The Internet is a significant gateway for people with disability to be connected with services, to give and receive information and enjoy their human rights, including to freedom of expression and opinion.[1080]

There is an opportunity to increase digital inclusion for people with disability through the implementation of an affordability scheme like that proposed by ACCAN. In assessing any such scheme, consideration should be given to people with disability who might be financially vulnerable, such as people who are studying, have high support needs or are on rehabilitation pathways.

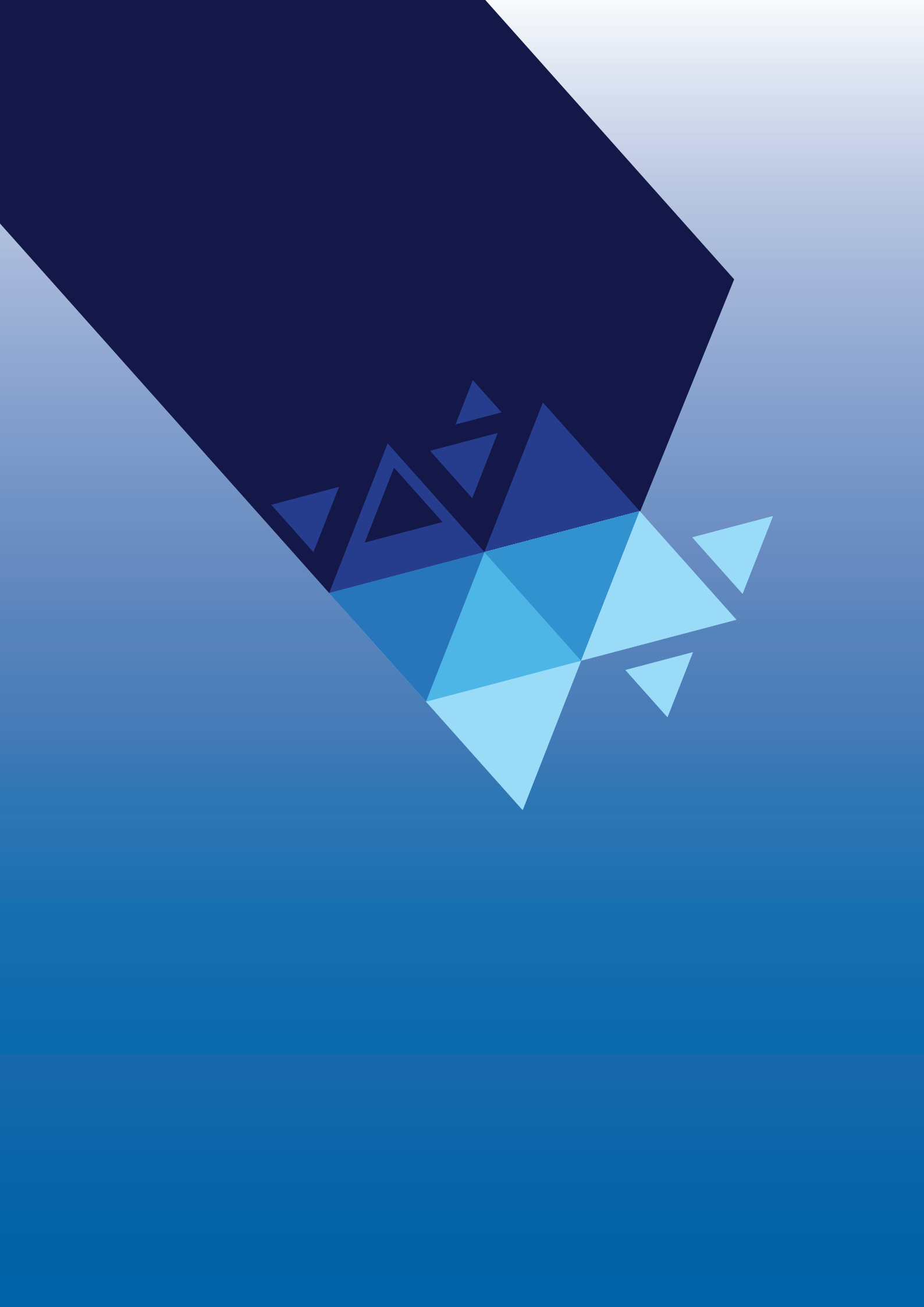
PROPOSAL 24:

The National Broadband Network should undertake economic modelling for the provision of a concessional wholesale broadband rate for people with disability who are financially vulnerable.

The Commission is interested in hearing about other measures that the private sector may take to address affordability barriers for people with disability and to improve digital inclusion.

QUESTION G:

What other measures could the private sector take to eliminate barriers to accessibility related to the affordability of Digital Technologies for people with disability?



CHAPTER 10

Design, education and capacity building

10.1 Introduction

This chapter focuses on how incorporating accessibility into the design and development process improves the functional accessibility of products and services that incorporate new technologies.

The Commission asked stakeholders about the role of *design* in ensuring new and emerging technologies are accessible for people with disability. Stakeholders overwhelmingly emphasised that the right to access technology needs to be a major consideration in the entire production process; from concept, research and design through to implementation, use and upgrading.

A broad cross-section of stakeholders agreed that the starting point is to address lack of knowledge about the right of people with disability to access technology and the obligations on businesses to design and develop accessible Digital Technology. Representatives from the disability and community sectors, industry and academia all pointed to inadequate awareness or understanding about accessibility, especially in the Digital Technology industry.

The Commission's consultations showed strong support for building the capacity of designers and engineers to understand the importance of 'human rights by design' in their role as technology developers. This included support for developing a better understanding, especially in the Digital Technology industry, of the commercial benefits of producing human rights compliant Digital Technology.

10.2 Human rights by design

An overarching theme from the consultation process was that we all benefit if the right to access technology is embedded across the entire production process—from product concept through to final product manufacture and upgrades.[1081]

Most obviously, people with disability would benefit from being able to access technology. However, benefits also would flow to other members of the community through more inclusive functionality. The technology industry would benefit from a bigger consumer market and a better reputation.

Stakeholders emphasised that the right of people with disability to access communications technology should be foundational, and considered in the earliest conceptual, research and design phases. For example, Google submitted:

> Access and accessibility need to be considered from the beginning of the development cycle in order to produce technologies and products that contribute positively to people's lives.[1082]

The focus is on innovation and best practice in design. Industry stakeholders provided positive examples of accessible design in their production processes, and disability advocacy groups also recognised the importance that some businesses have placed on accessible design and resulting benefits for the disability community.[1083] Digital Gap Initiative submitted:

> in instances where accessibility is provided, our experience is that accessibility is a design feature added on towards the end of the product lifecycle, as opposed to being considered in early design phases. The consequence is that accessibility is not coherently integrated into products, including websites, and thus provides an inconsistent experience for end users.[1084]

Significantly, embedding accessibility into the design process can reduce disputes, and indeed litigation, regarding unlawful discrimination.

(a) Design approaches

There are various frameworks for designing accessible Digital Technology. These principles may be broadly termed 'human rights by design'—a process that includes incorporating human rights norms into the concept, research, design, testing and manufacturing phases of technological innovation with the intended outcome of human rights compliance in production and use of the Digital Technology.

'Human rights by design' involves consideration of all the key human rights engaged in using the product or service that is being developed. Therefore, in addition to the right of people with disability to access technology, such an approach also would draw attention to the engagement of other rights, such as privacy and non-discrimination.[1085]

BSR suggests that 'human rights by design' principles could be aligned with 'privacy by design' principles—proactive rather than reactive to human rights abuses; embedded into the design and architecture of IT systems; and requiring the designers and operators to keep the interests of the individual uppermost.[1086]

Stakeholders suggested several design approaches. In this section, we assess four of these approaches: universal design; accessible design; inclusive design; and co-design. These approaches overlap in their operation and conceptual bases.

All of these approaches fall within the Commission's overarching 'human rights by design' rubric for two reasons. First, they share a goal of producing more and improved protection and fulfilment of human rights in product design (and in particular the right of people with disability to access technology).

Secondly, each approach focuses on the importance of embedding accessibility into the entire technology cycle—from concept, research and design, testing and production through to implementation and use—but focusing especially on the earliest phases of this cycle.

Table 4: Human rights by design and other design approaches

Human rights by design Considers and incorporates human rights into the whole design and development cycle with the goal of protecting human rights. This includes rights to access to technology, privacy, non-discrimination etc			
Universal design	**Accessible design**	**Inclusive design**	**Co-design**
Goal: Usable by all people to the greatest extent possible	Goal: Independent use by people with disability	Goal: Usable by all people through adapting to different needs of individuals	Goal: Greater accessibility through design by people with disability

(i) Universal design

'Universal design' is an approach to designing 'products, environments, programs and services, so that they are usable by all people, including people with disability, to the greatest extent possible, without the need for adaptation or specialised design'.[1087] The CRPD expressly endorses universal design.

Seven principles for universal design were developed in 1997 by the Centre for Universal Design, North Carolina State University. They are:

- equitable use
- flexibility in use
- simple and intuitive use
- perceptible information
- tolerance for error
- low physical effort and
- size and space for approach and use.[1088]

Universal design seeks to ensure that products are usable by the entire community with as little adjustment as possible, minimising the need for retrofitting accessibility.[1089]

Stakeholders generally agreed that applying universal design principles can be important in improving accessible Digital Technology.[1090] However, Harpur noted that universal design is not a panacea for all people in every circumstance:

> The definition of 'universal design' in the CRPD acknowledges that inclusive access cannot always be provided. Under universal design access should be provided '... to the greatest extent possible'. Where universal design cannot be achieved, then the second prong becomes relevant—the right to reasonable accommodation or adjustments.[1091]

(ii) Accessible design

Universal design focuses on increasing usability for all people. By comparison, accessible design focuses more specifically on the needs of people with disability.[1092]

For example, adhering to WCAG 2.1 during website design and development results in content that is more accessible to a wide range of people with disability, including accommodations for blindness and low vision, deafness and hearing loss, limited movement, speech disabilities, photosensitivity, and combinations of these; and some accommodation for learning disabilities and cognitive limitations.[1093]

There is no single solution that will meet the accessibility needs of all people with disability.[1094] Adobe submitted:

> Because different users have different needs, it is important to acknowledge there is no single perfect accessible final result… Design is not engineering, and you can't 'certify' designers the way you can products. The creative phases of product design are too abstract to pin down that way. What they need to do is create some kind of incentive to build it right from the start.[1095]

One stakeholder noted that human interaction—the user interface—is key. Within many technology companies, this requires the involvement of the product manager, user experience designer and the developer of the product.[1096]

WebKeyIT submitted that universal design and inclusive design principles are important and useful, but that accessibility requires giving special consideration to people with disability and compliance testing according to externally developed standards, such as WCAG 2.1.[1097]

(iii) Inclusive design

Inclusive design is similar to universal design, as it considers the full range of human diversity with respect to characteristics such as ability, language, gender and age. Unlike universal design, which is 'a common design that works for everyone', inclusive design is a 'design system that can adapt, morph, or stretch to address each design need presented by each individual'.[1098]

The Inclusive Design Research Centre emphasises three fundamental principles in inclusive design: recognising diversity and uniqueness; inclusive processes and tools; and broader beneficial impact.[1099]

The Centre for Inclusive Design submitted that every decision has the potential to include or exclude people, and inclusive design emphasises that understanding user diversity contributes significantly to informing these decisions—maximising inclusion.[1100] Adobe stated that learning from their experience, inclusive design is in everyone's interests.[1101]

Stakeholders provided examples of how universal, accessible and inclusive design principles overlap in theory and in practice. For example, the Australian Banking Association developed the 'Accessibility Principles for Banking Services', which builds on the foundation of the '7 Principles of Universal Design' and WCAG principles.[1102]

The Association reported that it adopts an inclusive approach in its principles, stating that an inclusive design methodology is critical to 'ensure banking services in Australia are optimally placed to deliver the best accessibility and inclusive experience for their users.'[1103] This overlap has been described as the application of universal and inclusive design principles to improve the usability and accessibility of technology for people with disability, and other users of technology.[1104]

Stakeholders suggested that inclusive design is likely to improve the functionality of new technologies for all users, because it encourages designers to be more creative and innovative.[1105] 'Edge users'—that is, people who are not considered to be among the mainstream users of a product or service—are included in the design process. They are 'less likely to defend and validate a current design that doesn't meet their needs', generating greater creativity in the design process.[1106] This can

> respond not only to disability needs but also enhance access for people of all ages and varying literacy abilities, thus addressing human rights principles regarding non-discrimination and equity of access for all.[1107]

Two stakeholders submitted that the functionality of AI technologies is enhanced for all users through inclusive design as it 'informs more adaptive and innovative AI systems',[1108] and 'outlier data is sought and valued, voices are heard and consultation is sought, myriads of unexpected users derive benefit from inherently accessible design.'[1109]

(iv) Co-design and workplace diversity

Some stakeholders emphasised the importance of including people with disability in all design phases.[1110] This gives rise to two related concepts: ensuring that workplaces are diverse and include people with disability, and developing co-design practices that include people with disability in the design process as consultants.

It was submitted that designers are more likely to consider the accessibility needs of people with disability if people with lived experience of disability form part of design teams and the broader businesses or organisations they sit within.[1111] Disability groups can also be included in the design process.[1112]

Advocates noted that people with disability are sometimes asked to be involved in the design of products but, unlike other industry consultants or experts, are generally not remunerated for their expert input.[1113] They reported the relatively common experience of a business engaging in the co-design method by involving one or more people with similar disability, missing out on a wider range of user experiences and needs.[1114] The Australian Centre for Health Law Research submitted:

> human rights principles of dignity and autonomy demand that [people with disability] are supported to participate in decisions around how technology is to be designed and delivered.[1115]

(b) Education, training and capacity building

Stakeholders broadly agreed on the importance of building the capacity of designers and engineers through increased education and awareness about the right to access technology. This view was shared among accessibility consultants, people from the disability and community sectors, academia and industry.

It was submitted that the technology industry could benefit from increased education, training and capacity building across their educational and professional life—from secondary and tertiary education through to ongoing professional development.[1116]

Stakeholders urged the introduction of a core university and professional development 'human rights by design' subject for technologists, designers, engineers, scientists and architects,[1117] and as part of any course that requires decision making on goods and services available to the public, including marketing, product management and social policy.[1118]

This course could include content on disability, international human rights obligations, the CRPD, accessible and other types of design, and best practice. ANZ submitted that accessibility would generally be improved if it was a core subject for university curricula so that those entering the workforce understand the issues and barriers, and how they might be able to play a role in improving accessible design.[1119]

Intopia suggested that accessibility training should be integrated into secondary courses on IT and in vocational training settings.[1120]

Accessibility consultants and disability advocates said that an understanding and awareness of accessible design could be enhanced through a national education campaign with technology industry stakeholders.[1121] It was said that professionals working in this space would have a greater understanding and appreciation of the issues if they are part of diverse workplaces that include people with disability,[1122] especially at senior levels.[1123]

Several stakeholders raised related ideas about expert accessibility leadership,[1124] or a national certification scheme.[1125] An independent body could accredit accessibility certifiers, to help build capacity in a sector by training and equipping professionals. This body could also monitor compliance with accessibility standards.

The University of Technology Sydney raised the possibility of an expert panel of user groups to provide resources to support 'human rights by design' strategies. The panel could be constituted by people with disability, available to contribute to design, testing and review of emerging technology.[1126] Intopia suggested the introduction of a 'National Reference Group', based on the US Access Board, a federal agency that leads the design and development of accessibility guidelines and standards in the United States.[1127]

A wide range of stakeholders agreed that education and awareness raising for the general community also would be beneficial in helping the public understand issues of inclusion and accessibility, and consequently inform consumer choices.[1128] Speech Pathology Australia noted the importance of raising understanding and awareness in areas of the public and private sectors where there are known to be significant communication barriers for people with disability—especially the justice system, primary health services and general practice, hospital systems, aged care systems, and local government consumer-facing services.[1129]

(c) Business case for human rights by design

Stakeholders across all sectors pointed to the commercial opportunities for businesses that implement 'human rights by design'. LexisNexis submitted that industry can understand the value of human rights compliant design for their shareholders, employees and customers through lessons learned from corporate social responsibility (CSR) frameworks.[1130]

In particular, stakeholders referred to the long-term financial value of gaining a competitive advantage, increased market share and reputation, and greater resilience to external forces such as market changes and law reform.[1131]

Stakeholders pointed to the opportunity to grow market share by strengthening existing markets, as well as creating and growing new markets.[1132] A recent report prepared by PwC on behalf of the Centre for Inclusive Design estimated that an inclusive design approach can increase the potential market of products and services by three to four times, at least in the three Australian industries assessed in the report.[1133]

Delivering innovative products to consumers and solving 'unmet need' can help businesses tap into network effects and increase customer base and market share.[1134] The University of Melbourne summarised the commercial imperative as follows:

> In the age of personalisation, a device which cannot be used in every environment is not competitive.[1135]

Technology products may be initially designed for people with disability but end up benefiting all customers. For example, captions on television and SMS or text messages were both initially developed to support people who are deaf or with hearing impairment receive and impart information, but are now more widely used by others as well.[1136]

Other CSR benefits for employees and shareholders include the positive links between organisational commitment to the social cause and high levels of in-role job performance,[1137] and increased shareholder returns from a more diverse employee and board membership to include people with disability.[1138]

On the other hand, industry and community stakeholders also addressed the costs associated with a 'human rights by design' approach. Some observed that the upfront costs of 'human rights by design' could be prohibitive or a disincentive for business, depending on the technology setting and their business capabilities.[1139] Commercial considerations can outweigh accessible design considerations in the fast-paced technology marketplace, where businesses are striving to be innovators and market leaders.[1140]

As discussed in Part 3, it is common for tech businesses to release a 'minimum viable product' (MVP). An MVP undergoes iterative improvements as it is tested and used by consumers, with the aim of producing a more effective and refined final product.[1141]

If the initial version of the product is not accessible for people with disability, they must then wait for further accessibility refinements, which are not guaranteed.[1142]

The Public Interest Advocacy Centre noted that this can result in people with disability being responsible for learning and re-learning successive iterations over time.[1143] An industry participant suggested that designers may benefit from a trial-and-error design process, which is directed towards the goal of accessibility and allows those designers to share what they learn about accessibility with other designers.[1144]

Several industry and community stakeholders agreed that a 'human rights by design' strategy is often more cost effective than retrofitting accessibility into a product after it has been developed and distributed.[1145] Digital Gap Initiative submitted that when accessibility is added to a product towards the end of the product life cycle, 'accessibility is not coherently integrated into products, including websites, and thus provides an inconsistent experience for end users'.[1146] Harpur said:

> Where universal design is adopted, many access barriers are not created in the first place and thus the need to engage in retrofitting is reduced or eliminated.[1147]

10.3 The Commission's preliminary view

The Commission considers there are significant benefits in a 'human rights by design' strategy in respect of new and emerging technologies, and proposes some ways of better integrating this strategy into practice. This approach is an effective way of fulfilling Australia's CRPD obligations to promote the design, development and production of accessible Digital Technology, and promote research and development of universally designed goods, services, equipment and facilities.[1148]

A 'human rights by design' strategy is consistent with some current government policy. For example, a future priority action area for the National Disability Strategy is to improve community awareness of the benefits of universal design.[1149] The consultation showed strong support for 'human rights by design' in Digital Technology production, and for building capacity in the technology industry to design and develop human rights compliant products and services.

There was also strong support among stakeholders across all sectors for building the capacity of technology designers in order to provide benefits for the whole community through more usable products and services. The Commission draws attention to the business case for the design, development and use of accessible technology.[1150]

(a) 'Human rights by design'

The Commission sees a 'human rights by design' strategy as encompassing principles contained in the various design approaches raised in the consultation process. Accessible design, universal design and inclusive design are intertwined, with accessibility the goal and the measure of how successful the particular design process is.[1151]

'Human rights by design' principles do not apply solely in the context of disability. They are also being integrated into design and development processes with a view to protecting a range of other human rights as well. These principles are particularly relevant in the use of AI,[1152] and to protecting personal information and privacy,[1153] themes which are further discussed at Chapter 6 in this report.

The decisions made by human designers in the design process can have positive or negative human rights impacts. These decisions will be informed by the individual life experiences and biases of the designers.[1154] Similarly, the positive and negative impacts on the right to access technology resulting from designer decisions will emerge throughout the production and use of technology.

'Human rights by design' introduced at the earliest stage not only benefits people with disability who rely on accessibility features. It also can have benefits for shareholders, employees and other customers. In addition, a strategy that enables accessibility issues to be dealt with early is generally more cost effective than retrofitting accessible features at a later stage.

Taking into account Australia's obligations under the CRPD, and this Project's consultation and research, the Commission observes that a 'human rights by design' strategy generally emphasises the following principles:

- The primary goal is accessibility to all people, to the greatest extent possible, without the need for adaptation or specialised design.[1155]
- People with disability and their representatives should be encouraged to provide meaningful input in the development process, in roles such as designers, co-designers and expert consultants.
- People with disability and their representatives should be encouraged to participate in all phases of the development process—concept, research, design, iterations, testing, production, manufacture and upgrades.

A 'human rights by design' strategy draws on principles that underpin the CRPD—individual autonomy, independence, non-discrimination, full and effective participation and inclusion in society, respect for difference, equality of opportunity, accessibility, equality of men and women, and respect for children.[1156]

The adoption of a 'human rights by design' strategy in government policies and procedures would be an important step in promoting accessible technology. The COAG Disability Reform Council is a forum for member Governments to discuss and progress key national reform in disability policy.[1157] The Council oversees the implementation of the NDIS, and National Disability Agreement and National Disability Strategy reforms, to support people with disability, their families and carers.

The Department of Social Services is engaging in community consultation on the development of a new National Disability Strategy, as COAG prepares to work on a new Strategy at the end of 2020.[1158] The preliminary findings from this Project consultation would suggest that accessible technology be considered as a priority for the next Strategy.

PROPOSAL 25:

The Council of Australian Governments Disability Reform Council should:

(a) lead a process for Australia's federal, state and territory governments to commit to adopting and promoting 'human rights by design' in the development and delivery of government services using Digital Technologies, and monitor progress in achieving this aim

(b) include policy action to improve access to digital and other technologies for people with disability as a priority in the next National Disability Strategy.

(b) Education, training and capacity building

Too often, products and services relying on new Digital Technology are designed and developed in ways that are inaccessible for people with disability. There are many possible causes, but the Commission considers that if a 'human rights by design' strategy were more common, this would partially address the problem.

'Human rights by design' is still a relatively new concept. More could be done to promote this strategy through education, training and capacity building initiatives. The Commission is interested in initiatives to promote 'human rights by design' in the technology sector.

Such initiatives could enhance Australia's compliance with art 9(2)(c) of the CRPD to provide training on accessibility issues. It would also meet a need identified by many stakeholders, including people with disability and representatives from industry, government and academia.

As outlined in Chapter 2, the UN Guiding Principles on Business and Human Rights impose obligations on the private sector to adhere to human rights in the design, production and implementation of new technologies.

Sector capacity building would improve awareness and understanding of the right of people with disability to access technology, and the obligation to design and develop human rights compliant tech products and services. There was much agreement across industry, government and the community about a lack of understanding in the technology industry of the access rights of people with disability. This problem could be addressed by encouraging developers of Digital Technology to incorporate 'human rights by design' principles through existing frameworks, such as CSR.

Consultation suggested two key focus areas: education and training for professionals and students; and national industry-wide and community leadership and guidance.

The first of these efforts could be targeted at tertiary and vocational students of technology, science and engineering. A 'human rights by design' course could cover different models of disability (eg, social, legal and medical), international and national legal frameworks such as the CRPD and DDA, accessible design methodologies, and best industry practice.

Development of a course of this nature would involve expert input from professional engineers, technologists and scientists, the Department of Education and Training, civil society representatives and tertiary and secondary education sectors. The Australian Council of Learned Academies (ACOLA) brings together four independent Learned Academies: the humanities; science; social sciences; and technology and engineering.[1159] ACOLA helps inform national policy through various activities, including coordinating multi-stakeholder groups and consulting on significant national issues.

‘Education providers should include “human rights by design” in relevant courses.’

PROPOSAL 26:

Providers of tertiary and vocational education should include the principles of ‘human rights by design’ in relevant degree and other courses in science, technology and engineering. With appropriate support, the Australian Council of Learned Academies should undertake consultation on how to achieve this aim most effectively and appropriately within the tertiary and vocational sector.

While scientists, technologists and engineers are the primary designers of Digital Technologies, there are other professionals and tradespeople who are involved in the implementation and deployment of these technologies. A ‘human rights by design’ course could be appropriately modified for different disciplines which intersect with the provision of goods, services and facilities to the public which include these Digital Technologies.

QUESTION H:

What other tertiary or vocational courses, if any, should include instruction on 'human rights by design'?

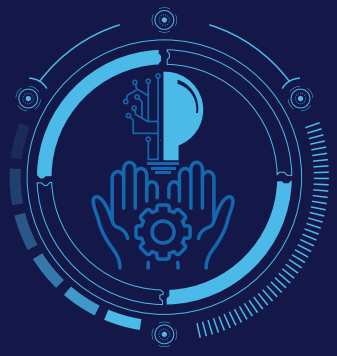
Education and training on 'human rights by design' would benefit professionals who are already practising as designers and engineers. Such training could be offered to Chartered Professional Engineers and Technologists, for example, who are required to undertake 150 hours of continuing professional development (CPD) over three years to maintain their Chartered Status.[1160]

CPD activities support an engineer as they carry out their technical and professional duties, via conferences and training courses, and also allows for the completion of tertiary or post-graduate courses.[1161]

The current CPD requirements include a minimum number of hours to be dedicated to the engineer's area of practice, risk management and business and management skills. A 'human rights by design' course could be a minimum requirement within an engineer's three-year CPD cycle.

PROPOSAL 27:

Professional accreditation bodies for engineering, science and technology should consider introducing mandatory training on 'human rights by design' as part of continuing professional development.



'Professional accreditation bodies should consider mandatory training on "human rights by design".'

The Commission also considers there would be value in targeted capacity building across the technology sector.

There were two main activities identified in the consultation process as vital for industry:

- education and training for public and private entities on 'human rights by design', and
- the creation of an accessibility accreditation scheme to support organisations implement and achieve nationally standardised accessibility benchmarks.

An entity tasked with these kinds of capacity building and accreditation roles could also support education and training efforts such as the development of a 'human rights by design' professional and educational unit of study. There are several suitable organisations that could perform these functions. The Commission invites feedback on the most appropriate entities to do so.

PROPOSAL 28:

The Australian Government should commission an organisation to lead the national development and delivery of education, training, accreditation, and capacity building for accessible technology for people with disability.

CHAPTER 11

Legal protections

11.1 Introduction

The Commission asked stakeholders whether changes are needed to law and policy to promote accessible Digital Technology for people with disability.

People with disability, community and advocacy groups, and academia expressed support for stronger legal protection in relation to discrimination against people with disability in accessing technology.[1162] Some NGOs advocating stronger legal protections also called for increased education and awareness of 'human rights by design' among the technology sector.[1163]

While many stakeholders agreed that a better understanding of accessibility across the technology industry would likely produce a longer-term cultural shift, law reform to promote human rights compliance is also needed.

Three key issues concerning legal protections were identified in the Commission's consultation and are discussed below.

First, if there is a breach of the DDA, responsibility lies with affected persons with a disability to complain and take legal or other action. In other words, there is no independent body that monitors compliance with the DDA, identifies problems and seeks to resolve them. Many stakeholders reported that this creates a heavy burden on the disability sector, which must attempt to perform these roles itself.

Secondly, there are currently no legally enforceable national standards or industry codes dealing specifically with the accessibility of Digital Technology. Some stakeholders saw this as a gap that should be filled.

Thirdly, some stakeholders reported inconsistencies between national and international digital accessibility standards. They argued that some international standards, which could be of benefit in Australia to people with disability, do not apply here.

11.2 Submissions and consultations

Several community and advocacy groups representing thousands of people with disability urged law reform to promote the right of people with disability to access technology.[1164] These stakeholders observed that the current voluntary guidelines, such as WCAG 2.1, are not legally enforceable.

More general obligations prescribed by law—in particular, the DDA—may require aggrieved parties to take legal action, placing the burden on the disability community. Digital Gap Initiative submitted:

> [F]rom first-hand experience, the burden imposed on individuals with disabilities to seek legal remedies under Australian law is far too onerous. Placing the onus on minority communities to lodge complaints and fight for change presents an unnecessary challenge.[1165]

At the same time, some industry and community representatives agreed that the law should not provide harsh punitive measures for non-compliance with legislated accessibility standards, but should be sufficient to enforce compliance effectively.[1166]

(a) Legislative framework

As discussed above, Australia has obligations under international human rights law regarding disability. Many of these obligations are incorporated into Australia's federal, state and territory laws. Some stakeholders emphasised Australia's commitment to uphold human rights according to the CRPD.[1167] They noted that although Australia has implemented some of these obligations through the National Disability Strategy, current Australian law does not fully reflect national obligations to protect the rights of people with disability.

The DDA prohibits discrimination on the basis of disability in employment, education, accommodation and in the provision of goods, services and facilities. An individual may complain to the Commission that they have suffered discrimination on the basis of their disability.[1168]

The Commission can then investigate and conciliate the complaint. If the matter is not resolved informally through conciliation, the complainant can take their matter to the Federal Circuit Court or Federal Court of Australia, which can determine the matter according to law.[1169] State and territory anti-discrimination laws contain similar protections and dispute resolution mechanisms for people with disability.[1170] Stakeholders focused on the DDA and other federal law.

(i) Power imbalance

There can be a power imbalance in any complaint or litigation process, including under the DDA. If a complaint of disability discrimination can be resolved in a no-cost informal complaints jurisdiction, such as the Commission's, this can mitigate the negative effects of any power imbalance between a complainant and respondent.

However, problems can arise where a disability discrimination complaint cannot be resolved in this informal way. It has been observed that, in practice, only people with structural, relational, financial and community support have the capacity to engage in a dispute resolution process such as litigation.[1171]

The emotional, social and financial costs to complainants under the DDA can have adverse impacts on people with disability.[1172] Advocacy fatigue develops for people who are already exposed to systemic inequalities and who are continually fighting for the preservation and advancement of their rights and autonomy.[1173] The University of Technology Sydney submitted:

> These large cases are rare, and they expose a plaintiff to considerable financial risk… Even if a plaintiff wins, they may have costs awarded against them. This is a significant risk to consumers and a barrier to people with disability challenging discriminatory places, products and services.[1174]

(ii) An obligation to provide accessible Digital Technology?

Many community, legal and advocacy groups urged a new legal requirement to provide accessible technology.[1175] This could be achieved through a mandatory industry code, or binding disability standards regulating Digital Technology under the DDA.

As discussed above, Australian Standards are voluntary. Standards Australia is not a government regulator, and does not enforce or monitor compliance with standards.[1176] Some stakeholders suggested that the voluntary nature of such standards presents more of a problem if the law does not provide sufficient protection. For example, the Digital Gap Initiative stated that they

> present a huge structural issue in Australia's framework. While voluntary standards are a positive start, no meaningful action will be taken without some formal mandate and even punitive measures for businesses which fail to comply with legislative requirements.[1177]

In view of the voluntary nature of standards overseen by Standards Australia, stakeholders made a number of suggestions. First, some urged that the DDA prohibit the provision of Digital Technology goods and services that do not comply with internationally recognised digital accessibility criteria such as those contained in the WCAG.

A Disability Standard made under s 31 of the DDA could relate to the provision of accessible Digital Technology.[1178] Such a standard could incorporate 'human rights by design' principles;[1179] be developed in consultation with people with disability[1180] and the technology industry;[1181] and be updated regularly to account for technological changes.[1182] The Digital Gap Initiative submitted that there should be civil penalties for non-compliance with any such standards.[1183]

Harpur submitted that the DDA should require technology designers to consider universal design principles in the design and manufacturing process.[1184]

Some stakeholders from the technology industry and community sector noted that it may be difficult for legislated accessibility standards to stay relevant in the fast-paced technology context.[1185] It was broadly recognised that a new standard, or set of standards, would need to be flexible to accommodate swift technological change, regularly reviewed, and principles-based so it is generally applicable.[1186] Any standard also would need to be expressed in specific or detailed terms in order to be useful to the developers of Digital Technology.[1187]

Secondly, some submitted that there should be compliance monitoring of accessible Digital Technology by an independent regulatory oversight body, which could help support the implementation of Australia's international treaty obligations and complement any industry self-regulatory measures.[1188]

Thirdly, some stakeholders urged the immediate review of current voluntary standards and the adoption of leading international accessibility standards.[1189]

Community, legal and advocacy groups broadly agreed that enforceable accessibility standards would help relieve the burden on the disability sector to redress inaccessible Digital Technology. Kingsford Legal Centre submitted:

> The reactive nature of the complaints system means that it is difficult to address inaccessible technologies once they have already been made available to the market.[1190]

Stakeholders anticipated the potential significant impacts of legislative amendments on the technology industry and suggested:

- an open, consultative approach between any regulatory body and technology service providers [1191]
- a progressive approach to implementation to achieve true compliance through building organisation awareness and capacity[1192]
- resourcing and funding dedicated to assisting organisations comply.[1193]

11.3 The Commission's preliminary view

The CRPD requires Australia to adopt appropriate measures in law for the full realisation of human rights for people with disability,[1194] and to effectively monitor implementation of its provisions.[1195] The UN Committee on the Rights of Persons with Disabilities suggests that national legislation 'should provide for the mandatory application of accessibility standards and for sanctions, including fines, for those who fail to apply them'.[1196]

As noted above, the DDA prohibits discrimination on the basis of disability in employment, education, accommodation and in the provision of goods, services and facilities. It has been rightly observed that the DDA's protection should be improved as it applies to Digital Technologies.

The Commission considers that this would be best achieved through the development of a Disability Standard for Digital Communication Technology. Concerns about the DDA's monitoring and compliance mechanisms should be addressed through the broader anti-discrimination law reform process that the Commission is leading.[1197]

People with disability encounter multiple barriers when accessing Digital Technologies, a conclusion consistent with other reports regarding the experience of people with disability accessing goods, services and facilities.[1198] Most of the consultation feedback from people with disability centred on their experience in accessing Digital Technologies that are used primarily for communication—ICT, smart digital devices, and the Internet. These technologies were identified as enabling people with disability to enjoy greater independence and participation in all areas of life.

The Commission acknowledges the financial and emotional burdens on individuals who make a complaint under the DDA, especially if a complaint cannot be resolved informally or by conciliation. The Commission has previously recorded similar concerns—for example, in the Disability Standards Public Transport Review.[1199]

Participants in that review highlighted the power imbalance between people with disability and the transport industry, and the benefit that would be gained for individuals in providing additional avenues of enforcement for individuals under the Transport Standards.[1200]

(a) Digital Communication Technology Standard

The Commission's preliminary view is that a legally-binding standard or set of standards should be developed to better protect the right of people with disability to access Digital Technologies used for communication.

Such a process should promote the importance of human rights compliant design. The proposed Standard should cover the provision of goods, services and facilities, which use Digital Technologies for communications purposes and which are available to the public. This would include ICT, VR and AR technologies, as the accessibility of these technologies have been a focus of particular concern from many stakeholders.

When Digital Technologies that are used for communication are inaccessible to people with disability, other human rights also can be limited, such as the right to work, education and freedom of expression.

The proposed Standard should apply only to goods, services and facilities that are used primarily for communications. It would not apply, for example, to the many IoT-enabled devices that have a display that enables information to be communicated to a user, such as an IoT-enabled fridge that communicates information about the products inside it.

While important accessibility issues are raised by such devices, this category of goods, services and facilities is sufficiently different to those used primarily for communications—such as mobile phones—that they would need to be considered separately.

Commercial incentives and industry capacity building are necessary, as they promote understanding of the issues and accessible and inclusive practices. However, law reform to enforce compliance also appears necessary to ensure protection of rights to access technology.

Section 31 of the DDA enables the Attorney-General of Australia to make Disability Standards that set out detailed requirements about accessibility in a range of areas.[1201] Disability Standards currently cover education, buildings and public transport domains.[1202] One purpose of the existing Standards is to give clarity to service providers regarding how to make their goods, services and facilities accessible for people with disability.[1203]

It is unlawful for a person to contravene a Disability Standard made under the DDA.[1204] However, where a person acts in accordance with a Disability Standard, the person is complying with the DDA,[1205] and can be certain that they have not unlawfully discriminated against a person with disability.

The introduction of a 'Digital Communication Technology Disability Standard' could benefit the technology industry and the broader community in many ways.

First, it could lead to greater availability of accessible Digital Technology for people with disability, potentially reducing the likelihood of unlawful discrimination.

Secondly, it would provide greater certainty for industry regarding their obligations to design and develop accessible Digital Technology, and guidance on how to undertake 'human rights by design' processes, enhancing compliance with human rights laws.

Thirdly, as discussed above, the development of new technology products and services which are more usable, may bring commercial benefits, as well as benefiting the broader community.

The development of a Digital Communication Technology Standard would raise a range of practical questions, including:

- should there be one or more Standards to cover different types of technologies?
- how should the balance be struck between principles, outcomes and guidelines-based content?
- how should such a Standard respond to swift technological change?

A Digital Communication Technology Standard would likely need to include principles, outcomes and guidelines-based criteria (for example, compliance with WCAG 2.1). It would require regular updates to keep pace with a rapidly changing technological environment. It would also need to ensure flexibility for different contexts, and yet be specific enough to be useful for technology developers.

If a Digital Communication Technology Standard is developed, assistance would be needed. Some providers are already prioritising and delivering accessible Digital Technology and would require minimal adjustments. Others might need more substantial support and capacity building for the task ahead.

As outlined throughout this chapter, any law reform should be complemented with strong capacity-building efforts. It should be supported through an open and consultative approach between the Australian Government, technology industry and the community. It could also allow for progressive compliance milestones, as has been the experience with some existing Standards under the DDA.[1206]

PROPOSAL 29:

The Attorney-General of Australia should develop a Digital Communication Technology Standard under section 31 of the *Disability Discrimination Act 1992* (Cth). In developing this new Standard, the Attorney-General should consult widely, especially with people with disability and the technology sector. The proposed Standard should apply to the provision of publicly available goods, services and facilities that are primarily used for communication, including those that employ Digital Technologies such as information communication technology, virtual reality and augmented reality.

QUESTION 1:

Should the Australian Government develop other types of Standards, for Digital Technologies, under the *Disability Discrimination Act 1992* (Cth)? If so, what should they cover?

(b) Compliance framework and data collection

For the reasons given above, the Commission's preliminary view is that a Digital Communication Technology Standard under the DDA would improve protections for people with disability against discrimination from the provision of inaccessible goods, services and facilities. However, some stakeholders submitted that this would not be a complete solution to the discrimination that people with disability experience, as there are no compliance monitoring mechanisms for accessible Digital Technologies available under Australian law.

The proposals for capacity building for the public and private sectors, together with the proposed Digital Communication Technology Standard, would likely increase the technology sector's provision of accessible goods, services and facilities. Nevertheless, it is reasonable to assume there will be challenges in giving full effect to any new Standards focused on Digital Communication Technologies.

For example, there has been difficulty in ensuring that rights are upheld with respect to transport accessibility for people with disability in the absence of additional enforcement mechanisms for the Transport Standards.[1207]

In its 2019 submission to the UN Committee on the Rights of Persons with Disabilities, the Commission noted the lack of measures to ensure nationally consistent implementation, enforceability, monitoring and compliance under the Building Standards and Transport Standards. [1208]

The Commission therefore recommended the Australian Government introduce a national data collection and reporting framework, coordinated between the federal, state and territory governments, to enable the measurement of progress and compliance against these standards.[1209]

Another Commission project, *Free and Equal: An Australian conversation on human rights*,[1210] aims to identify current limitations and barriers to better human rights protections, and builds consensus on actions to promote, protect and fulfil human rights.

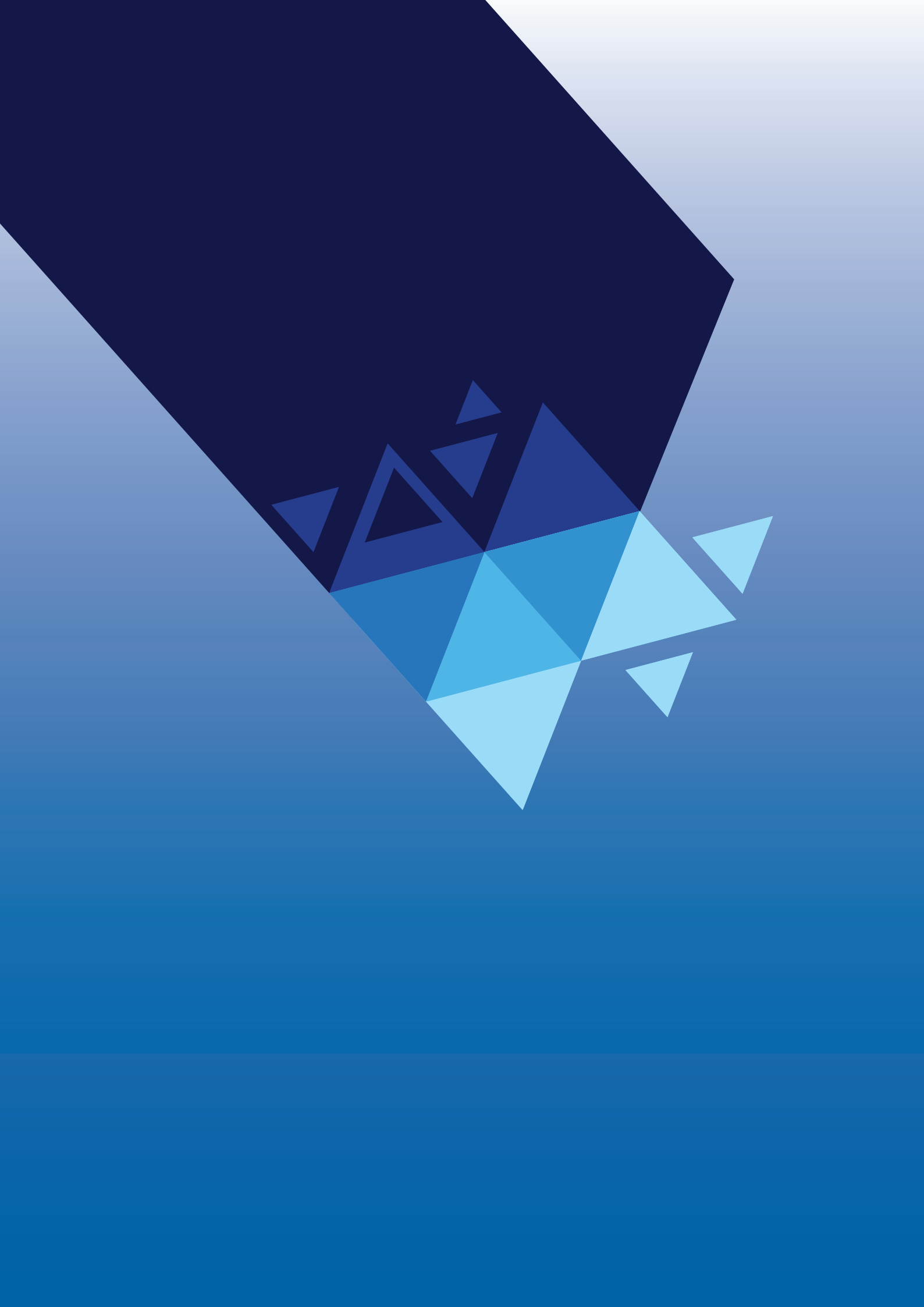
Among other things, that project is reviewing federal anti-discrimination law.[1211] Some of the systemic issues raised in this Project concerning the DDA and anti-discrimination legislation more broadly are being considered in that Commission-wide process.

Two of the priorities presented in the federal anti-discrimination law reform process are to: promote compliance and provide clarity about legal obligations; and to ensure the complaint handling process is accessible to the most vulnerable in the community.[1212]

Concerning the first priority, the Commission notes that current compliance measures under discrimination law (including the existing Disability Standards) have had variable outcomes. The current Standards have been helpfully used to convert general legal principles into measurable, outcome-focused requirements. However, there is a need for greater awareness-raising activities, industry support to promote compliance, and robust review processes to assess measures of compliance.[1213]

Secondly, the Commission has proposed that a priority for federal discrimination law reform should be to ensure that those most vulnerable in the community have access to the complaint-handling function of the Commission. There is scope for a greater emphasis on systemic discrimination and placing less pressure in individual complainants, as well as a review of the financial costs associated with court cases where the complaint is not resolved at the Commission level. This point was raised by some stakeholders in this Project.[1214]

The Commission considers that a proposed Digital Communication Technology Standard is more likely to achieve its aims if there are additional measures that can assist people and organisations to understand their responsibilities under the law and to provide increased certainty to them when seeking to comply.[1215] This could include a mix of mechanisms within an overall compliance framework.[1216] These reforms will be considered in the Free and Equal process, and be informed by stakeholder feedback gathered throughout this Project and the broader Free and Equal consultation process.



PART E: CONSULTATION

CHAPTER 12
Proposals and questions

Part A: Introduction and framework

Proposal 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.

Proposal 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.

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?

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.

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

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.

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.

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.

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.

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?

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.

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.

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?

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

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.

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

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.

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.

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?

Proposal 15: The Australian Government should consider establishing a regulatory sandbox to test AI-informed decision-making systems for compliance with 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 eg, 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 eg, transparency and protection of trade secrets?

(h) how should the regulatory sandbox be evaluated?

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 datapoints and professionals designing and developing AI-informed decision-making systems.

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.

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.

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.

Part D: Accessible technology

Proposal 20: Federal, state, territory and local governments should commit to using Digital Technology that complies with recognised accessibility standards, currently WCAG 2.1 and Australian Standard EN 301 549, and successor standards. To this end, all Australian governments should:

(a) Adopt an accessible procurement policy, promoting the procurement of goods, services and facilities that use Digital Technology in a way that meets the above accessibility standards. Such a policy would also favour government procurement from entities that implement such accessibility standards in their own activities.

(b) Develop policies that increase the availability of accessible communication services such as Easy English versions and human customer supports.

Proposal 21: The Australian Government should conduct an inquiry into compliance by industry with accessibility standards such as WCAG 2.1 and Australian Standard EN 301 549. Incentives for compliance with standards could include changes relating to taxation, grants and procurement, research and design, and the promotion of good practices by industry.

Proposal 22: The Australian Government should amend the *Broadcasting Services Act 1992* (Cth) to require national broadcasting services, commercial broadcasting services, and subscription broadcasting services to:

(a) audio describe content for a minimum of 14 hours per week for each channel, with annual increases

(b) increase the minimum weekly hours of captioned content on an annual basis.

Proposal 23: Standards Australia should develop an Australian Standard or Technical Specification that covers the provision of accessible information, instructional and training materials to accompany consumer goods, in consultation with people with disability and other interested parties.

Proposal 24: The National Broadband Network should undertake economic modelling for the provision of a concessional wholesale broadband rate for people with disability who are financially vulnerable.

Question G: What other measures could the private sector take to eliminate barriers to accessibility related to the affordability of Digital Technologies for people with disability?

Proposal 25: The Council of Australian Governments Disability Reform Council should:

(a) lead a process for Australia's federal, state and territory governments to commit to adopting and promoting 'human rights by design' in the development and delivery of government services using Digital Technologies, and monitor progress in achieving this aim

(b) include policy action to improve access to digital and other technologies for people with disability as a priority in the next National Disability Strategy.

Proposal 26: Providers of tertiary and vocational education should include the principles of 'human rights by design' in relevant degree and other courses in science, technology and engineering. With appropriate support, the Australian Council of Learned Academies should undertake consultation on how to achieve this aim most effectively and appropriately within the tertiary and vocational sector.

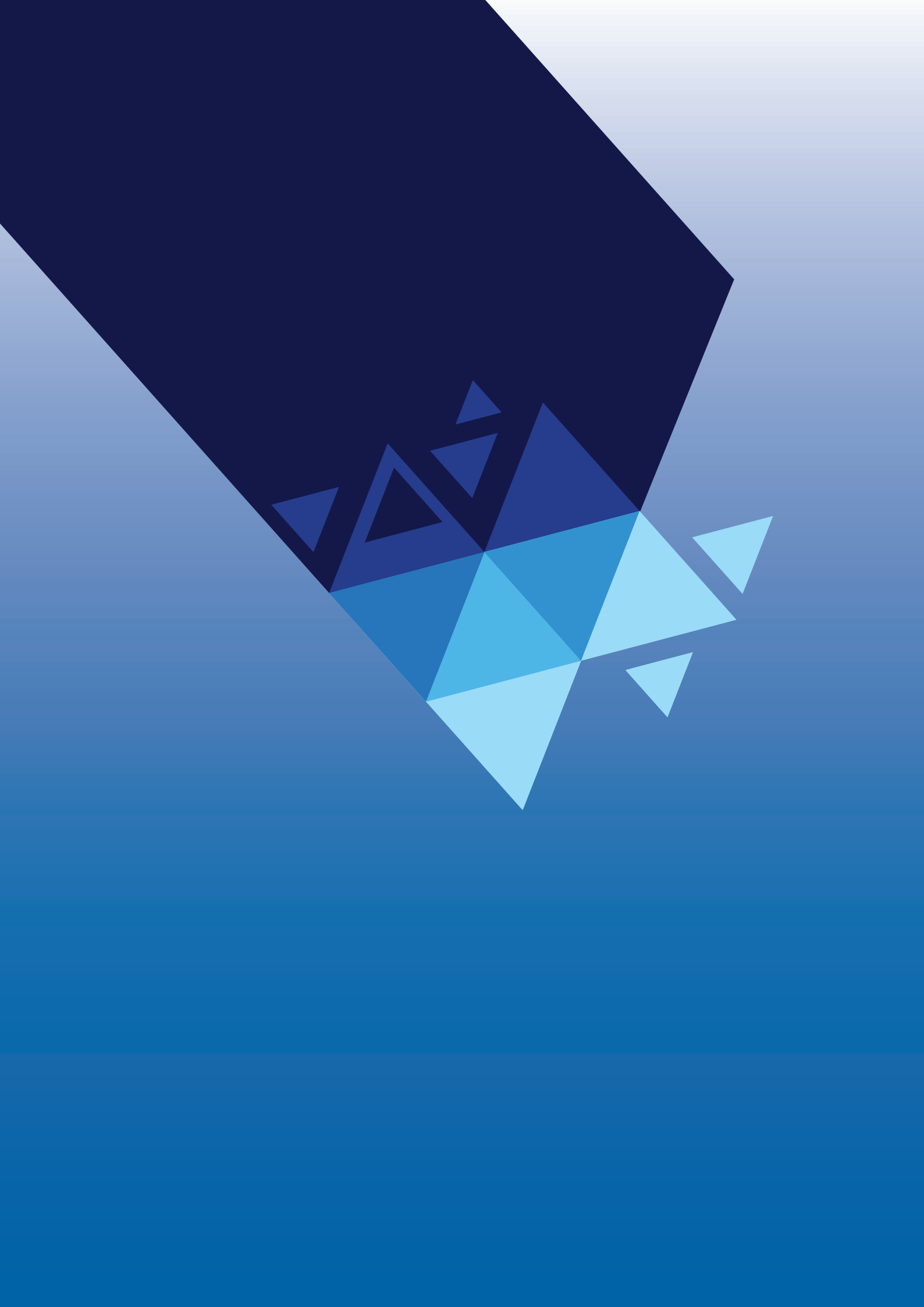
Question H: What other tertiary or vocational courses, if any, should include instruction on 'human rights by design'?

Proposal 27: Professional accreditation bodies for engineering, science and technology should consider introducing mandatory training on 'human rights by design' as part of continuing professional development.

Proposal 28: The Australian Government should commission an organisation to lead the national development and delivery of education, training, accreditation, and capacity building for accessible technology for people with disability.

Proposal 29: The Attorney-General of Australia should develop a Digital Communication Technology Standard under section 31 of the *Disability Discrimination Act 1992* (Cth). In developing this new Standard, the Attorney-General should consult widely, especially with people with disability and the technology sector. The proposed Standard should apply to the provision of publicly available goods, services and facilities that are primarily used for communication, including those that employ Digital Technologies such as information communication technology, virtual reality and augmented reality.

Question I: Should the Australian Government develop other types of Standards, for Digital Technologies, under the *Disability Discrimination Act 1992* (Cth)? If so, what should they cover?



CHAPTER 13

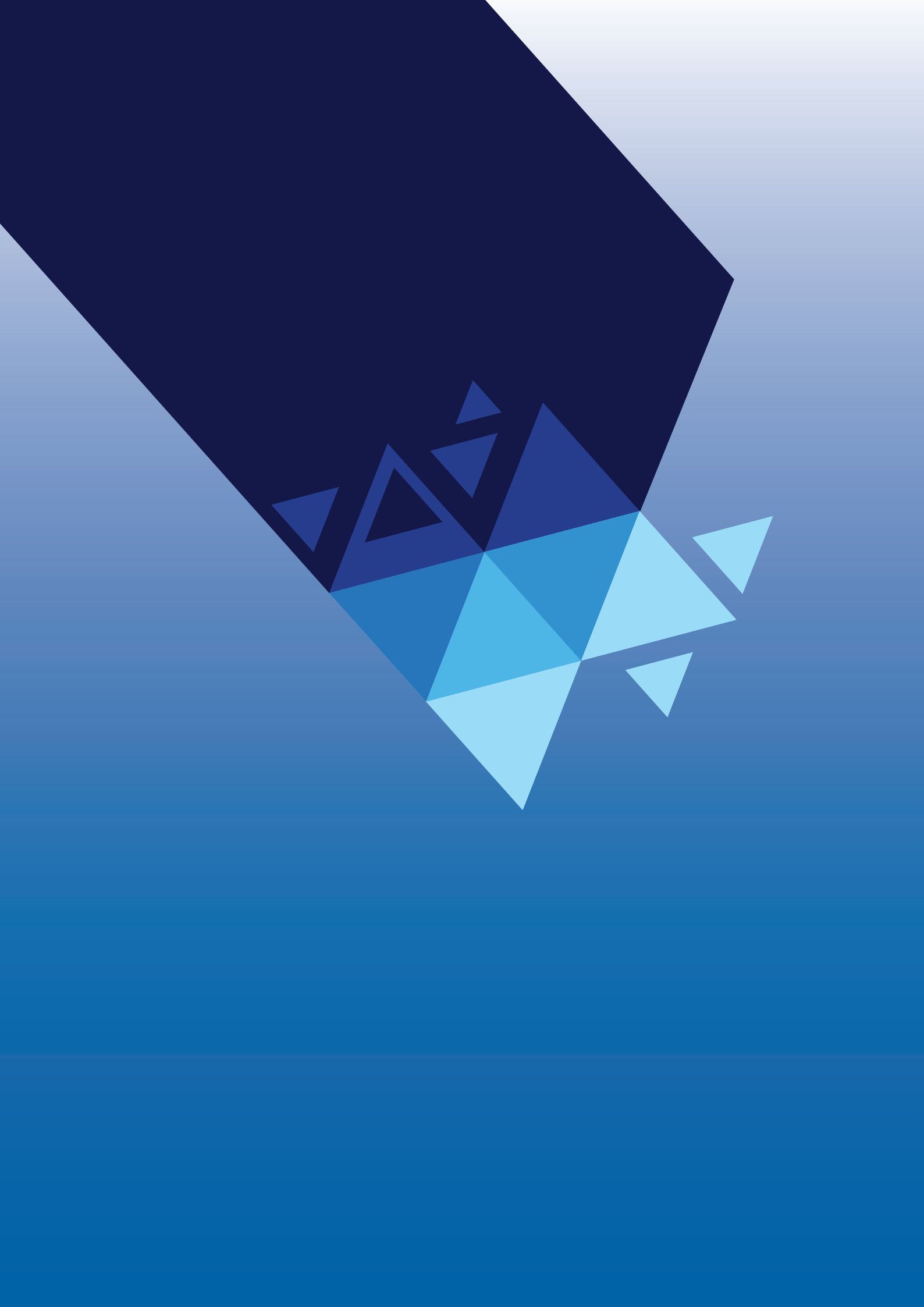
Making a submission

The Commission would like to hear your views on the proposals and questions in this Discussion Paper.

Written submissions may be formal or informal. They can address some or all the consultation questions and proposals. The information collected through the consultation process may be drawn upon, quoted or referred to in any Project documentation. You can elect to make your submission public or confidential.

Written submissions must be received by **Tuesday, 10 March 2020**. The submission form and details on the submission process can be found at https://tech.humanrights.gov.au

To contact the Human Rights and Technology Project team please email tech@humanrights.gov.au or phone (02) 9284 9600 or TTY 1800 620 241.



Appendix A – List of submissions

A. Submissions – Issues Paper

The Commission received 14 confidential submissions to the Issues Paper.

Submission No	Full name
1	Ryan Bryer
4	Kayleen Manwaring
6	Access Now
7	Paul Harpur
8	Melville Miranda
9	Dietitians Association of Australia
10	Dan Svantesson
11	Lisa Fowkes
12	Commonwealth Ombudsman
13	Australian Library and Information Association
14	Royal Australian and New Zealand College of Radiologists
15	Public Interest Advocacy Centre
19	Martha Browning, Megan Ellis, Kelly Yeoh with Tania Leiman
20	Rafael Calvo, Julian Huppert, Dorian Peters, Gerard Goggin
21	Izerobzero
22	Diarmaid Harkin and Adam Molnar
23	Chartered Accountants Australia & New Zealand
24	Marcus Wigan
25	Adam Johnston
26	Pymetrics
27	Australian Red Cross
28	Portable
29	Australian Privacy Foundation, the Queensland Council for Civil Liberties, and Electronic Frontiers Australia
30	The Australian Rehabilitation and Assistive Technology Association
31	Google
32	Uniting Church in Australia, Synod of Victoria and Tasmania
33	Moira Paterson
34	Australian Computer Society
35	Crighton Nichols
36	Office of the Victorian Information Commissioner
37	Mark Dean and Miguel Vatter

Submission No	Full name
39	CHOICE
41	Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University
42	Consumer Policy Research Centre
43	Andrew Normand
44	Australian Centre for Health Law Research
45	Caitlin Curtis, Marie Mangelsdorf, and James Hereward
46	Nicola Henry
47	Ability Research Centre
48	National Disability Services
49	Online Hate Prevention Institute
50	Sean Murphy
51	Deloitte
52	Financial Rights Legal Centre
53	John Corker
54	Normann Witzleb
55	Australian Women Against Violence Alliance
56	Scott McKeown
58	NSW Young Lawyers Communications, Entertainment and Technology Law Committee
59	Sumathy Ramesh
60	The Warren Centre
61	Paul Henman
62	Speech Pathology Australia
63	Roger Clarke
64	Katalin Fenyo
65	Maria O'Sullivan
66	Data Synergies
67	Domestic Violence Victoria
68	La Trobe LawTech
69	Australian Communications Consumer Action Network
70	The Allens Hub for Technology, Law and Innovation
71	Australian Council of Learned Academies
72	Michael Wildenauer
73	Intellectual Disability Rights Service
75	Pip Butt
76	Errol Fries

Submission No	Full name
77	ANZ
78	Carers Australia
79	University of Melbourne
80	Feral Arts/Arts Front
81	Digital Gap Initiative
82	Castan Centre for Human Rights Law, Monash University
83	Global Partners Digital
84	Northraine
85	ThinkPlace
86	EMR Australia
87	Marco Rizzi and David Glance
89	Kingsford Legal Centre
90	Shelley Bielefeld
91	Nicolas Suzor, Kim Weatherall, Angela Daly, Ariadne Vromen, Monique Mann
92	The Centre for Inclusive Design
93	The Ethics Centre
95	Legal Aid New South Wales
97	PwC Indigenous Consulting
98	Office of the Australian Information Commissioner
99	National Association of Community Legal Centres
100	Ron McLay
101	Law Council of Australia
103	University of Technology Sydney
104	Eleanor Salazar, Jerwin Parker Roberto, Leah Gelman, Angelo Gajo
105	Emily Mundzic, Llewellyn Thomas, Eleanor Salazar, Jerwin Parker Roberto
106	Dorotea Baljevic and Ksenija Nikandrova
107	Kate Mathews-Hunt
108	Jobs Australia
109	Ian Law
110	Migrant Worker Justice Initiative, UNSW, UTS
111	LexisNexis
112	Emma Jane and Nicole Vincent
113	Adobe
114	WebKeyIT
116	Emergency Services Levy Insurance Monitor

Submission No	Full name
117	Intopia
118	Claire Devenney and Christopher Mills
119	Jane Refshauge

B. Submissions – White Paper

The Commission received 7 confidential submissions to the White Paper.

Submission No	Full name
1	Kate Mathews-Hunt
2	University of Wollongong
4	Roger Clarke
5	Letter from the Commissioner for Children and Young People
7	Marie Johnson
8	Graham Greenleaf, Roger Clarke, David Lindsay
10	Australian Human Rights Institute
11	NTT Communications Cloud Infrastructure Services Inc
12	Robert Chalmers
13	Remi AI
15	Ethics for AI and ADM
16	Stacy Carter, Nehmat Houssami and Wendy Rogers
17	Renee Newman Knake
18	Data2X and the UN Foundation
20	National Archives of Australia
21	Australian Council of Trade Unions
22	Chartered Accountants Australia and New Zealand
23	Henry Dobson
24	FutureLab.Legal
25	Queensland Council for Civil Liberties
26	Electronic Frontiers Australia
27	Australian Services Union
28	Marcus Smith
29	Joylon Ford

Submission No	Full name
30	Michael Guihot and Matthew Rimmer
31	Office of the eSafety Commissioner
32	Interactive Games and Entertainment Association
33	Centre for Policy Futures, University of Queensland
34	Libby Young
35	The Australian Industry Group
36	Northraine
37	Social Innovation Research Institute, Swinburne University of Technology
38	The Allens Hub for Technology, Law and Innovation
39	Office of the Victorian Information Commissioner
40	The Royal Australian and New Zealand College of Radiologists
41	Ruth Lewis
42	Standards Australia
43	Hayden Wilkinson
44	Simon Moore
45	Law Council of Australia
46	Joanne Evans
47	Portable
48	The Montreal AI Ethics Institute
49	Australian Information and Industry Association
50	The University of Melbourne
51	Microsoft
53	Crighton Nichols
54	Julia Powles, Marco Rizzi, Fiona McGaughey, David Glance
56	Izerobzero
57	Access Now
58	Consumer Policy Research Centre
59	Blockchain Assets
60	Effective Altruism ANZ
61	Digital Industry Group Inc
62	Australian Research Data Commons
63	Office of the Australian Information Commissioner

Appendix B – Acronyms

ACCAN	Australian Communications Consumer Action Network
ACCC	Australian Competition and Consumer Commission
ACM	Association of Computing Machinery
ACMA	Australian Communications and Media Authority
ACOLA	Australian Council of Learned Academies
ADII	Australian Digital Inclusion Index
ADJR Act	*Administrative Decisions (Judicial Review) Act 1977* (Cth)
AIA	algorithmic impact assessment
ALRC	Australian Law Reform Commission
AI	artificial intelligence
ANAO	Australian National Audit Office
APRA	Australian Prudential Regulation Authority
AR	augmented reality
ASIC	Australian Securities and Investment Commission
ATO	Australian Taxation Office
BSA	*Broadcasting Services Act 1992* (Cth)
COAG	Council of Australian Governments
CPD	continuing professional development
CRPD	Convention on the Rights of Persons with Disabilities
CSR	corporate social responsibility
DDA	*Disability Discrimination Act 1992* (Cth)
DTA	Digital Transformation Agency
G20	Group of Twenty
G7	Group of Seven
GDPR	General Data Protection Regulation (European Union)
HRIA	human rights impact assessment
ICCPR	International Covenant on Civil and Political Rights
ICESCR	International Covenant on Economic, Social and Cultural Rights
ICT	information and communications technology
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things
IP	Human Rights and Technology Project Issues Paper
ISO	International Standards Organization
ML	machine learning
NDC	National Data Commissioner
NDIA	National Disability Insurance Agency
NDIS	National Disability Insurance Scheme

NGO	non-government organisation
NTS	National Transition Strategy
OAIC	Office of the Australian Information Commissioner
OECD	Organisation for Economic Cooperation and Development
STEM	science, technology, engineering and maths
UN	United Nations
UN Guiding Principles	United Nations Guiding Principles on Business and Human Rights
VR	virtual reality
WCAG	Web Content Accessibility Guidelines
WP	Human Rights and Technology White Paper

1 Klaus Schwab, *The Fourth Industrial Revolution* (World Economic Forum, 2016).
2 Stefan Hajkowicz and Dave Dawson, *Digital Megatrends: A perspective on the coming decade of digital disruption* (CSIRO Data 61, 2017) <https://www.data61.csiro.au/>; Innovation and Science Australia, *Australia 2030: prosperity through innovation*, Australian Government (2017) <https://www.industry.gov.au/>.
3 Australian Human Rights Commission, *Human Rights and Technology Project: Consultation* (24 June 2019) <https://tech.humanrights.gov.au>.
4 Australian Human Rights Commission, *Human Rights and Technology IP* (July 2018); Australian Human Rights Commission and World Economic Forum, *Artificial Intelligence: Governance and Leadership: White Paper* (January 2019): <https://tech.humanrights.gov.au>.
5 See, for example, Australian Government, Department of Industry, Innovation and Science, AI Ethics Principles (November 2019), <https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework>; Hajkowicz SA, Karimi S, Wark T, Chen C, Evans M, Rens N, Dawson D, Charlton A, Brennan T, Moffatt C, Srikumar S, Tong KJ (CSIRO Data 61) *Artificial Intelligence: Solving problems, growing the economy and improving our quality of life* (2019) <https://data61.csiro.au/en/Our-Research/Our-Work/AI-Roadmap>; Australian Government, Department of Home Affairs, Australian Signals Directorate, Australian Cyber Security Centre, *Draft Code of Practice: Securing the Internet of Things for Consumers* (November 2019) <https://www.homeaffairs.gov.au/reports-and-pubs/files/code-of-practice.pdf>.
6 United Nations Secretary General, *UN Secretary-General's Strategy on New Technologies* (September 2018) <https://www.un.org/en/newtechnologies/images/pdf/SGs-Strategy-on-New-Technologies.pdf>; Council of Europe Commissioner for Human Rights, 'Unboxing Artificial Intelligence: 10 Steps to Protect Human Rights' (Council of Europe, May 2019); OECD, 'OECD Principles on Artificial Intelligence' (22 May 2019) <https://www.oecd.org/going-digital/ai/principles/>; OECD Council Recommendations on Artificial Intelligence, OECD/Legal/0449 (adopted 22 May 2019).
7 See, for example domestic and international inquiry and reform processes in the United States, Denmark, Estonia, and the United Kingdom.
8 OECD, *OECD Principles on Artificial Intelligence* (22 May 2019) <https://www.oecd.org/going-digital/ai/principles/>; OECD Council Recommendations on Artificial Intelligence, OECD/Legal/0449 (adopted 22 May 2019).
9 OECD Council Recommendations on Artificial Intelligence, OECD/Legal/0449 (adopted 22 May 2019).
10 The Human Rights and Technology Expert Reference Group meets quarterly to discuss the work of the Project and forms an important part of Project governance. Members of the ERG and their biographies available at <https://tech.humanrights.gov.au/our-work>.
11 Written submissions responding to the IP were opened on 24 July 2019 and closed on 2 October 2019.
12 Stanford Center on Democracy, Development and the Rule of Law, 'The Future of Human Centered AI: Governance Innovations and Protection of Human Rights' (Conference, Stanford University, 16 April 2019) <https://cddrl.fsi.stanford.edu/content/human-centered-ai>; 'G7 Multistakeholder Conference on Artificial Intelligence: Final Summary Report' (Government of Canada, 6 December 2018) <http://www.ic.gc.ca/eic/site/133.nsf/eng/00007.html>.
13 See, for example, Australian Parliament Select Committee on the Future of Work and Workers, *Hope is not a strategy – our shared responsibility for the future of workers* (Report, 19 September 2018); Committee for Economic Development of Australia, *Australia's Future Workforce?* (June 2015) <https://www.ceda.com.au/Research-and-policy/All-CEDA-research/Research-catalogue/Australia-s-future-workforce>. This was raised in IP submissions, see for example: R McLay; 7; LexisNexis, 5 ; Feral Arts/Arts Front, 13; S McKeown, 1.
14 Related issues will be addressed to some extent in chapters related to AI. The issue was raised in IP submissions, see, for example: Law Council of Australia; University of Technology Sydney, WebKeyIT; Australian Privacy Foundation, the Queensland Council for Civil Liberties, and Electronic Frontiers Australia; Australian Rehabilitation and Assistive Technology Association; M Paterson; C Nichols; K Manwaring; S Murphy; Financial Rights Legal Centre; N Witzleb; NSW Young Lawyers Communications, Entertainment and Technology Law Committee; Access Now; R Clarke; M Rizzi, D Glance; D Baljevic, K Nikandrova, 4.
15 Australian Digital Inclusion Index (2019) <https://digitalinclusionindex.org.au/the-index-report/report/>.
16 See IP submissions: Kingsford Legal Centre, 9; Legal Aid NSW, 4; M Browning, M Ellis, K Yeoh, with T Leiman 5; M Paterson, Domestic Violence Victoria, 1; Online Hate Prevention Institute, 2.
17 Office of the eSafety Commissioner, submission to the WP, 2.
18 University of Technology Sydney, submission to the IP, 46.
19 See ACCC, *Digital Platforms Inquiry*, referred on 4 December 2017 <https://www.accc.gov.au/focus-areas/inquiries/digital-platforms-inquiry>; *Report of the Australian Taskforce to combat terrorist and extreme violent material online* (30 June 2019) <https://www.pmc.gov.au/resource-centre/national-security/report-australian-taskforce-combat-terrorist-and-extreme-violent-material-online>. In the United States, see, for example, Senate Subcommittee on Communications, Technology Innovation and the Internet, 'Optimizing for Engagement: Understanding the Use of Persuasive Technology on Internet Platforms' (Report, 25 June 2019) <https://www.commerce.senate.gov/2019/6/optimizing-for-engagement-understanding-the-use-of-persuasive-technology-on-internet-platforms>.
20 United Nations, *Charter of the United Nations,* 1 UNTS XVI (24 October 1945).
21 *Universal Declaration on Human Rights,* GA Res 217A(III), UN GAOR, 3rd sess, 183rd plen mtg, UN Doc A/810 (10 December 1948) art 1 (UDHR). In the identical Preambles of the *International Covenant of Civil and Political Rights* and the *International Covenant on Economic, Social and Cultural Rights*, it is stated that the rights to the treaties seek to protect 'derive from the inherent dignity of the human person'.
22 See, for example, Genevieve Bell, 'Boyer Lectures: What does it mean to be human in a digital world?' (delivered on 5 October 2017).
23 See, for example, IP submissions: Castan Centre for Human Rights Law; Law Council of Australia; CHOICE.
24 See, for example, *The Toronto Declaration on Protecting the Right to Equality and Non-Discrimination in Machine Learning Systems* (2018) <https://www.accessnow.org/>; OECD, *OECD Council Recommendations on Artificial Intelligence*, OECD/Legal/0449, adopted on 22 May 2019.
25 United Nations Human Rights Council, 'Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression', 29th Sess, A/HRC/29/32, Agenda Item 3 (22 May 2015); United Nations Office of the High Commissioner for Human Rights, '"Smart Mix" of Measures Needed to Regulate New Technologies – Bachelet' (24 April 2019) <https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=24509>; United Nations General Assembly, *The right to privacy in the digital age*, GA Res 61/167, 68th sess, UN Doc A/RES/68/167 (adopted 21 January 2014).
26 OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/Legal/0449 (adopted on 22 May 2019).
27 *G20 Ministerial Statement on Trade and the Economy, G20 2019 Japan* (30 July 2019) <https://g20trade-digital.go.jp/dl/Ministerial_Statement_on_Trade_and_Digital_Economy.pdf>.
28 Government of Canada, 'Canada-France Statement on Artificial Intelligence' on *Government of Canada* (7 June 2018) <http://international.gc.ca/world-monde/international_relations-relations_internationales/europe/2018-06-07-france_ai-ia_france.aspx?lang=eng>.
29 Microsoft, *Human Rights Statement – Microsoft Corporate Social Responsibility* <https://www.microsoft.com/en-us/corporate-responsibility/human-rights-statement>.
30 UN General Assembly, *Vienna Declaration and Programme of Action*, A/CONF.157/23 (12 July 1993) para 5.

31 Human rights in this table are contained in: *International Covenant on Civil and Political Rights*, opened for signature 16 December 1966, 999 UNTS 171 (entered into force 23 March 1976) (ICCPR); *International Covenant on Economic, Social and Cultural Rights*, opened for signature 16 December 1966, 993 (entered into force 3 January 1976) (ICESCR), *Convention on the Elimination of All Forms of Discrimination against Women*, opened for signature 18 December 1979, 189 UNTS 1429 (entered into force 3 September 1981) (CEDAW), *Convention on the Rights of the Child,* opened for signature 20 November 1989, E/CN.4/RES/1990/74 (entered into force 2 September 1990) (CRC), *International Convention on the Elimination of All Forms of Racial Discrimination,* opened for signature 21 December 1965, 660 UNTS 195 (entered into force 4 January 1969) (CERD), *Convention on the Rights of Persons with Disabilities,* opened for signature 13 December 2006, 2515 UNTS 3 (entered into force 3 May 2008) (CRPD).
32 UN General Assembly, *Vienna Declaration and Programme of Action*, A/CONF.157/23 (12 July 1993) 1.
33 The High Court of Australia has said the freedom of political communication is implied from the system of representative and responsible government set up by the terms and structure of the Australian Constitution. See, eg, *Lange v Australian Broadcasting Corporation* (1997) 189 CLR 520*, McCloy v New South Wales* [2015] HCA 34; (2015) 89 ALJR 857; and, most recently, *Brown v Tasmania* [2017] HCA 43.
34 *Charter of Human Rights and Responsibilities Act 2006* (VIC); *Human Rights Act 2019* (QLD); *Human Rights Act 2004* (ACT).
35 Australian Law Reform Commission, *Traditional Rights and Freedoms—Encroachments by Commonwealth Laws* (Report No 129, December 2015), ch 2.
36 Australian Human Rights Commission, 'Common law rights, human rights scrutiny and the rule of law' (19 July 2013) <https://www.humanrights.gov.au/our-work/rights-and-freedoms/common-law-rights-human-rights-scrutiny-and-rule-law>.
37 *Australian Human Rights Commission Act 1986* (Cth).
38 *Human Rights (Parliamentary Scrutiny) Act 2011* (Cth).
39 See Australian Human Rights Commission, 'Permissible limitations on rights' (6 June 2013) <https://www.humanrights.gov.au/our-work/rights-and-freedoms/permissible-limitations-rights#ICCPR>.
40 See, for example, *Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment*, opened for signature 10 December 1984, 1465 UNTS 85 (entered into force 26 June 1987); ICCPR arts 6,7, 8, 11,15, 16.
41 United Nations Economic and Social Council, *Siracusa Principles on the Limitation and Derogation Provisions in the International Covenant on Civil and Political Rights* UN Doc E/CN.4/1985/4, Annex (1985) (ICCPR); United Nations Human Rights Committee, *General Comment No. 31 [80]: The Nature of the General Legal Obligation Imposed on States Parties to the Covenant*, UN Doc CCPR/C/21/Rev.1/Add. 13 (26 May 2004) 6.
42 ICCPR, art 19(3).
43 ICCPR, art 4. For an explanation on non-derogable and absolute rights, see Attorney-General's Department, 'Absolute rights' <https://www.ag.gov.au>.
44 ICCPR, art 12.
45 ICCPR, art 9.
46 ICCPR, art 4.
47 See, for example, *Catt v the United Kingdom* (European Court of Human Rights, Application No 43514/15, 24 January 2019); *Big Brother Watch and Ors v the United Kingdom* (European Court of Human Rights, Application Nos 58170/30, 62322/14, and 24960/15, 13 September 2018); *Google Spain SL v Gonzalez* (Court of Justice of the European Union, C-131/12, ECLI:EU:C:2014:317, 13 May 2014) [80].
48 *Catt v the United Kingdom* [2019] European Court of Human Rights Application No 43514/15, (24 January 2019) 108.
49 *Catt v the United Kingdom* [2019] European Court of Human Rights Application No 43514/15, (24 January 2019) 120.
50 *Catt v the United Kingdom* [2019] European Court of Human Rights Application No 43514/15, (24 January 2019) 128.
51 N Suzor, K Weatherall, A Daly, A Vromen, M Mann, submission to the IP, 2.
52 See, for example, Eleonore Pauwels, 'The New Geopolitics of Converging Risks: The UN and Prevention in the Era of AI' (United Nations University Centre for Policy Research, 29 April 2019) <https://i.unu.edu/media/cpr.unu.edu/attachment/3472/PauwelsAIGeopolitics.pdf>.
53 See, for example, PwC Indigenous Consulting, submission to the IP, 21.
54 Noted in IP submissions: Red Cross, 4; Australian Computer Society, 1, 2; Public Interest Advocacy Centre, 8.
55 ICCPR, art 19.
56 ICCPR, arts 21-22.
57 Pymetrics, submission to the IP, 4.
58 ICCPR, art 2. See also: CEDAW, arts 11 and 14(2)(e); CERD, arts 5(e)(i) and (ii); CRPD, art 27.
59 ICESCR arts 6(1), 7 and 8(1)(a).
60 Wenya Linda Bi et al 'Artificial Intelligence in Cancer Imaging: Clinical Challenges and Applications' [2019] *CA: A Cancer Journal for Clinicians* Vol 69(2).
61 PwC Indigenous Consulting, submission to the IP, 15.
62 Maged Kamel Boulos, Guochao Peng and Trang VoPham, 'An Overview of GeoAI Applications in Health and Healthcare' (2019) 18(1) *International Journal of Health Geographics,* 7.
63 ICESCR art 12; Committee on Economic, Social, and Cultural Rights, *General Comment No. 14: The Right to the Highest Attainable Standard of Health,* 11 August 2000, 22nd sess E/C.12/2000/4, para 1.
64 Committee on Economic, Social, and Cultural Rights, *General Comment No. 14: The Right to the Highest Attainable Standard of Health,* 11 August 2000, 22nd sess E/C.12/2000/4, para 13.
65 See, especially, *United Nations Declaration on the Rights of Indigenous Peoples*, GA Res 61/295, UN GAOR, 61st sess, 107th plen mtg, Agenda Item 68, Supp No 49, UN Doc A/RES/61/295, Annex, (2 October 2007) art 21(1); see also arts 23, 24 and 29.
66 See, for example, United Nations Committee on the Rights of Persons with Disabilities, *General Comment No.2 (2014) on Accessibility,* UN Doc CRPD/C/GC/2 (22 May 2014) para 9. Correspondence from the National Disability Insurance Scheme, for example, has been sent in incompatible formats to people with vision impairment, Luke Henrique-Gomes 'NDIS sent letters to vision impaired and blind people in format they could not read' The Guardian (29 June 2019) <https://www.theguardian.com/australia-news/2019/jun/29/ndis-sent-letters-to-vision-impaired-and-blind-people-in-format-they-could-not-read>.
67 CRPD, arts 9 and 19.
68 IP submissions identified the danger of AI technologies, particularly automated decision-making systems to engage the right to non-discrimination; see, for example: Intopia, 5; Castan Centre for Human Rights Law, 8, 9; Australian Centre for Health Law Research, 7.
69 Jeffrey Dastin, 'Amazon scraps secret AI recruiting tool that showed bias against women' *Reuters* (10 October 2018) <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>.
70 Sam Levin, 'Facebook cracks down on discriminatory ads after years of backlash' *The Guardian* (20 March 2019) <https://www.theguardian.com/technology/2019/mar/19/facebook-advertising-discrimination-lawsuit-aclu-race-gender>.
71 These issues were raised in IP submissions made to the IP, including: University of Technology Sydney, 46; National Association for Community Legal Centres, 7; Australian Privacy Foundation, Queensland Council for Civil Liberties, Electronic Frontiers Australia, 19; PwC Indigenous Consulting, 25. See also, Raffaele Piccolo, 'AI in Criminal Sentencing: A Risk to Our Human Rights?' (2018) 40(11) *The Bulletin* 3, 17.
72 See, for example, Virginia Eubanks, *Automated Inequality* (St Martin's Publishing Group, 2018).

73 See, for example, Mark Latonero, 'Governing Artificial Intelligence: Upholding Human Rights & Dignity' (Data&Society, 2018); Mark Hodge and Dunstan Allison-Hope, 'Artificial Intelligence: A Rights-Based Blueprint for Business,' (Working Paper No 2, BSR, 2018); Filippo Raso et al, 'Artificial Intelligence & Human Rights: Opportunities & Risks' (Berkman Klein Center Research Publication No 2018-6). Petra Molnar and Lex Gill, *Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System* (University of Toronto, 2018), 13.

74 See, for example, Filippo Raso et al, 'Artificial Intelligence & Human Rights: Opportunities & Risks' (Berkman Klein Center Research Publication No 2018-6); Eileen Donahoe and Megan MacDuffee Metzger, 'Artificial Intelligence and Human Rights' (2019) 30(2) *Journal of Democracy* 115.

75 See, for example, IEEE, 'Ethically Aligned Design' (2018) <https://ethicsinaction.ieee.org/>.

76 In jurisdictions where there is human rights legislation, individuals have begun to challenge the use of a specific technology on the basis that it impermissibly breaches their human rights. See, for example, the challenge brought against the use of facial recognition technology by police in Wales Liberty Human Rights, 'Stand up to Police Spying' (14 February 2019) <https://www.libertyhumanrights.org.uk/our-campaigns/stand-police-spying>; Hannah Couchman, 'Policing by Machine: Predictive Policing and the Threat to Our Rights' (Text, Liberty Human Rights, January 2019) <https://www.libertyhumanrights.org.uk/policy/report-policing-machine>.

77 This approach reflects a common way of applying a human rights approach via the PANEL Principles of participation, accountability, non-discrimination and equality, empowerment and legality. See, for example, Australian Human Rights Commission, *Human Rights Based Approaches* <https://www.humanrights.gov.au/our-work/rights-and-freedoms/human-rights-based-approaches>; Scottish Human Rights Commission, *A Human Rights Based Approach: An Introduction* <http://www.scottishhumanrights.com/rights-in-practice/human-rights-based-approach/>.

78 See, for example, IP submissions: Australian Centre for Health Law Research, 2; Kingsford Legal Centre; Castan Centre for Human Rights Law, 5; University of Melbourne, 7; La Trobe LawTech, 8; E Jane, N Vincent, 44; University of Technology Sydney, 8; Royal Australian and New Zealand College of Radiologists, 5; Australian Centre for Health Law Research, 44; National Association of Community Legal Centres, 9; Australian Privacy Foundation, Queensland Council for Civil Liberties, Electronic Frontiers Australia, 30; CHOICE, 2; Society on Social Impacts of Technology, IEEE and Law Futures Centre Griffith University, 6; Deloitte, 5; NSW Young Lawyers Communications, Entertainment and Technology Law Committee, 10.

79 Australian Centre for Health Law Research, submission to the IP, 2.

80 Kingsford Legal Centre, submission to the IP, 2.

81 Castan Centre for Human Rights Law, submission to the IP, 6.

82 University of Technology Sydney, submission to the IP, 8.

83 See, for example, IP submissions: Australian Privacy Foundation, Queensland Council for Civil Liberties, Electronic Frontiers Australia, 12; University of Technology Sydney, 27; Access Now, 4.

84 Josh Bavas, 'Facial recognition quietly switched on at Queensland stadiums, sparking privacy concerns,' *ABC News* (6 June 2019) <https://www.abc.net.au/news/2019-06-05/facial-recognition-quietly-switched-on-at-queensland-stadiums/11178334>; Josh Bavas, 'The Facial Recognition Security Issue Police Tried to Keep Secret' *ABC News* (6 May 2019) <https://www.abc.net.au/news/2019-05-06/australias-biggest-facial-recognition-roll-out-rushed/11077350>; The Hon Lisa Neville MP, Acting Premier, Minister for Policy and Emergency Services, Minister for Water, 'Media Release: New Police Eye in the Sky to Keep Victorians Safe' (10 July 2019) <https://www.premier.vic.gov.au/new-police-eye-in-the-sky-to-keep-victorians-safe/>.

85 See, for example, Bernard Marr, 'Chinese Social Credit Score: Utopian Big Data Bliss or Black Mirror on Steroids?' *Forbes* (21 January 2019) <https://www.forbes.com/sites/bernardmarr/2019/01/21/chinese-social-credit-score-utopian-big-data-bliss-or-black-mirror-on-steroids/#65f4bfe748b8>.

86 Clause 14 of the Identity-Matching Services Bill 2019 provides for the development, operation and maintenance of an 'interoperability hub', that supports electronic communications for the purpose of requesting and providing identity-matching services.

87 See the Explanatory Memorandum for the Identity-Matching Services Bill 2019, at [16].

88 See Australian Human Rights Commission, Submission No 11 to the Parliamentary Joint Committee on Intelligence and Security, *Review of the Identity-Matching Services Bill 2018 and the Australian Passports Amendment (Identity-Matching Services) Bill 2018* (29 March 2018), <https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/Intelligence_and_Security/Identity-Matching2019/Submissions>.

89 Parliamentary Joint Committee on Intelligence and Security, 'Advisory report on the Identity-matching Services Bill 2019 and the Australian Passports Amendment (Identity-matching Services) Bill 2019' (2019), <https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/Intelligence_and_Security/Identity-Matching2019/Report>.

90 Australian Law Reform Commission, *For Your Information: Privacy Law and Practice* (Report 108, May 2008), 406 [9.64], <https://www.alrc.gov.au/publication/for-your-information-australian-privacy-law-and-practice-alrc-report-108/>.

91 Council of Europe Article 29 Data Protection Working Party, *Opinion 3/2012 on developments in biometric technologies,* Doc 00720/12/EN WP193 (27 April 2012), 20-21.

92 Peter Fussey and Daragh Murray, 'Independent Report on the London Metropolitan Police Service's Trial of Live Facial Recognition Technology' (Human Rights Centre, University of Essex, July 2019), 36 <https://www.essex.ac.uk/news/2019/07/03/met-police-live-facial-recognition-trial-concerns>.

93 UN Special Rapporteur on the Right to Privacy, *Right to Privacy* (Advanced Unedited Version) (27 February 2019); Human Rights Council, 40th sess, 25 February to 22 March 2019, Agenda item 3, 3. Also see IP submissions: Australian Privacy Foundation, Queensland Council for Civil Liberties, Electronic Frontiers Australia, 9; Law Council of Australia, 20; University of Technology Sydney, 27; Access Now; M Dean and M Vatter, 2; Red Cross, 5. See also the Telecommunications and Other Legislation Amendment (Assistance and Access) Bill 2018; Australian Human Rights Commission, Submission No 47 to the Parliamentary Joint Committee on Intelligence and Security, *Review of the Telecommunications and Other Legislation Amendment (Assistance and Access Bill) 2018* (12 October 2018); Access Now, Submission No 33 to the Parliamentary Joint Committee on Intelligence and Security, *Review of the Telecommunications and Other Legislation Amendment (Assistance and Access Bill) 2018* (12 October 2018).

94 See, for example, the United Kingdom and China.

95 See, for example, Maya Wang and Human Rights Watch, *China's Algorithms of Repression: Reverse Engineering a Xinjiang Police Mass Surveillance App* (2019), 15 <https://www.hrw.org/report/2019/05/01/chinas-algorithms-repression/reverse-engineering-xinjiang-police-mass-surveillance>.

96 See, for example, Joy Buolamwini and Timinit Guru, 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification' (2018) 81 *Proceedings of Machine Learning Research* 1-15 <http://proceedings.mlr.press/v81/buolamwini18a.html>.

97 Kate Conger, Richard Fausset and Serge F Kovaleski, 'San Francisco Bans Facial Recognition Technology,' *The New York Times* (14 May 2019) <https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html>.

98 For a discussion on data ownership see Václav Janeček, 'Ownership of Personal Data in the Internet of Things' (2018) 34(5) *Computer Law & Security Review* 1039.

99 IP submissions raised several issues around liability. See, for example, N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 9; M Dean and M Vatter, 5; E Jane and N Vincent, 61; K Mathews-Hunt, 137; Global Partners Digital, 3, 7; Castan Centre for Human Rights Law, 4.

100 United Nations Human Rights Council, *Resolution 17/4 Human rights and transnational corporations and other business enterprises,* UN Doc A/HRC/RES/17/4 (16 June 2011).

101 United Nations Office of the High Commissioner for Human Rights, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework* (2011). Endorsed by the Human Rights Council on 16 June 2011 in Resolution 17/4, <https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr_en.pdf>.
102 *Modern Slavery Act 2018* (Cth); *Modern Slavery Act 2015* (UK).
103 United Nations Office of the High Commissioner for Human Rights, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework* (2011), Principles 1-6.
104 United Nations Office of the High Commissioner for Human Rights, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework* (2011), Principles 11-14.
105 United Nations Office of the High Commissioner for Human Rights, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework* (2011), Principles 15-24.
106 United Nations Office of the High Commissioner for Human Rights, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework* (2011), Principles 29-31.
107 Noted in IP submissions: Law Council of Australia, 24; Google, 2; Royal Australian and New Zealand College of Radiologists, 5; Australian Centre for Health Law Research, 6.3; Data Synergies, 8; Castan Centre for Human Rights Law, 6; Global Partners Digital, 9; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 10; PwC Indigenous Consulting, 6.
108 Law Council of Australia, submission to the IP, 24.
109 See, for example, the Council of Europe's ICT Sector Guide on Implementing the UN Guiding Principles on Business and Human Rights; Karen Yeung, *Committee of Experts on Human Rights Dimensions of Automated Data Processing and Different Forms of Artificial Intelligence* (Council of Europe, 2018), 4.
110 Office of the High Commissioner for Human Rights, '41st session of the Human Rights Council – Opening statement by the UN High Commissioner for Human Rights Michele Bachelet' (24 June 2019).
111 United Nations Human Rights Council, 'Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression', 29th Sess, A/HRC/29/32, Agenda Item 3 (22 May 2015).
112 United Nations Human Rights Council, 'Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression', 29th Sess, A/HRC/29/32, Agenda Item 3 (22 May 2015).
113 See, for example, IP submissions: N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 9; Royal Australian and New Zealand College of Radiologists , 4; K Mathews-Hunt, 41; University of Technology Sydney, 34; Castan Centre for Human Rights Law, 4; D Harkin, A Molnar, 3; M Dean and M Vatter, 5; K Manwaring, 273; N Henry; R McLay, 11; Global Partners Digital, 3.
114 Mark Hodge and Dunstan Allison-Hope, 'Artificial Intelligence: A Rights-Based Blueprint for Business,' (Working Paper No 3, BSR, 2018).
115 See Nick Clegg, 'Charting a Course for an Oversight Board for Content Decisions' *Facebook Newsroom* <https://newsroom.fb.com/news/2019/01/oversight-board/>; Brent Harris, Director of Governance and Global Affairs, 'Global Feedback and Input on the Facebook Oversight Board for Content Decisions, *Facebook Newsroom* (27 June 2019) <https://newsroom.fb.com/news/2019/06/global-feedback-on-oversight-board/>.
116 See, for example, Amy Binns, 'Cambridge Analytica Scandal: Facebook's User Engagement and Trust Decline,' *The Conversation* <http://theconversation.com/cambridge-analytica-scandal-facebooks-user-engagement-and-trust-decline-93814>.
117 See, for example, 'Australia 2030: Prosperity through Innovation' (Australian Government, November 2017) <https://www.industry.gov.au/data-and-publications/australia-2030-prosperity-through-innovation>.
118 See IP submissions: D Svantesson, 4; Red Cross, 9; Warren Centre, 11; University of Technology, Sydney, 36; N Vincent, 65; Castan Centre for Human Rights Law, 4; Office of the Australian Information Commissioner, 4; Royal Australian and New Zealand College of Radiologists, 7; M Browning, M Ellis, K Yeoh, 22; Intopia, 5; D Harkin and A Molar, 3. See WP submissions: Allens Hub for Technology, Law & Innovation, 5; Data2X and the UN Foundation, 1; Access Now, 7; Robert Chalmers, 2; RemiAI, 3; Microsoft, 3.
119 OECD, *OECD Council Recommendations on Artificial Intelligenc*e, OECD/Legal/0449, adopted on 22 May 2019.
120 United Nations Human Rights Committee 'New and emerging digital technologies and human rights', resolution adopted by the Human Rights Council A/HRC/RES/41/11 (11 July 2019).
121 UN Office of the High Commissioner for Human Rights "'Smart mix" of measures needed to regulate new technologies – Bachelet" *Media release* (24 April 2019) <https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=24509&LangID=E>.
122 For further information, see Office of the High Commissioner for Human Rights, 'General Comment on children's rights in relation to the digital environment' (2019) <https://www.ohchr.org/EN/HRBodies/CRC/Pages/GCChildrensRightsRelationDigitalEnvironment.aspx>.
123 Office of the High Commissioner for Human Rights, "'Smart mix" of measures needed to regulate new technologies – Bachelet' (24 April 2019) <https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=24509&LangID=E>.
124 See Principles 1 to 6 of the United Nations General Principles of Business and Human Rights; United Nations Office of the High Commissioner for Human Rights, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework* (2011). Endorsed by the Human Rights Council on 16 June 2011 in Resolution 17/4.
125 See for example, The Allens Hub for Technology, Law, and Innovation, submission to the IP, 3.
126 Civil Aviation Safety Authority, 'The Rules' <https://www.casa.gov.au/drones/rules>.
127 See for example, IP submissions: Castan Centre for Human Rights Law, 4; Law Council of Australia, 22; CHOICE, 18. Submissions to the WP: Queensland Council of Civil Liberties, 2; University of Melbourne, 2.
128 See, for example, IP submissions: Law Council of Australia, 22; Northraine, 2.
129 For example, IP submissions: Digital Gap Initiative, 5; Australian Communications Consumer Action Network, 4.
130 See, for example, IP submissions: M O'Sullivan, 7; Access Now, 3; Castan Centre for Human Rights Law,15. See also: Remi AI, submission to the WP, 4.
131 See for example: University of Melbourne, submission to the WP, 5.
132 See, for example, WP submissions: Law Council of Australia, 4; Social Innovation Research Institute, 2; G Greenleaf, R Clarke, D Lindsay, 2; Access Now, 2; Queensland Council of Civil Liberties, 3; Centre for Policy Futures, 4.
133 Castan Centre for Human Rights Law, submission to the IP, 7.
134 See, for example, IP submissions: Data Synergies, 1; University of Technology Sydney, 15; Royal Australian and New Zealand College of Radiologists, 2, 4. See also: C Nichols, submission to the WP, 6.
135 E Fries, submission to the IP, 1.
136 Australian Productivity Commission, *Data Availability and Use: Productivity Commission Inquiry Report* (No. 82, 31 March 2017).
137 Australian Information and Industry Association, submission to the WP, 2, 5.
138 Data Synergies, submission to the IP, 1.
139 See for example, Data Synergies, submission to the IP, 1.
140 Commonwealth Ombudsman, submission to the WP, 6.
141 Digital Industry Group, submission to the WP, 4, 5.
142 Digital Industry Group, submission to the WP, 5.

143 See Chapter 2. See also, for example, IP submissions: Global Partners Digital, 20; Australian Centre for Health Law Research, 5-6; Kingsford Legal Centre, 13; Castan Centre for Human Rights Law, 5; University of Melbourne, 7; La Trobe LawTech, 8; E Jane, N Vincent, 44; University of Technology Sydney, 8; Royal Australian and New Zealand College of Radiologists, 5; National Association of Community Legal Centres, 9; Australian Privacy Foundation, Queensland Council for Civil Liberties, Electronic Frontiers Australia, 30; CHOICE, 6; Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University, 6; Deloitte, 5; NSW Young Lawyers Communications, Entertainment and Technology Law Committee, 10.
144 University of Technology Sydney, submission to the IP, 35.
145 See, for example, IP Submissions: Access Now, 5; Australian Library and Information Association, 2; PwC Indigenous Consulting, 7; University of Melbourne, submission to the WP, 5.
146 PwC Indigenous Consulting, submission to the IP, 5.
147 PwC Indigenous Consulting, submission to the IP, 32.
148 See, for example: Committee on the Rights of Persons with Disabilities, *General Comment No. 2 (2014) Article 9: Accessibility*, 11th sess, CRPD/C/GC/2 (11 April 2014); Committee on the Rights of Persons with Disabilities, *General Comment No 6 (2018) Article 5: Equality and Non-discrimination*, CPRD/C/GC/6 (26 April 2018); Intopia, submission to the IP, 5.
149 See, for example, IP submissions: LexisNexis, 12; Deloitte, 5,12; Law Council of Australia, 23; University of Technology Sydney, 52.
150 See for example, IP submissions: University of Melbourne, 2, ThinkPlace, 9; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 6; K Matthews-Hunt, 251. See also Microsoft, submission to the WP, 3.
151 See, for example, IP submissions: University of Technology Sydney, 36; R Clarke, 8; M Paterson; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 7. See also M Guihot and M Rimmer, submission to the WP, 27. Braithwaite's Regulatory Pyramid models a hierarchy of regulatory interventions from least interventionist at the base of the pyramid, up to the most interventionist at its peak. The model was first put forward by Braithwaite in J Braithwaite, *To Punish or Persuade: Enforcement of Coal Mine Safety* (SUNY Press,1985).
152 See, for example, IP submissions: University of Technology Sydney, 24; Chartered Accountants Australia and New Zealand, 3.
153 United Nations Office of the High Commissioner for Human Rights, '"Smart mix" of measures needed to regulate new technologies – Bachelet' (24 April 2019), <https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=24509&LangID=E>.
154 See for example N Suzor, K Weatherall, A Daly, A Vromen, M Mann, submission to the IP, 6.
155 N Suzor, K Weatherall, A Daly, A Vromen, M Mann, submission to the IP, 6.
156 Administrative Review Council *Administrative Accountability in business areas subject to complex and specific regulation* (November 2008), 7 <https://www.arc.ag.gov.au/Publications/Reports/Pages/Reportfiles/ReportNo49.aspx>.
157 Administrative Review Council *Administrative Accountability in business areas subject to complex and specific regulation* (November 2008), 9 <https://www.arc.ag.gov.au/Publications/Reports/Pages/Reportfiles/ReportNo49.aspx>.
158 Australian Law Reform Commission, *"For Your Information" Privacy Law and Practice* (Report No 108, 2008), 235 [4.7] <https://www.alrc.gov.au/publications/report-108>.
159 *Privacy Act 1988* (Cth). See also: Office of the Australian Information Commissioner, submission to the IP, 3.
160 Office of the Australian Information Commissioner, 'Australian Entities and the EU General Data Protection Regulation (GDPR)' (8 June 2018) <https://www.oaic.gov.au/privacy/guidance-and-advice/australian-entities-and-the-eu-general-data-protection-regulation/>.
161 See, for example, IP submissions: Castan Centre for Human Rights Law, 15, 18; Office of the Australian Information Commissioner, 1. See also WP submissions: Australian Information and Industry Association, 1; Chartered Accountants Australia and New Zealand, 1; R Chalmers, 1.
162 See, for example, IP submissions: Australian Privacy Foundation, Queensland Council for Civil Liberties, Electronic Frontiers Australia, 5; Kingsford Legal Centre, 3; Office of the Australian Information Commissioner, 4; University of Melbourne, 8, 9; P Harpur, 10; P Henman, 8; M Wildenauer, 2; Castan Centre for Human Rights Law, 28.
163 Julia Black, Martyn Hopper, and Christa Band, 'Making a Success of Principles-Based Regulation' (2007) Vol 1(3) *Law and Financial Markets Review* 191–206; University of Technology Sydney, submission to the IP, 57.
164 N Witzleb, submission to the IP, 7.
165 Harry Armstrong, Chris Gorst, and Jen Rae, 'Renewing regulation 'Anticipatory regulation' in an age of disruption' (Nesta, March 2019), 5, 6 <https://media.nesta.org.uk/documents/Renewing_regulation_v3.pdf>
166 Harry Armstrong, Chris Gorst, and Jen Rae, 'Renewing regulation 'Anticipatory regulation' in an age of disruption' (Nesta, March 2019), 5-6 <https://media.nesta.org.uk/documents/Renewing_regulation_v3.pdf>.
167 Portable, submission to the WP, 1.
168 Australian Securities and Investments Commission, 'Fintech regulatory sandbox' <https://asic.gov.au/for-business/innovation-hub/fintech-regulatory-sandbox/>; Nesta 'Anticipatory regulation' <https://www.nesta.org.uk/feature/innovation-methods/anticipatory-regulation/>.
169 See, for example, the following WP submissions: R Newman Knake, 2; H Wilkinson, 4; Social Innovation Research Institute, 3; Castan Centre for Human Rights Law, submission to the IP, 3.
170 Ryan Hagemann, Jennifer Huddleston Skees, and Adam Thierer, 'Soft law for hard problems: The governance of emerging technologies in an uncertain future' (2018) 17(1) *Colorado Technology Law Journal* 94, 59.
171 Consultation Roundtables; Australian Women Against Violence Alliance, submission to the IP, 9-10.
172 P Harpur, submission to the IP, 42.
173 See, for example, IP submissions: Castan Centre for Human Rights Law, 2; University of Melbourne, 2. See WP submissions: Microsoft; Digital Industry Group, 4. See also Klaus Schwab, *The Fourth Industrial Revolution* (World Economic Forum, 2016).
174 Remi AI, submission to the WP, 4.
175 Digital Gap Initiative, submission to the IP, 5 [15].
176 Microsoft, submission to the WP, 3.
177 NTT Communications Cloud Infrastructure Services, submission to the WP, 5.
178 See, for example the eSafety Commissioner's experience of managing illegal and offensive online content hosted overseas: Office of the eSafety Commissioner, submission to the WP, 4; Digital Industry Group, 5. See also IP submissions: Castan Centre for Human Rights Law, 3; D Svantesson, 4.
179 See, for example, IP submissions: Allens Hub for Technology, Law and Innovation 43; Castan Centre for Human Rights Law, 4; E Fries, 1; Law Council of Australia, 2. See also R Lewis, submission to the WP, 3.
180 Law Council of Australia, submission to the IP, 2.
181 Consumer Policy Research Centre, submission to the WP, 4; Ethics for AI and ADM, submission to the WP.
182 See, for example, IP submissions: N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 5; Kingsford Legal Centre, 6; Allens Hub for Technology, Law and Innovation 8.
183 R Calvo, J Huppert, D Peters, G Goggin, 2; Uniting Church in Australia, Synod of Victoria & Tasmania, 32 [3].
184 See, for example, IP submissions: The Ethics Centre, 38; K Matthew-Hunt, 98; Castan Centre for Human Rights Law, 3. H Wilkinson, submission to the WP, 4.
185 Essential Media, *The Essential Report – Human Rights Commission* (26 July 2019), 3.
186 World Economic Forum and Ipsos, 'Evolving Technology: A Global Survey on Data Privacy/Security and Automation/Artificial Intelligence' 20.
187 *Royal Commission into Misconduct in the Banking, Superannuation and Financial Services Industry* (Final Report, February 2019) vol 1, 2.

188 Global Digital Policy Incubator, 'The Future of Human-Centered AI: Governance Innovation and Protection of Human Rights' (24 April 2019) <https://medium.com/stanfords-gdpi/the-future-of-human-centered-ai-governance-innovation-and-protection-of-human-rights-5c371f195232>
189 *General Data Protection Regulation*, (European Union) 2016/679.
190 Makena Kelly, 'Where do the top presidential candidates stand on breaking up Big Tech,' *The Verge* (14 March 2019) <https://www.theverge.com/2019/3/14/18264252/presidential-candidates-breaking-up-big-tech-stance-warren-sanders-booker-klobuchar>.
191 Matthew Miller, Selina Wang and Sarah Frier, 'Zuckerberg calls for global internet regulations,' *Australian Financial Review* (31 March 2019) <https://www.afr.com/news/world/north-america/zuckerberg-calls-for-global-internet-regulations-20190331-p519at>; Paul Smith, 'Microsoft president says big tech regulation must learn from history,' *Australian Financial Review* (2 April 2019) <https://www.afr.com/technology/technology-companies/microsoft-president-says-big-tech-regulation-must-learn-from-history-20190329-p518v2>.
192 See for example Dillon Reisman, Jason Schultz, Kate Crawford, Meredith Whittaker, 'Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability' (AI Now Institute, April 2018) <https://ainowinstitute.org/aiareport2018.pdf>; David J Briggs, 'A Framework for Integrated Environmental Health Impact Assessment of Systemic Risks' (2008) 7(1) *Environmental Health* 61.
193 For discussion see WP submissions: Microsoft, 3; J Powles, M Rizzi, F McGaughey, D Glance, 1, 2. See also University of Technology Sydney, submission to the IP, 97.
194 See, for example, IP submissions: C Nichols, 4; N Witzleb, 10; K Fenyo, 2; M Wildenauer, 3; Castan Centre for Human Rights Law 11; Kingsford Legal Centre,14. Queensland Council of Civil Liberties to WP, 5.
195 See IP submissions: M Miranda, 1; Lisa Fowkes, 5. See WP submissions: Law Council of Australia, 2; Blockchain Assets 12-15. For a general overview of privacy protection, see Australian Law Reform Commission, *Serious Invasions of Privacy in the Digital Era* (ALRC Report 123, June 2014) ch 3.
196 Australian Law Reform Commission, *Serious Invasions of Privacy in the Digital Era* (ALRC Report 123, June 2014) ch 3; Department of Prime Minister and Cabinet, *A Commonwealth Statutory Cause of Action for Serious Invasion of Privacy* (September 2011).
197 See IP submissions: D Harkin and A Molnar; N Witzleb, 7.
198 See IP submissions: Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University, 7; Financial Rights Legal Centre, 7; Castan Centre for Human Rights Law, 9; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 16; Office of the Australian Information Commissioner, 1.
199 Australian Competition and Consumer Commission, *Digital Platforms Inquiry: Final Report* (June 2019), 24 <https://www.accc.gov.au/focus-areas/inquiries-ongoing/digital-platforms-inquiry>.
200 See IP submissions: Consumer Policy Research Centre; University of Technology Sydney; University of Melbourne.
201 See D Svantesson, submission to the IP, 2, 3. See Australian Research Data Commons, submission to the WP, 1.
202 See the IP submissions: M Miranda, 2, 3; L Fowkes; CHOICE.
203 See WP submissions: Access Now, 3; Montreal AI Ethics Institute, 6; K Fenyo.
204 See the following submissions to the IP: Access Now, 5; A Johnston; Red Cross; Consumer Policy Research Centre.
205 See IP submissions: M Wigan, 2; Red Cross, 5; J Corker, 2; Australian Women Against Violence Alliance, 8; Speech Pathology Australia, 6.
206 See IP submissions: M Paterson, 1; CHOICE, 7.
207 See, for example, IP submissions: Law Council of Australia, 6; M Paterson, 33; National Association for Community Legal Centres, 4; Australian Centre for Health Law Research, 3.1; Deloitte, 2; La Trobe LawTech, 3; Australian Privacy Foundation, Queensland Council of Civil Liberties, Electronic Frontiers Australia, 25.
208 See IP submissions: Domestic Violence Victoria, 1; E Jane, N Vincent, 24; Australian Women Against Violence Alliance.
209 See IP submissions: N Henry, 3; Australian Women Against Violence Alliance, 11; Domestic Violence Victoria, 1.
210 See IP submissions: University of Technology Sydney, 53; LexisNexis, 17.
211 See IP submissions: K Manwaring, 267; Castan Centre for Human Rights Law, 12; *Competition and Consumer Act 2010* (Cth) sch 2 (Australian Consumer Law).
212 Castan Centre for Human Rights Law, submission to the IP, 11, 12.
213 *General Data Protection Regulation,* Regulation (EU) 2016/679.
214 *General Data Protection Regulation,* Regulation (EU) 2016/679.
215 Office of the Gene Technology Regulator, 'About the Regulator' (8 November 2016) <http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/Content/about-regulator-1>.
216 Australian Radiation Protection and Nuclear Safety Agency, 'About Regulatory Services' <https://www.arpansa.gov.au/regulation-and-licensing/regulation/about-regulatory-services>.
217 Civil Aviation Safety Authority, 'Who we are' <casa.gov.au/about-us/who-we-are>.
218 Australian Competition and Consumer Commission, *Digital Platforms Inquiry* (Final Report, June 2019).
219 Australian Securities and Investments Commission, 'Innovation Hub' <https://asic.gov.au/for-business/innovation-hub/>.
220 National Data Commissioner <https://www.datacommissioner.gov.au/>.
221 National Data Commissioner, 'Advisory Council' <https://www.datacommissioner.gov.au/advisory-council>.
222 *Ombudsman Act 1976* (Cth); Commonwealth Ombudsman, 'Our Legislation' <http://www.ombudsman.gov.au/Our-responsibilities/our-legislation>.
223 Commonwealth Ombudsman, submission to the IP.
224 Office of the eSafety Commissioner, 'Legislation' <https://www.esafety.gov.au/about-the-office/legislation>.
225 Office of the eSafety Commissioner, 'Legislation' <https://www.esafety.gov.au/about-the-office/legislation>.
226 *Australian Prudential Regulation Authority Act 1998* (Cth).
227 Geoff Summerhayes, 'Computer terminal velocity: APRA's response to an accelerating risk' (Speech delivered at the Insurance Council of Australia Annual Forum, 8 March 2018) <https://www.apra.gov.au/news-and-publications/computer-terminal-velocity-apras-response-to-an-accelerating-risk>.
228 Department of Finance and Administration, Australian Government Information Management Office, Commonwealth Ombudsman, Australian National Audit Office, Office of the Privacy Commissioner *Automated Assistance in Administrative Decision-Making, Better Practice Guide* (February 2007), 4.
229 Digital Transformation Agency, 'Digital Service Standard Criteria' (4 October 2019) <https://www.dta.gov.au/help-and-advice/digital-service-standard/digital-service-standard-criteria>.
230 Australian Securities & Investments Commission, 'Providing digital financial product advice to retail clients' (August 2016) <https://download.asic.gov.au/media/3994496/rg255-published-30-august-2016.pdf>.
231 Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University, submission to the IP, 11.
232 See, for example, Standards Australia, submission to the WP, 3.
233 See, for example, International Standards Organisation, '#hightechstandards' (8 October 2019) <https://www.iso.org/sites/hitechstandards/>; S Ramesh, submission to the IP, 3.
234 International Organization for Standardization, 'Standards' (10 October 2019) <https://www.iso.org /standards.html>.
235 See, for example, IP submissions: Deloitte, 6; P Butt, 7; E Fries, 1.
236 Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University, submission to the IP, 10.
237 See IP submissions: P Butt, 6; Consumer Policy Research Centre, 5; Deloitte, 5.
238 Microsoft, submission to the WP, 3.
239 Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University, submission to the IP, 9.
240 See IP submissions: Office of the Australian Information Commissioner, 9; Deloitte, 6.
241 See, for example, the following IP submissions: University of Melbourne, 10; Digital Gap Initiative, 15.
242 See, for example, Deloitte, submission to the IP, 6.
243 UN Global Compact <https://www.unglobalcompact.org/>.
244 See, for example, IP submissions: PwC Indigenous Consulting, 7; Castan Centre for Human Rights Law, 24.

245 Australian Competition and Consumer Commission, *Digital Platforms Inquiry* (Final Report, June 2019), <https://www.accc.gov.au/focus-areas/inquiries-ongoing/digital-platforms-inquiry>.
246 Australian Competition and Consumer Commission, *Digital Platforms Inquiry* (Final Report, June 2019), <https://www.accc.gov.au/focus-areas/inquiries-ongoing/digital-platforms-inquiry>.
247 Department of Finance and Administration, Australian Government Information Management Office, Commonwealth Ombudsman, Australian National Audit Office, Office of the Privacy Commissioner *Automated Assistance in Administrative Decision-Making, Better Practice Guide* (February 2007), 4.
248 Greg Medcraft, 'Digital disruption and data: an ASIC update' (September 2017) <https://download.asic.gov.au/media/4451299/greg-medcraft-finsia-published-8-september-2017.pdf>.
249 Note IP submissions: University of Technology Sydney, 9, 33, 40; R McLay,14; Google, 2; Australian Privacy Foundation, Queensland Council for Civil Liberties, and Electronic Frontiers Australia, 5, 25; E Salazar, J Parker Roberto, L Gelman, A Gajo, 22, 23; K Matthews-Hunt, 23; Migrant Worker Initiative, 6; Royal Australian and New Zealand College of Radiologists, 1-2; M Browning, M Ellis, K Yeoh, with T Leiman, 21, 22; R Calvo, J Huppert, D Peters, G Goggin, 2; Red Cross, 7; M Paterson, 7; Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University, 8; Consumer Policy Research Centre, 26; NSW Young Lawyers Communications, Entertainment and Technology Law Committee, 10; R Clarke, 5; Castan Centre for Human Rights Law, 4; Think Place, 8; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 5; The Ethics Centre, 40; PwC Indigenous Consulting, 30; Global Partners Digital,13; Confidential submission 88, 1; confidential submission 71, 4.
250 Matthew Beard and Simon Longstaff, *Ethical by Design: Principles for Good Technology* (The Ethics Centre, 2018), 16.
251 The Ethics Centre, submission to the IP, 16.
252 See, eg, M Beard and S Longstaff, *Ethical by Design: Principles for Good Technology* (The Ethics Centre, 2018); The Ethics Centre, submission to the IP, 19.
253 The Gradient Institute, 'Whose Ethics?' (Blog post, 25 March 2019) <https://gradientinstitute.org/blog/4/>.
254 See the visual mapping undertaken by the Cyberlaw Clinic, Berkman Klein Center for Internet and Society, Harvard University: Jessica Fjeld, Hannah Hilligoss, Nele Achten, Maria Levy Daniel, Joshua Feldman, Sally Kagay *Principled Artificial Intelligence: A Map of Ethical and Rights-Based Approaches* (Draft Report, 4 July 2019) <https://ai-hr.cyber.harvard.edu/primp-viz.html>.
255 Australian Government, Department of Industry, Innovation and Science, AI Ethics Principles (November 2019) <https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework>.
256 Australian Government, Department of Industry, Innovation and Science, *AI Ethics Framework* (November 2019).
257 Australian Government, Department of Industry, Innovation and Science, *AI Ethics Framework* (November 2019).
258 Australian Government, Department of Industry, Innovation and Science, *AI Ethics Framework* (November 2019).
259 Centre for Data Ethics and Innovation, *Two Year Strategy* (Online, 21 March 2019) <https://www.gov.uk/government/publications/the-centre-for-data-ethics-and-innovation-cdei-2-year-strategy/centre-for-data-ethics-cdei-2-year-strategy>.
260 European Commission, 'Ethics Guidelines for trustworthy AI' (8 April 2019) <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.
261 Hannah Hilligoss and Jessica Field, 'Introducing the Principled Artificial Intelligence Project' (7 June 2019) <https://cyber.harvard.edu/story/2019-06/introducing-principled-artificial-intelligence-project>.
262 Hannah Hilligoss and Jessica Field, 'Introducing the Principled Artificial Intelligence Project' (7 June 2019) <https://cyber.harvard.edu/story/2019-06/introducing-principled-artificial-intelligence-project>.
263 Facebook, 'Draft Charter: An Oversight Board for Content Decisions' <https://fbnewsroomus.files.wordpress.com/2019/01/draft-charter-oversight-board-for-content-decisions-2.pdf>.
264 Sundar Pichai, 'AI at Google: our principles' (2018) <https://www.blog.google/technology/ai/ai-principles/?_ga=2.143074357.1664373363.1558576496-1493895775.1558576496>.
265 Kent Walter, 'An external advisory council to help advance the responsible development of AI,' (Online, 26 March 2018) <https://blog.google/technology/ai/external-advisory-council-help-advance-responsible-development-ai/>.
266 Sundar Pichai, 'AI at Google: our principles' (2018) <https://www.blog.google/technology/ai/ai-principles/?_ga=2.143074357.1664373363.1558576496-1493895775.1558576496>.
267 Sam Levin, 'Google scraps AI ethics council after backlash: 'Back to the drawing board,' *The Guardian* (5 April 2019).
268 Microsoft, 'Microsoft AI principles' (Webpage) <https://www.microsoft.com/en-us/AI/our-approach-to-ai>.
269 Facebook, 'Facebook and the Technical University in Munich Announce New Independent TUM Institute for Ethics in Artificial Intelligence' (20 January 2019) <https://newsroom.fb.com/news/2019/01/tum-institute-for-ethics-in-ai/>.
270 PayPal, 'Code of Business Conduct & Ethics' (2018) <https://www.paypalobjects.com/digitalassets/c/website/marketing/global/about/responsible-practices/Code-of-Conduct-2018.pdf>.
271 Akamai, 'Code of Ethics' <https://www.akamai.com/us/en/privacy-policies/code-of-ethics.jsp>.
272 IEEE, 'Code of Ethics' <https://www.computer.org/education/code-of-ethics>.
273 The Association of Computing Machinery, 'ACM Code of Ethics and Professional Conduct' (2018) <https://www.acm.org/binaries/content/assets/about/acm-code-of-ethics-booklet.pdf>.
274 The Association of Computing Machinery, 'ACM Code of Ethics and Professional Conduct' (2018) 4.
275 Accenture, 'Responsible AI and Robotics: An Ethical Framework' <https://www.accenture.com/gb-en/company-responsible-ai-robotics>.
276 Accenture, 'Responsible AI: A Framework for Emailing Trust in Your AI Solutions' <https://www.accenture.com/_acnmedia/PDF-92/Accenture-AFS-Responsible-AI.pdf>.
277 Association of Information Science and Technology, 'ASIS&T Professional Guidelines' <https://www.asist.org/about/asist-professional-guidelines/>.
278 The Future of Life Institute, 'Asilomar AI Principles' (2017) <https://futureoflife.org/ai-principles/?cn-reloaded=1>. The Future of Life Institute has an objective to safeguard life and develop optimistic vision of the future.
279 Access Now, 'The Toronto Declaration: Protecting the right to equality and non-discrimination in machine learning systems' (2018).
280 Eileen Donahoe and Megan Macduffee Metzger, 'Artificial Intelligence and Human Rights' (2019) 30 (2) *Journal of Democracy* 115-126, 118.
281 For detailed case studies see Legal Aid New South Wales, submission to the IP.
282 Luke Stark and Anna Lauren Hoffman, 'Data is the New What? Popular Metaphors and Professional Ethics in Emerging Data Culture' *Journal of Culture Analytics* (3 May 2019), 4.
283 See Andrew McNamara, Justin Smith and Emerson Murphy-Hill, *Does ACM's Code of Ethics Change Ethical Decision Making in Software Development?* ESEC/FSE (November 4–9, 2018, United States). <https://people.engr.ncsu.edu/ermurph3/papers/fse18nier.pdf>; James Arvanitakis, 'What are tech companies doing about ethical use of data? Not much' *The Conversation* (Online, 28 November 2018) <https://theconversation.com/what-are-tech-companies-doing-about-ethical-use-of-data-not-much-104845>.
284 Meredith Whittaker, Kate Crawford, Roel Dobbe, Genevieve Fried, Elizabeth Kaziunas, Varoon Mathur, Sarah Myers West, Rashida Richardson, Jason Shutlz, Oscar Schwarts, *AI Now 2018 Report* (AI Now Institute, 2018), 29 <https://ainowinstitute.org/AI_Now_2018_Report.pdf>.
285 Luke Stark and Anna Lauren Hoffman, 'Data is the New What? Popular Metaphors and Professional Ethics in Emerging Data Culture' *Journal of Culture Analytics* (3 May 2019), 20.

286 See, for example, Australian Privacy Foundation, Electronic Frontiers Australia, Queensland Council for Civil Liberties, submission to the IP.
287 See, for example: Meredith Whittaker, Kate Crawford, Roel Dobbe, Genevieve Fried, Elizabeth Kaziunas, Varoon Mathur, Sarah Myers West, Rashida Richardson, Jason Shutlz, Oscar Schwarts, *AI Now 2018 Report* (AI Now Institute, 2018), 30; Ben Wagner, 'Ethics as an Escape from Regulation: From ethics-washing to ethics-shopping?' in M. Hildebgrant (ed) *Being Profiling. Cogitas ergo sum* (Amsterdam University Press, 2018); Thomas Metzinger, 'Ethics washing made in Europe' *Der Tagesspiegel* <https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/24195496.html>; David Watts, 'How Big Tech Designs Its Own Rules of Ethics to Avoid Scrutiny and Accountability' *The Conversation* (28 March 2019) <http://theconversation.com/how-big-tech-designs-its-own-rules-of-ethics-to-avoid-scrutiny-and-accountability-113457>.
288 Thomas Metzinger, 'Ethics washing made in Europe' *Der Tagesspiegel* <https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/24195496.html>; David Matthews, 'Big Tech Funding AI Ethics Research to "Delay and Avoid" Regulation' *Times Higher Education* (18 April 2019) <https://www.timeshighereducation.com/news/big-tech-funding-ai-ethics-research-delay-and-avoid-regulation>.
289 Thomas Metzinger, 'Ethics washing made in Europe' *Der Tagesspiegel* <https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/24195496.html>.
290 See IP submissions: University of Melbourne, 10; Castan Centre for Human Rights Law, 6, 7; Digital Gap Initiative, para 50; K Mathews-Hunt, 222; Australian Privacy Foundation, Queensland Council for Civil Liberties, and Electronic Frontiers Australia, 29; University of Technology Sydney, 20. These views were also reflected in the roundtable consultations.
291 Mark Latonero, *Governing Artificial Intelligence: Upholding Human Rights and Dignity* (Data&Society, 2018); Filippo Raso, Hannah Hilligoss, Vivek Krishnamurthy, Christopher Bavitz, Levin Kim, *Artificial Intelligence & Human Rights* (Berkman Klein Center, 2018); European Commission, 'Ethics Guidelines for trustworthy AI' (8 April 2019) <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.
292 Stark and Hoffman (2019), for example, note the possibility of approaching data governance through the lens of human rights, but also recognise that 'the suitability and effectiveness of human rights discourse in mobilising good corporate governance and chance for social justice is an open question'. Luke Stark and Anna Lauren Hoffman, 'Data is the New What? Popular Metaphors and Professional Ethics in Emerging Data Culture' *Journal of Culture Analytics* (3 May 2019), 17.
293 Eileen Donahoe and Megan MacDuffee Metzger, 'Artificial Intelligence and Human Rights' (2019) 30(2) *Journal of Democracy* 115-126, 121.
294 Office of the High Commissioner for Human Rights, '"Smart mix" of measures needed to regulate new technologies- Bachelet' (24 April 2019) <https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=24509>.
295 OECD, *The OECD Principles on AI* (9 June 2019) <https://www.oecd.org/going-digital/ai/principles/>.
296 See, for example, IP submissions: Australian Centre for Health Law Research, 6; Kingsford Legal Centre; Castan Centre for Human Rights Law, 6, 7; University of Melbourne, 7; La Trobe LawTech, 8; Emma Jane and Nicole Vincent, 44; University of Technology Sydney, 8; Royal Australian and New Zealand College of Radiologists,5; Australian Centre for Health Law Research,44; National Association of Community Legal Centres, 9.
297 Castan Centre for Human Rights Law, submission to the IP, 6, 7.
298 See, for example, IP submissions: Australian and New Zealand College of Radiologists, 5; Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University, 3; Castan Centre for Human Rights Law, 6; Data Synergies, 5.
299 University of Technology Sydney, submission to the IP, citing the Human Rights, Big Data, and Technology Project (HRBDT), "The Human Rights, Big Data and Technology Project – Written Evidence (AIC0196), Submission to the House of Lords Select Committee on Artificial Intelligence.
300 For example: Library of Congress, 'Regulation of Artificial Intelligence: East/South Asia and the Pacific' <https://www.loc.gov/law/help/artificial-intelligence/asia-pacific.php>.
301 Julia Angwin, Jeff Larson, Surya Mattu, Lauren Kirchner, *Machine Bias* (23 May 2016) Propublica. At <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
302 CSIRO, 'AI and Machine Learning' *Data 61* (Research focus area) (15 June 2018) <https://data61.csiro.au/en/Our-Research/Focus-Areas/AI-and-Machine-Learning>.
303 House of Lords Select Committee on Artificial Intelligence, *AI in the UK: ready, willing and able?* (House of Lords, Paper No HL100, Report of Session 2017-19, 16 April 2017), 13, <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/10002.htm>.
304 See, for example, Bo Seo, 'Fears Australia will be left behind with minimal AI spending,' *Financial Review* (28 January 2019) <https://www.afr.com/technology/fears-australia-will-be-left-behind-with-minimal-ai-spending-20190116-h1a4oo>; Elsa Kania, 'Technological entanglement: Cooperation, competition and the dual-use dilemma in artificial intelligence,' *Australian Strategic Policy Institute* (online, 28 June 2018) <https://www.aspi.org.au/report/technological-entanglement>.
305 Essential Research 'The Essential Report – Human Rights Commission' (26 July 2019).
306 Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Pearson, 3rd ed, 2016), 1.
307 AI Now, *The AI Now 2016 Report: The Social and Economic Implications of Artificial Intelligence Technologies in the Near-Term* (22 September 2016), 2, <https://ainowinstitute.org/AI_Now_2016_Report.pdf>. See also Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Pearson, 3rd ed, 2016), 2.
308 OECD *Artificial Intelligence in Society (Summary in English)* (2019), 1 <https://read.oecd-ilibrary.org/science-and-technology/artificial-intelligence-in-society/summary/english_9f3159b8-en#page1>.
309 Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Pearson, 3rd ed, 2016), 27. David Kaye, Special Rapporteur, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, 73rd sess, UN Doc A/73/348 (29 August 2018), 6, <https://www.un.org/ga/search/view_doc.asp?symbol=A/73/348>.
310 David Kaye, Special Rapporteur, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, 73rd sess, UN Doc A/73/348 (29 August 2018), 4 <https://www.un.org/ga/search/view_doc.asp?symbol=A/73/348>.
311 Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Pearson, 3rd ed, 2016), 1038.
312 Centre for Public Impact, *Destination Unknown: Exploring the impact of Artificial Intelligence on Government* (Working Paper, September 2017); Toby Walsh, Neil Levy, Genevieve Bell, Anthony Elliott, James Maclaurin, Iven Mareels and Fiona Wood, *The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve our Wellbeing* (Report for the Australian Council of Learned Academies, 2019), 16 <https://acola.org/hs4-artificial-intelligence-australia/>.
313 The issue was raised in IP submissions; see, for example, R Clarke, 2. It has also been raised by external stakeholders, the IEEE prefers the term 'autonomous and intelligent systems' rather than AI, in IEEE, *Ethically Aligned Design: A vision for prioritizing human well-being with autonomous and intelligent systems* (First Edition, 2019), 16 <https://ethicsinaction.ieee.org>.
314 See IP submissions: The Allens Hub for Technology, Law, and Innovation, 6; Castan Centre for Human Rights Law, 2; R McLay, 3.
315 Executive Office of the President of the United States, *Big Data: Seizing Opportunities, Preserving Values* (May 2014) <https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf>.
316 Executive Office of the President of the United States and the President's Council of Advisors on Science and Technology, Report to the President: *Big Data and Privacy – a Technological Perspective* (May 2014), 19, 4 <https://bigdatawg.nist.gov/pdf/pcast_big_data_and_privacy_-_may_2014.pdf>.
317 Executive Office of the President of the United States, *Big Data: Seizing Opportunities, Preserving Values* (May 2014), 4 <https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf>.

318 Consumer Policy Research Centre, *A Day in the Life of Data* (29 May 2019), 15 <https://cprc.org.au/wp-content/uploads/CPRC-research-finds-Australian-consumers-in-the-dark-when-it-comes-to-data-free-for-all_29-May-2019.pdf>. For a discussion about the commercialisation of data, see also Productivity Commission (Australian Government), 'Opportunities for Business' *Data Availability and Use, Productivity Commission Inquiry Report* (No. 82, 31 March 2017), 106 <https://www.pc.gov.au/inquiries/completed/data-access/report>.

319 Article 22 of the *General Data Protection Regulation*, (European Union) 2016/679 provides a right for a decision not to be wholly automated where the decision will have this impact.

320 Information Commissioner's Office (UK) 'What does the GDPR say about automated decision-making and profiling' <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-does-the-gdpr-say-about-automated-decision-making-and-profiling/>.

321 In its recent report, ACOLA defined decision-making as 'decision-making assisted by AI techniques such as ML'. Australian Council of Learned Academies *The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve our Wellbeing* (30 July, 2019), <https://acola.org/hs4-artificial-intelligence-australia/>.

322 David Kaye, Special Rapporteur, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, 73rd sess, UN Doc A/73/348 (29 August 2018), 4 <https://www.un.org/ga/search/view_doc.asp?symbol=A/73/348>.

323 The Royal Society, 'Can machines really learn?' (Web page) <https://royalsociety.org/topics-policy/projects/machine-learning/what-is-machine-learning-infographic/>. See also David Kaye, Special Rapporteur, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expressio*n, 73rd sess, UN Doc A/73/348 (29 August 2018), 4, <https://www.un.org/ga/search/view_doc.asp?symbol=A/73/348>; Tom Mitchell, *The Discipline of Machine Learning* (Carnegie Mellon University, July 2006), 2. <http://www.cs.cmu.edu/~tom/pubs/MachineLearning.pdf>; Executive Office of the President of the United States and the President's Council of Advisors on Science and Technology, *Report to the President: Big Data and Privacy – a Technological Perspective* (May 2014), 24 <https://bigdatawg.nist.gov/pdf/pcast_big_data_and_privacy_-_may_2014.pdf>.

324 Allens Linklaters, *AI Toolkit: Ethical, safe, lawful; Practical guidance for AI projects* (Allens, 2019), 9, <https://www.allens.com.au/campaigns/data-driven-business/ai-toolkit/>.

325 David Kaye, Special Rapporteur, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, 73rd sess, UN Doc A/73/348 (29 August 2018), 4 <https://www.un.org/ga/search/view_doc.asp?symbol=A/73/348>.

326 Infographic based on: David Kaye, Special Rapporteur, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, 73rd sess, UN Doc A/73/348 (29 August 2018), 5 <https://www.un.org/ga/search/view_doc.asp?symbol=A/73/348>.

327 Miguel Carrasco, Steven Mills, Adam Whybrew and Adam Jura, *The Citizen's Perspective on the Use of AI in Government* (1 March 2019) <https://www.bcg.com/en-au/publications/2019/citizen-perspective-use-artificial-intelligence-government-digital-benchmarking.aspx>; The Senate Finance and Public Administration References Committee, Parliament of Australia, *Digital delivery of government services* (Final Report, June 2018), <https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Finance_and_Public_Administration/digitaldelivery/Report>.

328 Martin Stewart-Weeks and Simon Cooper, *Are we there yet? The digital transformation of government and the public sector in Australia* (Longueville Media Pty Ltd, 2019).

329 See Parliament of Australia Library, Parliament of Australia, *Public Sector Digital Transformation: A Quick Guide* (Research Paper Series 2018-19, 2 April 2019) <https://parlinfo.aph.gov.au/parlInfo/download/library/prspub/6594516/upload_binary/6594516.pdf>.

330 The Digital Transformation Agency's 'Digital Sourcing Framework for ICT Procurement', for example, has developed best practice principles to guide government departments when sourcing digital products and services. These principles are: encourage competition; be innovative and iterate often; be structured in a way that enables small and medium enterprises to compete fairly to provide components of large ICT projects; be outcomes focused; use open standards and cloud first; minimise cyber-security risks; avoid duplication by not building platforms that other agencies have already built <https://www.dta.gov.au/help-and-advice/ict-procurement/digital-sourcing-framework-ict-procurement>.

331 See IP submissions: R McLay, 27; Law Council of Australia, 20; Royal Australian and New Zealand College of Radiologists, 2; Dieticians Association of Australia, 3.

332 See IP submissions: Law Council of Australia, 5; Pymetrics, 4.

333 See IP submissions: Law Council of Australia, 8; Kingsford Legal Centre, 2.

334 See, for example, IP submissions: Law Council of Australia, 16; I Law, 8; Commonwealth Ombudsman; Australian Privacy Foundation, Queensland Council for Civil Liberties, Electronic Frontiers Australia, 8; M Paterson, 4; P Henman, 6; M O'Sullivan, 3; Kingsford Legal Centre, 7; National Association of Community Legal Centres, 5.

335 As noted in IP submissions: PwC Indigenous Consulting, 23; Australian Privacy Foundation, the Queensland Council for Civil Liberties, and Electronic Frontiers Australia, 12.

336 As noted in IP submissions: University of Technology Sydney,45; Australian Privacy Foundation, 20.

337 P Henman, submission to the IP, 3. See also New Zealand Human Rights Commission *Privacy, Data and Technology: Human Rights Challenges in the Digital Age* (May 2018), 21 <https://www.hrc.co.nz/files/5715/2575/3415/Privacy_Data_Technology_-_Human_Rights_Challenges_in_the_Digital_Age_FINAL.pdf>.

338 As noted in IP submissions: National Association of Community Legal Centres, 5; Law Council of Australia, 16; Australian Privacy Foundation, the Queensland Council for Civil Liberties, and Electronic Frontiers Australia, 9; Kingsford Legal Centre, 7; P Henman, 4.

339 Commonwealth Ombudsman, submission to the IP, 4; Senate Community Affairs References Committee, 'Design, scope, cost-benefit analysis, contracts awarded and implementation associated with the Better Management of the Social Welfare System initiative' (21 June 2017).

340 Errors and discrepancies arose when an assumption was made about income, and, consequently, incorrect information being included in the OCI's calculation: see Senate Standing Committee on Community Affairs *Design, scope, cost-benefit analysis, contracts awarded and implementation associated with the Better Management of the Social Welfare System initiative* (21 June 2017), [2.85]-[2.101] <https://www.aph.gov.au/parliamentary_business/committees/senate/community_affairs/socialwelfaresystem/Report/c02>.

341 See, for example, Australian Council of Social Service, Submission 31 to the Senate Standing Committee on Community Affairs inquiry into the *Design, scope, cost-benefit analysis, contracts awarded and implementation associated with the Better Management of the Social Welfare System Initiative* (March 2017) <https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/SocialWelfareSystem/Submissions>.

342 Senate Standing Committee on Community Affairs *Design, scope, cost-benefit analysis, contracts awarded and implementation associated with the Better Management of the Social Welfare System initiative* (21 June 2017), [1.23] <https://www.aph.gov.au/parliamentary_business/committees/senate/community_affairs/socialwelfaresystem/Report/c02>.

343 Senate Standing Committee on Community Affairs, *Design, scope, cost-benefit analysis, contracts awarded and implementation associated with the Better Management of the Social Welfare System initiative* (21 June 2017), 107.

344 See Senate Standing Committee on Community Affairs 'Chapter 3 – Communicating with Centrelink' *Design, scope, cost-benefit analysis, contracts awarded and implementation associated with the Better Management of the Social Welfare System initiative* (21 June 2017) <https://www.aph.gov.au/parliamentary_business/committees/senate/community_affairs/socialwelfaresystem/Report/c02>.

345 Paul Farrell 'Government to review all robodebts based on ATO income averaging' *ABC News*, (20 November 2019) <https://www.abc.net.au/news/2019-11-19/robodebt-scheme-human-services-department-halts-existing-debts/11717188/>.
346 ICESCR, art 9; UDHR, art 22.
347 Committee on Economic, Social and Cultural Rights, *General Comment No 19. The Right to Social Security (Art. 9)*, E/C.12/GC/19 4 February 2008 para 24.
348 Committee on Economic, Social and Cultural Rights, *General Comment No 19. The Right to Social Security (Art. 9)*, E/C.12/GC/19 4 February 2008 para 24.
349 Noted in IP submissions: LexisNexis,16, 21; Law Council of Australia, 14, 15; University of Melbourne, 6, 7.
350 Noted in IP submissions: National Association of Community Legal Centres, 7; Australian Privacy Foundation, the Queensland Council for Civil Liberties, and Electronic Frontiers Australia,19; PwC Indigenous Consulting, 25.
351 Michael McGowan 'More than 50% of those on secretive NSW police blacklist are Aboriginal' *The Guardian* (11 November 2017) <https://www.theguardian.com/australia-news/2017/nov/11/more-than-50-of-those-on-secretive-nsw-police-blacklist-are-aboriginal>.
352 Noted in IP submissions: Public Interest Advocacy Centre, 11; PwC Indigenous Consulting, 33; National Association of Community Legal Centres, 7.
353 Vicki Sentas and Camilla Pandolfini, 'Policing Young People in NSW: A Study of the Suspect Targeted Managment Plan. A Report of the Youth Justice Coalition of NSW', 8 <https://www.piac.asn.au/wp-content/uploads/2017/10/17.10.25-YJC-STMP-Report.pdf>.
354 Portfolio Committee No. 4 – Legal Affairs, 'Budget estimates 2017-2018 Uncorrected Proof' (Parliament House, Sydney, 9 November 2017), 6.
355 Sarah Marsh 'UK police use of computer programs to predict crime sparks discrimination warning' *The Guardian* (4 February 2019) <https://www.theguardian.com/uk-news/2019/feb/03/police-risk-racial-profiling-by-using-data-to-predict-reoffenders-report-warns>.
356 See, for example Albert Meijer and Martijn Wessels, 'Predictive Policing: Review of Benefits and Drawbacks' [2019] *International Journal of Public Administration* 1, 8. Hannah Couchman, *Policing by Machine: Predictive Policing and the Threat to Our Rights* (Liberty, 2019), 19.
357 *Brennan Center for Justice at New York University School of Law v New York City Police Department* 160541/2016.
358 M Browning, M Ellis, K Yeoh, T Leiman, submission to the IP, 2.
359 Pymetrics, submission to the IP, 2.
360 Red Cross, submission to the IP, 4.
361 M Rizzi, D Glance, submission to the IP, 2.
362 Noted in IP submissions: Consumer Policy Research Centre, 25; The Warren Centre, 6, 7; Consultation Roundtable 9.
363 Noted in the following IP submissions: Consumer Policy Research Centre, 25; The Warren Centre, 6–7; Consultation Roundtable 9.
364 M Dean and M Vatter, submission to the IP, 2, 3.
365 University of Technology Sydney, submission to the IP, 47.
366 Noted in IP submissions: Law Council of Australia, 16; CHOICE, 8; Kingsford Legal Centre, 7; R McLay, 13.2; University of Technology Sydney,12; Australian Red Cross, 10; Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University, 7; Financial Rights Legal Centre, 8; Global Partners Digital, 5; Australian Privacy Foundation, 6; Emergency Services Levy Monitor, 21.
367 Sam Levin, 'Facebook Cracks down on Discriminatory Ads after Years of Backlash' *The Guardian*, (19 March 2019) <https://www.theguardian.com/technology/2019/mar/19/facebook-advertising-discrimination-lawsuit-aclu-race-gender>.
368 Jeffrey Dastin, 'Amazon Scraps Secret AI Recruiting Tool That Showed Bias against Women,' *Newswire* (10 October 2018) <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-iduskcn1mk08g>.
369 ICCPR, arts 2, 16 and 26.
370 ICESCR, art 6.
371 For a detailed expert discussion on the use of 'persuasive technologies' see the US Senate Committee on Commerce, Science & Transportation, hearing on *Optimizing for Engagement: Understanding the use of persuasive technology on internet platforms* (25 June 2019). <https://www.commerce.senate.gov/public/index.cfm/2019/6/optimizing-for-engagement-understanding-the-use-of-persuasive-technology-on-internet-platforms>.
372 Amy Binns, 'Cambridge Analytica Scandal: Facebook's User Engagement and Trust Decline,' *The Conversation* (28 March 2018) <http://theconversation.com/cambridge-analytica-scandal-facebooks-user-engagement-and-trust-decline-93814>.
373 See, for example, IP submissions: R McLay, 26; R Calvo, J Huppert, D Peters, G Goggin, 5; Law Council of Australia, 22; Financial Rights Legal Centre, 10; Data Synergies, 8; P Butt, 3; The Ethics Centre, 25.
374 Noted in IP submissions: Royal Australian and New Zealand College of Radiologists; C Devenney, C Mills, 5; Google, 1.
375 Royal Australian and New Zealand College of Radiologists, submission to the IP.
376 See IP submissions: D Glance and M Rizzi, 5; Northraine, 8.
377 See, for example, the following IP submissions: Australian Privacy Foundation, 16; F McDonald; M Wildenauer.
378 See Chapter 2 on Human Rights.
379 See, for example, submissions to IP: Law Council of Australia; N Suzor, K Weatherall; A Daly, A Vromen, M Mann; D Glance and M Rizzi; LexisNexis; University of Technology Sydney.
380 D Glance and M Rizzi, submission to IP, 5. See also: Northraine, 8.
381 N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 6, 7; Intopia, 9; WebKey IT, 5.
382 University of Technology Sydney, submission to IP, 14.
383 N Suzor, K Weatherall, A Daly, A Vromen, M Mann, submission to the IP, 12.
384 Consultation roundtables.
385 The precautionary principle has been adopted in certain contexts, for example, environmental regulation. See IP submissions: K Mathews-Hunt ; EMR Australia, 8; R Clarke, 6.
386 As noted in R Clarke, submission to IP, 6.
387 See IP submissions: N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 3; PwC Indigenous Consulting, 6.
388 See, for example, IP submissions: ThinkPlace, 1; Law Council of Australia, 23; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 12; M Rizzi, D Glance, 5; LexisNexis, 22; University of Technology Sydney, 17.
389 University of Melbourne, submission to the IP, 7.
390 *The Queen (on application of Edward Bridges) v Chief Constable of South Wales Police* [2019] EWHC 2341 (Admin), per Lord Justice Haddon-Cave and Mr Justice Swift at [84].
391 D Glance, M Rizzi, submission to the IP, 5.
392 See, for example, IP submissions: Castan Centre for Human Rights Law, 2; University of Melbourne, 2. See WP submissions: Microsoft; Digital Industry Group, 4.
393 See WP submissions: The Allens Hub for Technology, Law, and Innovation, 3; Microsoft; Digital Industry Group Inc.
394 Public Interest Advocacy Centre, submission to the IP, 8.
395 PwC Indigenous Consulting, submission to the IP, 19.
396 University of Technology Sydney, submission to the IP, 36.
397 See IP submissions: Australian Red Cross, 7 ,9; University of Technology Sydney, 60.
398 Royal Australian and New Zealand College of Radiologists, submission to the IP, 3.
399 As noted by several IP submissions, including: N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 2, 7; NSW Young Lawyers Communications, Entertainment and Technology Law Committee, 10; National Association of Community Legal Centres, 9; M Rizzi, D Glance, 2; Intopia, 8.
400 See IP Submissions: N Suzor, K Weatherall, A Daly, A Vromen, M Mann; University of Technology Sydney, 43.
401 See, for example, University of Technology Sydney, submission to the IP, 43, 45, 49.
402 University of Technology Sydney, submission to the IP, 36.
403 Noted in IP submissions: Chartered Accountants Australia and New Zealand, 1; Castan Centre for Human Rights Law, 18; R Chalmers, submission to the WP, 1; Australian Information and Industry Association, submission to the WP, 5.

404 See IP submissions: Australian Privacy Foundation, 5; Kingsford Legal Centre, 3; Office of the Australian Information Commissioner, 4.
405 See IP submissions: University of Melbourne, 8, 9; P Harpur, 10.
406 See IP submissions: P Henman, 1; M Wildenauer, 2; Castan Centre for Human Rights Law, 28.
407 The importance of transparency was discussed in a number of IP submissions including: The Warren Centre, 2; LexisNexis, 13.
408 LexisNexis, submission to the IP, 13.
409 Noted in IP submissions: Office of the Australian Information Commissioner; 4; C Nichols; Consumer Policy Research Centre, 40; K Mathews-Hunt, 52; University of Technology Sydney, 54; PwC Indigenous Consulting, 25; Australian Red Cross, 9.
410 Matthew Beard and Simon Longstaff, *Ethical by Design: Principles for Good Technology* (The Ethics Centre, 2018) 7; IP submissions Pymetrics, 2; Commonwealth Ombudsman, 2; PwC Indigenous Consulting, 25; The Warren Centre, 2.
411 CHOICE, submission to the IP, 20.
412 University of Melbourne, submission to the IP, 9.
413 Centre for Inclusive Design, submission to the IP, 4.
414 Public Interest Advocacy Centre, 8-9; Australian Red Cross; CHOICE, 18. See also Peter Leonard, 'AI and Technology Ethics: Notes for Discussions of the ACS AI Ethics Committee,' *Data Synergies* (6 September 2018), 26.
415 OECD, 'OECD Principles on AI' (May 2019) <https://www.oecd.org/going-digital/ai/principles/>. See also the G20, 'Ministerial Statement on Trade and Digital Economy' (June 2019) <https://www.mofa.go.jp/files/000486596.pdf>.
416 G20, 'Ministerial Statement on Trade and Digital Economy' (June 2019) [19] <https://www.mofa.go.jp/files/000486596.pdf>.
417 Council of Europe Commissioner for Human Rights, 'Unboxing Artificial Intelligence: 10 Steps to Protect Human Rights' (Council of Europe, May 2019).
418 Council of Europe Commissioner for Human Rights, 'Unboxing Artificial Intelligence: 10 Steps to Protect Human Rights' (Council of Europe, May 2019).
419 Berkman Klein Center for Internet and Society *Introducing the Principled Artificial intelligence Project* (Harvard, 2019) <https://clinic.cyber.harvard.edu/2019/06/07/introducing-the-principled-artificial-intelligence-project/>.
420 Berkman Klein Center for Internet and Society, *The role of "Principles" in the governance of Artificial Intelligence* (Harvard, 2019), 3. Accessed through the 8th World Peace Forum Conference Pack.
421 House of Lords Select Committee on Artificial Intelligence, *AI in the UK: ready, willing and able? Report of Session 2017-19* (16 April 2018), 116 <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>.
422 Department for Business, Energy & Industrial Strategy, Council for Science and Technology, The Rt Hon Greg Clark MP, Sam Gyimah MP and The Rt Hon Claire Perry MP, *Press Release, Business Secretary hosts first cross-government working group on future regulation* (24 October 2018) <https://www.gov.uk/government/news/business-secretary-hosts-first-cross-government-working-group-on-future-regulation>.
423 See Chapter 2 on Human Rights.
424 See, for example, Filippo Raso et al, 'Artificial Intelligence & Human Rights: Opportunities & Risks' (Berkman Klein Center Research Publication No 2018-6); Eileen Donahoe and Megan MacDuffee Metzger, 'Artificial Intelligence and Human Rights' (2019) 30(2) *Journal of Democracy* 115.
425 Some commentators and submissions are calling for new rights. See, for example IP submissions: R Calvo, J Huppert, D Peters, G Goggin, 5; N Witzleb, 27.
426 At the 2018 FAT (Fairness, Accountability and Transparency) conference, A/Prof Arvind Narayanan delivered a paper with the title 'Twenty-one Fairness Definitions and their Politics'. Cited and discussed in Ellen Broad *Made by Humans: The AI Condition* (2018), 92.
427 Article 2(3) of the *International Covenant on Civil and Political Rights* (ICCPR) requires each state party to ensure a person whose Covenant rights have been violated has an effective remedy, and that this remedy will be enforced. See also Human Rights Committee, *General Comment No. 31 [80] The Nature of the General Legal Obligation Imposed on State Parties to the Covenant* (CCPR/C/21/Rev/1/Add/ 1326 May 2004) (Adopted 29 March 2004).
428 Accountability is the second limb of the PANEL principles, discussed in Chapter 2.
429 This comment was made in the first major court decision on the use of automated facial recognition technology. See: *Bridges, R v The Chief Constable of South Wales Police* [2019] EWHC 2341 (Admin), at 84 (per Lord Justice Haddon-Cave and Mr Justice Swift).
430 The right to an effective remedy is contained in Art 2(3) of the ICCPR; Art 14 of CAT; Art 2 of CEDAW; Art 6 of CERD.
431 Accountability is the second limb of the PANEL principles, outlined above in Chapter 2. See also, UN Office of the High Commissioner for Human Rights and the Centre for Economic and Social Rights, *Who will be Accountable? Human rights and the post-2015 Development Agenda HR Pub/13/ 1* (United Nations, 2013), 10, <https://www.ohchr.org/Documents/Publications/WhoWillBeAccountable.pdf>.
432 UN Office of the High Commissioner for Human Rights and the Centre for Economic and Social Rights, *Who will be Accountable? Human rights and the post-2015 Development Agenda HR Pub/13/ 1* (2013) (United Nations, New York and Geneva), 10.
433 See IP Submissions: Law Council of Australia, 15, 19; University of Technology Sydney, 51, 52; LexisNexis, 17; Public Interest Advocacy Centre, 10; Intopia, 7, 8; K Manwaring, 58, 82; Financial Rights Legal Centre, 22; D Watts, 11; Castan Centre for Human Rights Law, 18; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 4, 6; Kingsford Legal Centre, 6, 7. See WP submissions: M Guihot and M Rimmer, 5; Montreal AI Ethics Institute, 7; R Clarke, 9, 11.
434 While the legal dividing line between government and non-government decisions is generally obvious, there are some situations where it is less clear (for example, where a non-government entity makes a decision on behalf of the government). See, generally, *NEAT Domestic Trading Pty Ltd v AWB* (2003) 216 CLR 277.
435 Section 5 *Administrative Decisions (Judicial Review) Act 1977* (Cth).
436 Justice Melissa Perry '*i*Decide: Administrative Decision-Making in the Digital World' (2017) 91(1) *Australian Law Journal* 29.
437 Justice Melissa Perry '*i*Decide: Administrative Decision-Making in the Digital World' (2017) 91(1) *Australian Law Journal* 29.
438 See IP submissions: R McLay, 11.4-11.7; Law Council of Australia, 13; University of Technology Sydney, 44, 53; LexisNexis, 16; WebKeyIT, question 6; Emergency Services Levy Insurance Monitor, 36; Public Interest Advocacy Centre, 12; C Devenney, C Mills, T Leiman, 9; Intopia, 6; M Paterson; 2; CHOICE, 7; Consumer Policy Research Centre, 26; N Witzleb, para 29; La Trobe LawTech, 5; Kingsford Legal Centre, 6; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 4.
439 See IP submissions: Law Council of Australia, 17; University of Technology Sydney, 45, 52; L Fowkes, 13; LexisNexis, 10; WebKeyIT, question 1; Emergency Services Levy Insurance Monitor, 34; Royal Australian and New Zealand College of Radiologists, 4; M Browning, M Ellis, K Yeoh, with T Leiman, 4, 6; Cl Devenney, C Mills, with T Leiman, 3, 5; D Harkin and A Molnar, question 2; A Johnston; Australian Privacy Foundation, Queensland Council for Civil Liberties, Electronic Frontiers Australia, 21; Australian Rehabilitation and Assistive Technology Association, 6; M Paterson. 2; Office of the Victorian Information Commissioner, 7; CHOICE, 18; K Manwaring,1; Consumer Policy Research Centre, 8; F McDonald; Deloitte, 51, 3; J Corker; NSW Young Lawyers Communications, Entertainment and Technology Law Committee, 11; La Trobe LawTech, 10; University of Melbourne, 8; M Miranda; Castan Centre for Human Rights Law, 12; Kingsford Legal Centre, 6; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 6.
440 Law Council of Australia, submission to the IP, 22.
441 See, for example, IP submissions: University of Technology Sydney, 53; LexisNexis, 17.
442 See, for example, IP submissions: Jobs Australia, 2; L Fowkes, 13; confidential submission, 115(2).
443 See IP submissions: Consumer Policy Research Centre, 40; Castan Centre for Human Rights Law, 11.

444 See IP submissions: LexisNexis, 19; R McLay, 20; Church in Australia, Synod of Victoria & Tasmania, 9.
445 See IP submissions: M Wigan, question 2; Australian Privacy Foundation, Queensland Council for Civil Liberties, Electronic Frontiers Australia, 4; M Paterson 3; C Nichols, 4; CHOICE, 18; N Witzleb, para 26; NSW Young Lawyers Communications, Entertainment and Technology Law Committee, 10; Access Now, 5; M Wildenauer,72; M Miranda; Castan Centre for Human Rights Law, 11; M Rizzi, D Glance, 87, 1; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 8; Confidential submission, 2.
446 See IP submissions: K Matthews-Hunt, 117; Confidential submission, recommendation 4. Also: consultation roundtable 8.
447 See IP submissions: Financial Rights Legal Centre, 14; Law Council of Australia, 23; University of Technology Sydney, 13, 52; M Browning, M Ellis, K Yeoh, T Leiman, 6; M Paterson, 4; Office of the Victorian Information Commissioner, 8; CHOICE, 19; Consumer Policy Research Centre, 2.
448 See IP submissions: M Wigan, question 1; M O'Sullivan, 3.
449 See IP submissions: Red Cross, 6; Uniting Church in Australia, Synod of Victoria & Tasmania, 8.
450 See IP submissions: Access Now, 4; Global Partners Digital; National Association of Community Legal Centres, 5; Australian Privacy Foundation, the Queensland Council for Civil Liberties, and Electronic Frontiers Australia, 11; Warren Centre, 5; PwC Indigenous Consulting, 25. See WP submissions: M Smith, 1; Law Council of Australia, 4; Montreal AI Ethics Institute, 8; M Guihot and M Rimmer, 26.
451 United States Congress, H.R.2231 – Algorithmic Accountability Act 2019 Bill (116th Congress, 2019-2020) <https://www.congress.gov/bill/116th-congress/house-bill/2231>.
452 United States Congress, H.R.2231 – Algorithmic Accountability Act 2019 Bill (116th Congress, 2019-2020), clause 2, <https://www.congress.gov/bill/116th-congress/house-bill/2231>.
453 *Racial Discrimination Act 1975* (Cth); *Sex Discrimination Act 1984* (Cth); *Australian Human Rights Commission Act 1986* (Cth); *Disability Discrimination Act 1992* (Cth); *Age Discrimination Act 2004* (Cth); *Fair Work Act 2009* (Cth).
454 This point was noted in IP submissions: Adobe, 2; University of Technology Sydney, 31, 43, 44; N Suzor, K Weatherall, A Daly, A Vromen, M Mann,18; University of Melbourne, 9; Law Council of Australia, 14.
455 In 2019, the United States Equal Opportunity Commission concluded the deliberate exclusion of women and older people from seeing job advertisements on a social media platform was discriminatory: United States Equal Opportunity Commission, Washington Office, *Determination* (Charge Number 570-2018-01036, 5 July 2019). See also: Ariana Tobin 'Employers used Facebook to keep women and older workers from seeing job ads. The Federal Government thinks that's illegal' *ProPublica* (24 September, 2019), <https://www.propublica.org/article/employers-used-facebook-to-keep-women-and-older-workers-from-seeing-job-ads-the-federal-government-thinks-thats-illegal?utm_source=pardot&utm_medium=email&utm_campaign=dailynewsletter>.
456 Maddalena Favaretto, Eva De Clercq and Bernice Simone Elger 'Big data and discrimination: perils, promises and solutions: A systematic review (2019) *Journal of Big Data* 6:12, 14.
457 Anja Lambrecht and Catherine Tucker, 'Algorithmic bias? An empirical study into apparent gender-based discrimination in the display of STEM career ads' (9 March 2018) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2852260>.
458 See also Maddalena Favaretto, Eva De Clercq and Bernice Simone Elger 'Big data and discrimination: perils, promises and solutions: A systematic review (2019) *Journal of Big Data* (2019) 6:12, 18; Leese M 'The new profiling: algorithms, black boxes, and the failure of anti-discriminatory safeguards in the European Union' (2014) 45(5) *Secure Dialogue* 494-511.
459 Maddalena Favaretto, Eva De Clercq and Bernice Simone Elger 'Big data and discrimination: perils, promises and solutions: A systematic review (2019) *Journal of Big Data* 6:12, 15. See also, Ziliobaite I, Custers B, 'Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models' (2016) 24(2) *Artificial Intelligence Law* 183-201.
460 See, for example, Google's 'Privacy and Security Principles', <https://safety.google/principles/>.
461 See IP submissions: La Trobe LawTech, 10; University of Technology Sydney, 43; Law Council of Australia, 17.
462 *Naperville Smart Meter Awareness v City of Naperville*, No. 16-3766 (7th Cir. 2018). See also, Privacy International 'Privacy Win! US court says Fourth Amendment protects smart meter data' (24 August 2018) <https://privacyinternational.org/blog/2234/privacy-win-us-court-says-fourth-amendment-protects-smart-meter-data>; and CHOICE, submission to the IP, 10.
463 Office of the Australian Information Commissioner 'What is personal information?' (Web page, 5 May 2017), <https://www.oaic.gov.au/privacy/guidance-and-advice/what-is-personal-information/#ftn4>.
464 See Office of the Australian Information Commissioner, submission to the IP, 7. The OAIC also produces relevant guidance for organisations dealing with data, and conducting data analytics; see, for example, Office of the Australian Information Commissioner, *Guide to data analytics and the Australian Privacy Principles*, (March 2018) <https://www.oaic.gov.au/privacy/guidance-and-advice/guide-to-data-analytics-and-the-australian-privacy-principles/>.
465 APP 6.1 provides a right to access 'personal information' that is 'about an individual'. The Full Court of the Federal Court of Appeal in *Privacy Commissioner v Telstra Corporation Limited* [2017] FCAFC 4 determined that in order for information to be considered 'personal information…about an individual' within the parameters of the Privacy Act, it will need to be determined on a case by case basis whether the information requested is an about an individual, which 'will require an evaluative conclusion, depending upon the facts of any individual case, just as a determination of whether the identity can *reasonably* be ascertained will require an evaluative conclusion' [63].
466 Australian Law Reform Commission *Serious Invasions of Privacy in the Digital Era* (ALRC Report 123), 3.50 <https://www.alrc.gov.au/publication/serious-invasions-of-privacy-in-the-digital-era-alrc-report-123/3-overview-of-current-law-2/filling-the-gaps-in-existing-law/>.
467 Australian Law Reform Commission *For your information: Australian privacy law and practice* (ALRC Report 108, 12 August 2008) <https://www.alrc.gov.au/publication/for-your-information-australian-privacy-law-and-practice-alrc-report-108/>.
468 See IP submissions: K Matthews-Hunt, 242; N Witzleb; National Association of Community Legal Centres, 9; M Browning, M Ellis, K Yeoh, with T Leiman, 5; Australian Privacy Foundation, the Queensland Council for Civil Liberties, and Electronic Frontiers Australia, 22; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 6; Law Council of Australia, 18.
469 Australian Competition and Consumer Commission, 'Recommendation 19: Statutory tort for serious invasions of privacy,' *Digital Platforms Inquiry: Final Report* (June 2019), 37, <https://www.accc.gov.au/focus-areas/inquiries-ongoing/digital-platforms-inquiry>.
470 See IP submissions: Financial Rights Legal Centre, 14; Law Council of Australia, 23; University of Technology Sydney, 13, 52; M Browning, M Ellis, K Yeoh, T Leiman, 6; M Paterson, 4; Office of the Victorian Information Commissioner, 8; CHOICE, 19; Consumer Policy Research Centre, 2.
471 See *Social Security Act 1999* (Cth) s 6A; *Superannuation (Government Co-contribution for Low Income Workers) 2003* (Cth) s 48; *My Health Records Act 2012* (Cth) s 13A; *Child Support Assessment Act 1989* (Cth) s12A. See Simon Elvery 'How algorithms make important government decisions – and how that affects you,' *ABC News* (21 July 2017) <https://www.abc.net.au/news/2017-07-21/algorithms-can-make-decisions-on-behalf-of-federal-ministers/8704858#legislation>.
472 *Migration Legislation Amendment (Electronic Transactions and Methods of Notification) Act 2001* (Cth), s 9.
473 Parliament of Australia *Migration Legislation Amendment (Electronic Transactions and Methods of Notification) Bill 2001, Explanatory Memorandum*, para 9.
474 Administrative Review Council *Automated Assistance in Administrative Decision Making, Report to the Attorney-General* (Report No. 46) (November 2004), viii, <https://www.ag.gov.au/LegalSystem/AdministrativeLaw/Pages/publications/report-46.aspx>.

475 Submissions identified the need for review and oversight of the use of AI-informed decision making by government. See IP submissions: M O'Sullivan; Allens Hub for Technology, Law and Innovation, 8; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 7.
476 *Australian Constitution*, s 75(v).
477 *Pintarich v Deputy Commissioner of Taxation* [2018] FCAFC 79.
478 *Pintarich v Deputy Commissioner of Taxation* [2018] FCAFC 79, [140]; approving the decision of Finn J in *Semunigus v Minister for Immigration and Multicultural Affairs* [1999] FCA 422 at [19], affirmed in *Seamunigus v Minister for immigration and Multicultural Affairs* [2000] FCA 240.
479 *Pintarich v Deputy Commissioner of Taxation* [2018] FCAFC 79, [46], [49]. Compare *Gramar Pty Ltd (t/as Valley View Nursing Home) v Workcover Corporation of South Australia* [2011] SASC 237; the Supreme Court of South Australia considered the plaintiff's argument that the imposition of a supplementary levy, calculated by an automated computer generated process, was not a decision by the delegated authority under the legislation in question and therefore was an invalid decision that should be set aside. Justice Kelly determined that as the automated process was based on a formula adopted and maintained for the purposes of the levy scheme some years prior: 'Effectively therefore the rate was arrived at by the defendant in accordance with its policies, even though the precise calculation was performed by a computer, using the formula'. [33], see para's [27]-[34].
480 Josh Bavas 'Facial recognition quietly switched on at Queensland stadiums, sparking privacy concerns' *ABC News* (6 June 2019) <https://www.abc.net.au/news/2019-06-05/facial-recognition-quietly-switched-on-at-queensland-stadiums/11178334>.
481 See, for example, the implementation of the 'Social Credit System' in China: Vicky Xiuzhong Xu and Biang Xiao, 'China's Social Credit System seeks to assign citizens scores, engineer social behaviour,' *ABC News* (2 April 2018) <https://www.abc.net.au/news/2018-03-31/chinas-social-credit-system-punishes-untrustworthy-citizens/9596204>; Darren Byler 'China's hi-tech war on its Muslim minority,' The Guardian (11 April 2019) <https://www.theguardian.com/news/2019/apr/11/china-hi-tech-war-on-muslim-minority-xinjiang-uighurs-surveillance-face-recognition>.
482 See <www.gendershades.org>; Tom Simonite 'The Best Algorithms Struggle to Recognise Black Faces Equally' *Wired* (22 July 2019) <https://www.wired.com/story/best-algorithms-struggle-recognize-black-faces-equally/>.
483 Parliamentary Joint Committee on Intelligence and Security *Advisory report on the Identity-matching Services Bill 2019 and the Australian Passports Amendment (Identity-matching Services) Bill 2019* (24 October 2019), Recommendation 1, 5.7 <https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/Intelligence_and_Security/Completed_inquiries>.
484 Parliamentary Joint Committee on Intelligence and Security *Advisory report on the Identity-matching Services Bill 2019 and the Australian Passports Amendment (Identity-matching Services) Bill 2019* (24 October 2019), 75 <https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/Intelligence_and_Security/Completed_inquiries>.
485 Professor Pete Fussey and Dr Daragh Murray, The Human Rights, Big Data and Technology Project *Independent report on the London Metropolitan Police Service's Trial of Live Facial Recognition Technology* (Economic & Social Research Council and the Human Rights Centre, University of Essex, July 2019), 5.
486 Science and Technology Committee of the United Kingdom Parliament, *Biometric strategy and forensics services* (Fifth Report of Session 2017-19) (HC 800) (25 May 2018), 4 <https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/800/800.pdf>.
487 *Bridges, R v The Chief Constable of South Wales Police* [2019] EWHC 2341 (Admin), at 97.
488 House of Commons Science and Technology Committee (United Kingdom Parliament) *The work of the Biometrics Commissioner and the Forensic Science Regulator* (House of Commons Paper 1970, Nineteenth Report of Session 2017-19, 18 July 2019), 4 <https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/1970/197003.htm#_idTextAnchor000>.
489 See, for example, Carly Kind (Ada Lovelace Institute), 'Biometrics and facial recognition technology – where next?' (Web page, 2 July 2019) <https://www.adalovelaceinstitute.org/biometrics-and-facial-recognition-technology-where-next/>.
490 AP 'San Francisco becomes first US city to ban police use of facial recognition technology' *ABC News* (15 May 2019), <https://www.abc.net.au/news/2019-05-15/san-francisco-bans-police-use-of-facial-recognition-technology/11116204>.
491 Kate Crawford, 'Halt the use of facial-recognition technology until it is regulated,' *Nature* (27 August 2019), <https://www.nature.com/articles/d41586-019-02514-7>.
492 Monika Zalnieriute, Lyria Bennett Moses and George Williams 'The Rule of Law and Automation of Government Decision-Making' (2019) 82(3) *The Modern Law Review* 425-455, 429.
493 See IP Submissions: L Fowkes, 7; M Paterson; M Wildenauer; Allens Hub for Technology, Law and Innovation; CHOICE, 2.
494 See IP submissions: University of Melbourne, 8; PwC Indigenous Consulting, 32; M Rizzi, D Glance, 3; LexisNexis, 10, 12; Northraine, 2, 11; Law Council of Australia, 23; Emergency Services Levy Insurance Monitor, 34; University of Technology Sydney, 44.
495 See IP submissions: Royal Australian and New Zealand College of Radiologists, 2; University of Technology Sydney, 52; Northraine,11; S Carter, N Houssami, W Rogers, submission to the WP, 5.
496 Commonwealth Ombudsman, submission to the IP, 8.
497 Toby Walsh, Neil Levy, Genevieve Bell, Anthony Elliott, James Maclaurin, Iven Mareels and Fiona Wood, *The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve our Wellbeing* (Report for the Australian Council of Learned Academies, 2019), 10 <https://acola.org/hs4-artificial-intelligence-australia/>.
498 See IP submissions: University of Technology Sydney, 45; Commonwealth Ombudsman, 7; M Browning, M Ellis, K Yeoh, with T Leiman, 15; R McLay, 18; University of Technology Sydney, 44, 45; P Henman, 7; M O'Sullivan, 5; Data Synergies, 66, 10; Allens Hub for Technology, Law and Innovation, 6, 8; P Butt, 7; University of Melbourne, 8; Castan Centre for Human Rights Law, 3; Northraine, 11; M Rizzi, D Glance, 4; Kingsford Legal Centre, 6; Confidential submission, 5.
499 There are exceptions where reasons are not required. See, for example, *Taxation Administration Act 1953* (Cth) s 14ZZB.
500 See, for example, s 13 of the *Administrative Decisions (Judicial Review) Act 1977* (Cth), which states a statement of reasons must be given on request to a person who is entitled to apply for judicial review of a decision by the Federal Court or Federal Circuit Court.
501 Commonwealth Ombudsman, 'How we can help; Australian Government Agencies and Services' (Webpage) <https://www.ombudsman.gov.au/How-we-can-help/australian-government-agencies-and-services>.
502 Administrative Review Council *Decision Making: Reasons (Best-practice guide 4)* (Commonwealth of Australia, August 2007), 7 <https://www.ag.gov.au/LegalSystem/AdministrativeLaw/Pages/administrative-review-council-publications.aspx>.
503 High-Level Expert Group on Artificial Intelligence (Set up by the European Commission), *Ethics Guidelines for Trustworthy AI* (European Commission, April 2019), 18 <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.
504 High-Level Expert Group on Artificial Intelligence (Set up by the European Commission), *Ethics Guidelines for Trustworthy AI* (European Commission, April 2019), 18 <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.
505 Sandra Wachter, Brent Mittelstadt and Luciano Floridi 'Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation' 2017 7(2) *International Data Privacy Law* 76.
506 T Walsh, K Miller, J Goldenfein, F Chen, J Zhou, R Nock, B Rubinstein, M Jackson, *Closer to the machine; Technical, social and legal aspects of AI* (Office of the Victorian Information Commissioner, August, 2019), 56, <https://ovic.vic.gov.au/closer-to-the-machine-ai-publication/>.
507 European Commission for the Efficiency of Justice (CEPEJ) and Council of Europe, *European Ethical Charter on the use of Artificial Intelligence in judicial systems and their environment* (Adopted at the 31st plenary meeting, of the CEPEJ 3-4 December 2018), 55 <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>.

508 Danielle Keats Citron and Frank Pasquale 'The Scored Society: Due Process for Automated Predictions' (2014) 89, *Washington University Law Review* 1, 33.

509 Toby Walsh, Neil Levy, Genevieve Bell, Anthony Elliott, James Maclaurin, Iven Mareels and Fiona Wood, *The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve our Wellbeing* (Report for the Australian Council of Learned Academies, 2019), 124, <https://acola.org/hs4-artificial-intelligence-australia/>.

510 *Houston Federal of Teaches Local 2415 v Houston Independent School District* (28:2201 Declaratory Judgment, 4 May 2017). See also, AI Now Institute, *AI Now Report 2018*, (December 2018) 29-21, <https://ainowinstitute.org/reports.html>.

511 The parties ultimately settled the claim, following a partially successful application by the defendant for summary judgment; the court refused to give summary judgement on the constitutional questions: See American Federation of Teachers *Media Release: Federal suit settlement: End of Value-Added Measures for Teacher Termination in Houston* (10 October, 2017) <https://www.aft.org/press-release/federal-suit-settlement-end-value-added-measures-teacher-termination-houston>.

512 IBM Watson OpenScale, for example, is a product that promises its users 'more fair and explainable outcomes for customers'; see IBM *Manage AI, with trust and confidence in business outcomes: IBM Watson OpenScale is an open platform that helps remove barriers to enterprise-scale AI*' (Product Brief), 4 <https://www.ibm.com/cloud/watson-openscale/>.

513 Executive Office of the President of the United States *The national artificial intelligence research and development strategic plan: 2019 update* (A Report by the Select Committee on Artificial Intelligence of the National Science and Technology Council) (June 2019), 5, <https://www.whitehouse.gov/wp-content/uploads/2019/06/National-AI-Research-and-Development-Strategic-Plan-2019-Update-June-2019.pdf>.

514 Executive Office of the President of the United States *The national artificial intelligence research and development strategic plan: 2019 update* (A Report by the Select Committee on Artificial Intelligence of the National Science and Technology Council) (June 2019), 'Strategy 1', 7, <https://www.whitehouse.gov/wp-content/uploads/2019/06/National-AI-Research-and-Development-Strategic-Plan-2019-Update-June-2019.pdf>.

515 Nichol Turner Lee, Paul Resnick and Genie Barton *Report Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms* (The Brookings Institution, 22 May 2019), <https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/#footnote-7>.

516 Russell and Norvig note 'Throughout the 60-year history of computer science, the emphasis has been on the *algorithm* as the main subject of study. But some recent work in AI suggests that for many problems, it makes more sense to worry about the *data* and be less picky about what algorithm to apply. This is true because of the increasing availability of very large data sources'. Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Pearson, 3rd ed, 2016), 27. See IP submissions: Webkey IT, 4; University of Technology Sydney, 43, 44; Office of the Australian Information Commissioner, 7; PwC Indigenous Consulting, 32, 33; Northraine, 6.

517 Executive Office of the President of the United States, *Big Data: A report on algorithmic systems, opportunity, and civil rights* (May 2016), 7, <https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf>. See also IP submissions: Adobe, 2; University of Technology Sydney, 44; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 18; Intopia, 7; University of Melbourne, 8; LexisNexis, 16, 17; Office of the Australian Information Commissioner; University of Melbourne, 8.

518 The European Union Agency for Fundamental Rights, for example, stated in its 2018 report on big data, that if algorithms are fed biased data 'discrimination will be replicated, perpetuated and potentially even reinforced', see European Union Agency for Fundamental Rights, *#BigData: Discrimination in data-supported decision making* (FRA Focus, 2018), 10. This point has also been recognised by the UK House of Lords Select Committee on AI: *AI in the UK: ready, willing and able?* (2018), 5; Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (Pearson, 3rd ed, 2016), 495; Executive Office of the President of the United States, *Big Data: A report on algorithmic systems, opportunity, and civil rights* (May 2016), 8 <https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf>.

519 Jeffrey Dastin, 'Amazon scraps secret AI recruiting tool that showed bias against women,' *Reuters* (10 October 2018) <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>.

520 Matthias Leese, 'The new profiling: algorithms, black boxes, and the failure of anti-discriminatory safeguards in the European Union' (2014) 45(5) *Security Dialogue* 494-511.

521 Maddalena Favaretto, Eva De Clercq and Bernice Simone Elger 'Big data and discrimination: perils, promises and solutions: A systematic review (2019) *Journal of Big Data* 6:12, 14.

522 Australian Law Reform Commission, 'Other methods of proof: Assessors, court experts, pre-sentence reports,' *Recognition of Aboriginal Customary Laws* (ALRC Report 31,18 August 2010) <https://www.alrc.gov.au/publication/recognition-of-aboriginal-customary-laws-alrc-report-31/26-other-methods-of-proof-assessors-court-experts-pre-sentence-reports/>.

523 See Therapeutic Goods Administration, Department of Health, Australian Government 'Evaluation of a new medicine – The TGA's lifecycle approach to regulation' (30 June 2011) <https://www.tga.gov.au/behind-news/evaluation-new-medicine-tgas-lifecycle-approach-regulation>; US Food and Drug Administration *Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device, Discussion Paper and Request for Feedback* (May, 2019), 15 <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>.

524 See IP submissions: Law Council of Australia, 23; University of Technology Sydney, 52, 53; Migrant Worker Initiative, 35; LexisNexis, 12; M Browning, M Ellis, K Yeoh, T Leiman, 4; Australian Privacy Foundation, 4, 8; M Paterson, 4; CHOICE, 19; N Witzleb, 8; Roger Clarke, 5; M O'Sullivan, 12; Castan Centre for Human Rights Law, 18; Office of the Australian Information Commissioner, 5. Submissions to the WP: Confidential submission, 1; Montreal AI Ethics Institute, 10; Roger Clarke, 12; The University of Melbourne, 6; Access Now, 3.

525 *General Data Protection Regulation*, (European Union) 2016/679 arts 13-15.

526 *General Data Protection Regulation*, (European Union) 2016/679 art 22(1).

527 *General Data Protection Regulation*, (European Union) 2016/679 art 22(3).

528 Section 14(4) of the *Data Protection Act 2018* (UK) allows for automated decisions to be challenged. It requires the data subject to be notified that a decision has been made solely on automated processing, and gives the data subject the opportunity to ask for the decision to be reconsidered, or ask the data controller to take a new decision that is not based solely on automated processing.

529 Sandra Wachter, Brent Mittelstadt and Luciano Floridi 'Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation' *International Data Privacy Law* (2017) Vol 7(2).

530 Andrew D Selbst and Julia Powles 'Meaningful information and the right to explanation' (2017) *International Data Privacy Law 233*, 7(4), 20.

531 Margot E Kaminski 'The right to explanation, explained' (2019) 34(189) *Berkeley Technology Law Journal* 190, 215.

532 N Suzor, K Weatherall, A Daly, A Vromen, M Mann, submission to the IP, 3, 10.

533 See IP submissions: R McLay, 9.1; University of Technology Sydney, 34; LexisNexis, 16; Royal Australian and New Zealand College of Radiologists, 5; Intopia, 6, 7; P Henman, 6; NSW Young Lawyers Communications, Entertainment and Technology Law Committee, 11; N Suzor, K Weatherall, A Daly, A Vromen and M Mann, 3. See WP submissions: Law Council of Australia, 2; Montreal AI Ethics Institute, 7; Royal Australian and New Zealand College of Radiologists, 3; R Clarke, 17; Microsoft, 5. See also, Toby Walsh, Neil Levy, Genevieve Bell, Anthony Elliott, James Maclaurin, Iven Mareels and Fiona Wood, *The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve our Wellbeing* (Report for the Australian Council of Learned Academies, 2019), 137 <https://acola.org/hs4-artificial-intelligence-australia/>.

534 Government of Canada '4. Objectives and Expected Results' *Directive on Automated Decision-Making* (2 May 2019) <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592#appA>.

535 Scott R Peppet 'Regulating the internet of things: first steps toward managing discrimination, privacy, security and consent' *Tex L Rev* 2014; 93:85. See also Australian Competition and Consumer Commission, *Digital Platforms Inquiry: Final Report* (June 2019), section 8.2, <https://www.accc.gov.au/focus-areas/inquiries-ongoing/digital-platforms-inquiry>.

536 *Ewert v Canada 2018* SCC 30 [2018] 2 S.C.R 165 (13 June, 2018).

537 *Ewert v Canada 2018* SCC 30 [2018] 2 S.C.R 165 (13 June, 2018), [65]. The court drew this conclusion in the context of requirements under the relevant legislation to use reliable and up to date information, and to be responsive to the needs of particular groups of offenders, including Indigenous Canadians: [44], [49], [53].

538 Alessandro Mantelero 'Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection' *Computer Law & Security Review* (Vol. 32, Issue 2) 238-255, 238.

539 Virginia Eubanks, *Automating Inequality, How high tech tools profile, police and punish the poor* (2017) (New York: St Martin's Press), 6, 7.

540 M Wildenauer, submission to the IP, answer to question 2.

541 See PwC Indigenous Consulting, submission to the IP.

542 Australian Human Rights Commission, *Discussion Paper on Discrimination Law Reform* (2019) <https://www.humanrights.gov.au/our-work/rights-and-freedoms/publications/discussion-paper-priorities-federal-discrimination-law>.

543 See IP submissions: National Association of Community Legal Centres, 5; LexisNexis, 9; Law Council of Australia, 16; Intopia, 7; University of Technology Sydney, 43, 45; Office of the Australian Information Commissioner, 7; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 12.

544 High-Level Expert Group on Artificial Intelligence (Set up by the European Commission) *Ethics Guidelines for Trustworthy AI* (European Commission, April 2019), 18 <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

545 Jake Goldenfein 'Algorithmic transparency and decision-making accountability', in T Walsh, K Miller, J Goldenfein, F Chen, J Zhou, R Nock, B Rubinstein, M Jackson *Closer to the machine; Technical, social and legal aspects of AI* (Office of the Victorian Information Commissioner, August, 2019), 56 <https://ovic.vic.gov.au/closer-to-the-machine-ai-publication/>.

546 Monika Zalnieriute, Lyria Bennett Moses and George Williams 'The Rule of Law and Automation of Government Decision-Making' (2019) *Modern Law Review* 82(3) 425-455.

547 Monika Zalnieriute, Lyria Bennett Moses and George Williams 'The Rule of Law and Automation of Government Decision-Making' (2019) *Modern Law Review* 82(3) 425-455, 454.

548 Justice R French, 'Judicial review rights,' AIAL Forum, no. 28 (2011), 30, cited in Administrative Review Council Automated Assistance in Administrative Decision Making, Report to the Attorney-General (Report no. 46) (November 2004), vii <https://www.arc.ag.gov.au/Documents/AAADMreportPDF.pdf>.

549 Professor Aarvind Narayanan, 'Tutorial: 21 fairness definitions and their politics' (Conference paper, presented at the Conference on Fairness, Accountability and Transparency, Atlanta, United States, 23 February, 2018) <https://docs.google.com/document/d/1bnQKzFAzCTcBcNvW5tsPuSDje8WWWY-SSF4wQm6TLvQ/edit>.

550 These issues are currently being considered in the Commission's National Conversation on Human Rights. See 'Free and Equal: An Australian Conversation on Human rights' (Web page, Australian Human Rights Commission, 31 May 2019) <https://www.humanrights.gov.au/free-and-equal#uCh68>.

551 Office of the Australian Information Commissioner 'Australian attitudes to privacy survey 2017' (updated November 2019) <https://www.oaic.gov.au/updates/videos/australian-community-attitudes-to-privacy-survey-2017/>. See also, Office of the Australian Information Commissioner, submission to the IP, 6.

552 Victorian Law Reform Commission *Surveillance in Public Places (Final Report 18)* (1 June, 2010), <https://www.lawreform.vic.gov.au/projects/surveillance-public-places/surveillance-public-places-final-report>; NSW Law Reform Commission *Invasion of Privacy (Report 120)* (April 2009), <https://www.lawreform.justice.nsw.gov.au/Pages/lrc/lrc_completed_projects/lrc_privacy.aspx>; Australian Law Reform Commission, *Serious Invasions of Privacy in the Digital Era* (Report 123, June 2014) <https://www.alrc.gov.au/wp-content/uploads/2019/08/summary_report_whole_pdf_.pdf>.

553 Australian Law Reform Commission, *Serious Invasions of Privacy in the Digital Era* (Report 123, June 2014), para 5.58 <https://www.alrc.gov.au/wp-content/uploads/2019/08/summary_report_whole_pdf_.pdf>.

554 Australian Human Rights Commission *Discussion Paper: Priorities for federal discrimination law reform* (2019) <https://www.humanrights.gov.au/our-work/rights-and-freedoms/publications/discussion-paper-priorities-federal-discrimination-law>.

555 AI Now Institute *AI Now Report 2018*, (December 2018) 22 <https://ainowinstitute.org/reports.html>.

556 Senate Standing Committee on Community Affairs, *Design, scope, cost-benefit analysis, contracts awarded and implementation associated with the Better Management of the Social Welfare System initiative* (21 June 2017), <https://www.aph.gov.au/parliamentary_business/committees/senate/community_affairs/socialwelfaresystem/Report/c02>; see also Commonwealth Ombudsman, submission to the IP.

557 The Hon Robert French AC, 'Rationality and reason in administrative law – would a roll of the dice be just as good?' Australian Academy of Law Annual Lecture (Speech delivered 29 November 2017, Perth), citing *Wainohu v New South Wales* (2011) 243 CLR 181 and *Public Service Board (NSW) v Osmond* (1986) 159 CLR 656.

558 Essential Research, *The Essential Report – Human Rights Commission* (26 July 2019).

559 ARC, '2020 ARC Centre for Automated Decision-Making and Society' <https://www.arc.gov.au/2020-arc-centre-excellence-automated-decision-making-and-society>.

560 *Houston Federal of Teaches Local 2415 v Houston Independent School District* (28:2201 Declaratory Judgment, 4 May 2017).

561 *Data Protection Act 2018* (UK) s 14(4).

562 See, for example, Min Kyung Lee, 'Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management' (Big Data & Society, 2018).

563 See, for example, Berkeley Dietvorst, Joseph Simmons and Cade Massey, 'Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err' *Journal of Experimental Psychology* (2014).

564 See, for example, Jennifer Logg, Julia Minson, Don Moore, 'Algorithm Appreciation: People Prefer Algorithmic to Human Judgment' (Working Paper 17-086, Harvard Business School); Rebecca Williams, 'Rethinking Deference for Algorithmic Decision-Making (Oxford Legal Studies Research Paper No.7/2019, 31 August 2018.

565 Niels Wouters, Ryan Kelly, Eduardo Velloso, Katrin Wolf, Hasan Shahid Ferdous, Joshua Newn, Zaher Joukhadar, Frank Vetere, 'Biometric mirror: Exploring ethical opinions towards facial analysis and automated decision-making' (2019) Proceedings of the 2019 on Designing Interactive Systems Conference, 447.

566 *S and Marper v United Kingdom* (2009) 48 EHRR 50, at 112.

567 Paul Farrell 'Government halting key part of robodebt scheme, will freeze debts for some welfare recipients' ABC News (20 November, 2019) <https://www.abc.net.au/news/2019-11-19/robodebt-scheme-human-services-department-halts-existing-debts/11717188>.

568 Australian Law Reform Commission, *Serious Invasions of Privacy in the Digital Era* (Summary Report 123, June 2014) <https://www.alrc.gov.au/wp-content/uploads/2019/08/summary_report_whole_pdf_.pdf>.

569 United Nations Office of the High Commissioner for Human Rights '"Smart mix" of measures needed to regulate new technologies – Bachelet' (24 April 2019),<https://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=24509>.
570 Administrative Review Council *Administrative Accountability in business areas subject to complex and specific regulation* (November 2008), 5 <https://www.ag.gov.au/LegalSystem/AdministrativeLaw/Documents/publications/report-49.pdf>.
571 Administrative Review Council *Administrative Accountability in business areas subject to complex and specific regulation* (November 2008), 9 <https://www.ag.gov.au/LegalSystem/AdministrativeLaw/Documents/publications/report-49.pdf>.
572 Infographic based on: UK Information Commissioner's Office (ICO) *An overview of the Auditing Framework for Artificial Intelligence and its core components* (26 March 2019) <https://ai-auditingframework.blogspot.com/2019/03/an-overview-of-auditing-framework-for_26.html>.
573 See IP submissions: University of Technology Sydney, 58; LexisNexis, 18; WP submissions: Law Council of Australia, 1; Allens Hub for Technology, Law and Innovation, 5.
574 See IP submissions: University of Technology Sydney, 58; Castan Centre for Human Rights Law, 19.
575 See IP submissions: University of Technology Sydney, 58; Centre for Inclusive Design, 92; Office of the Australian Information Commissioner, 5; Access Now, submission to the WP, 2.
576 See WP submissions, Standards Australia, 3; Australian Information and Industry Association, 3; Microsoft, 3; Digital Industry Group, 4; G Greenleaf, R Clarke, D Lindsay, 4; University of Technology Sydney, submission to the IP, 61.
577 Such as the Institute of Electrical and Electronics Engineers (IEEE), referred to by Castan Centre for Human Rights Law, submission to the IP, 22.
578 See, for example, IP submissions: Chartered Accountants Australia and New Zealand, 23, 3.
579 University of Technology Sydney, submission to the IP, 15. Social Innovation Research Institute, Swinburne University of Technology, submission to the WP, 6.
580 See IP submissions: University of Technology Sydney, 15; Chartered Accountants Australia and New Zealand, 3; Castan Centre for Human Rights Law, 24; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 13.
581 See, for example, N Suzor, K Weatherall, A Daly, A Vromen, M Mann, submission to the IP, 12.
582 See IP submissions: WebKeyIT, 7(b); Confidential, Submission 18, response to question 7(b); Confidential, Submission 2,7(b); S Ramesh 7(b); S Murphy, 7(b); P Henman, 8; Castan Centre for Human Rights Law, 19, 28; EMR Australia, 14.
583 See IP submissions: University of Technology Sydney, 61; WebKeyIT, 7(b); S Murphy, 7(b); Castan Centre for Human Rights Law, 20, 28.
584 See for example: Australian Computer Society, submission to the IP, 1.
585 University of Technology Sydney, submission to the IP, 61.
586 United Nations Office of the High Commissioner for Human Rights, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework* (2011). Endorsed by the Human Rights Council on 16 June 2011 in Resolution 17/4. See Castan Centre for Human Rights Law, submission to the IP, 22. See Chapter 2 (Human Rights).
587 Office of the Australian Information Commissioner, submission to the IP, 4.
588 See, for example, IEEE, *Ethically Aligned Design: A vision for prioritizing human well-being with autonomous and intelligent systems* (First Edition, 2019), 56 <https://ethicsinaction.ieee.org>.
589 Council of Europe *Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data* (January 2017), 4 <https://rm.coe.int/16806ebe7a>.
590 Office of the eSafety Commissioner *Safety by Design* (undated) <https://www.esafety.gov.au/esafety-information/safety-by-design>.
591 See, for example, BSR *Artificial Intelligence: A Rights-Based Blueprint for Business* (August 2018), 13 <https://www.bsr.org/en/our-insights/report-view/artificial-intelligence-a-rights-based-blueprint-for-business>.
592 Gradient Institute 'Helping machines to help us' *Gradient Institute Blog* (6 December 2018) <https://gradientinstitute.org/blog/1/>.
593 See IP submissions: N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 91, 17, 18; Northraine, Submission 84(a), 12; S Ramesh, 7(c); Castan Centre for Human Rights Law, 27; ThinkPlace, 5; University of Technology Sydney, 41. See WP submissions: Australian Information Industry Association, 4; Access Now, 3.
594 Montreal AI Ethics Institute, submission to the WP, 6, 10.
595 See IP submissions: for example, University of Technology Sydney, 14. See WP submissions: NTT Commissions Cloud Infrastructure Services, 2; Royal Australian and New Zealand College of Radiologists, 3.
596 See IP submissions: Australian Women Against Violence Alliance, answer to question 7(c); Centre for Inclusive Design, section VI. See WP submissions: Portable, 1; University of Melbourne, 3; Crighton Nichols, 2.
597 See IP submissions: WebKeyIT, 7(c).
598 See IP submissions: C Nichols, 7(c).
599 See IP submissions: LexisNexis, 21.
600 See IP submissions: LexisNexis, 21; S Murphy, 7(c); Castan Centre for Human Rights Law, 28, 82; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 5, 10.
601 Standards Australia, *Discussion Paper: Developing standards for Artificial Intelligence: Hearing Australia's Voice* (June 2019) <www.standards.org.au>; See also, Standards Australia, submission to the WP, 3.
602 Department of Industry, Innovation and Science *AI Ethics Principles* (2019) <https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework/ai-ethics-principles>.
603 Australian Government, Department of Home Affairs, *Voluntary Code of Practice: A Call for views; Securing the Internet of Things for Consumers* (November 2019), <https://www.homeaffairs.gov.au/reports-and-publications/submissions-and-discussion-papers/code-of-practice>.
604 Digital Transformation Agency, *Digital Service Standard* (10 May 2017) <https://www.dta.gov.au/standard/>. The Standards were identified by the Commonwealth Ombudsman, submission to the IP, 11.
605 IEEE, *Ethically Aligned Design: A vision for prioritizing human well-being with autonomous and intelligent systems* (First Edition, 2019), <https://ethicsinaction.ieee.org>.
606 IEEE, *Ethically Aligned Design: A vision for prioritizing human well-being with autonomous and intelligent systems* (First Edition, 2019), 2, <https://ethicsinaction.ieee.org>.
607 Global Network Initiative, 'Preamble' *Principles on Freedom of Expression and Privacy* (May 2017), <https://globalnetworkinitiative.org/gni-principles/>.
608 See IP submissions: University of Technology Sydney, 37; E Jane and N Vincent, 43; Australian Red Cross, 9; Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University, 9; Australian Centre for Health Law Research, para 24; Roger Clarke, 473; Anonymous, submission 71, 3; Global Partners Digital, answer to question 4; Suzor et al, 11.
609 LexisNexis, submission to the IP, 17.
610 See, for example, The Warren Centre, submission to the IP, 7.
611 See IP submissions: LexisNexis, 14; Deloitte, 6.
612 P Butt, submission to the IP, 5.
613 See IP submissions: K Matthews-Hunt 107, 12; CHOICE, 19; Financial Rights Legal Centre, 19.
614 See, for example, Royal Australian and New Zealand College of Radiology, submission to the IP, 7.
615 See, for example, Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University, submission to the IP, 10, 11.
616 Consumer Policy Research Centre, submission to the IP, 46.
617 *General Data Protection Regulation*, (European Union) 2016/679, Article 42.
618 See IP submissions Deloitte, 6; Google, 2.
619 For discussion on a potential Turing Stamp see IP submissions: E Fries; University of Melbourne, 10; Castan Centre for Human Rights Law, 82; Deloitte, 6.
620 Office of the Australian Information Commissioner, submission to the IP, 9.

621 See, for example, University of Melbourne, submission to the IP, 10; J Powles, M Rizzi, F McGaughey, D Glance, submission to the WP, 2.
622 Castan Centre for Human Rights Law, submission to the IP, 25.
623 There is no universally accepted format of what an impact assessment should contain or even when it should be conducted. For a detailed comparative discussion of impact assessments across a range of jurisdictions and subject areas, see James Harrison and Mary-Ann Stephenson *Human Rights Impact Assessment: Review of Practice and Guidance for Future Assessments* (1 June 2010) (Scottish Human Rights Commission) <www.scottishhumanrights.com>.
624 *Human Rights (Parliamentary Scrutiny) Act 2011* (Cth) s 8; Attorney-General's Department Statements of Compatibility <https://www.ag.gov.au/RightsAndProtections/HumanRights/Human-rights-scrutiny/Pages/Statements-of-Compatibility.aspx>.
625 Office of the Australian Information Commissioner (OAIC), 'A guide to undertaking privacy impact assessments' (5 May 2014) <https://www.oaic.gov.au/privacy/guidance-and-advice/guide-to-undertaking-privacy-impact-assessments/#undertaking-a-pia>.
626 Office of the Australian Information Commissioner, submission to the IP, 4.
627 Government of Canada, *Directive on Automated Decision* Making (1 April 2019) <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>.
628 Government of Canada, 'Appendix B – Impact Assessment Levels' and 'Appendix C – Impact Level Requirements,' *Directive on Automated Decision Making* (1 April 2019) <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>.
629 Microsoft, *Annual Human Rights Report* (2018), 21 <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE2FMZY>.
630 See, for example WP submissions: R Clarke, 23; Montreal AI Ethics Institute, 12; Law Council of Australia, 5. See IP submissions: Office of the Australian Information Commissioner, 5; Castan Centre for Human Rights Law, 6; M O'Sullivan, 14.
631 See IP submissions: LexisNexis, 12; Law Council of Australia, 20.
632 See IP submissions: University of Technology Sydney, 63; Access Now, 5; R Clarke, 12; Data Synergies, 4; Castan Centre for Human Rights Law, 18; Global Partners Digital, question 3(c).
633 United Nations Office of the High Commissioner for Human Rights, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework* (2011), 20. Endorsed by the Human Rights Council on 16 June 2011 in Resolution 17/4.
634 Dillon Reisman, Jason Schultz, Kate Crawford, Meredith Whittaker, *Algorithmic Impact Assessments: Accountability for Framework for Public Agency and Accountability* (April 2018) (AI Now), 4 <https://ainowinstitute.org/aiareport2018.pdf>.
635 See, for example A Mantelero, 'AI and Big Data: A Blueprint for a Human Rights, Social and Ethical Impact Assessment' (2018) 34(4) *Computer Law & Security Review* 754.
636 Dillon Reisman, Jason Schultz, Kate Crawford, Meredith Whittaker, *Algorithmic Impact Assessments: Accountability for Framework for Public Agency and Accountability* (April 2018) (AI Now), 15 <https://ainowinstitute.org/aiareport2018.pdf>.
637 Standards Australia, *Discussion Paper: Developing standards for Artificial Intelligence: Hearing Australia's Voice* (June 2019) <www.standards.org.au>.
638 IEEE, *Ethically Aligned Design: A Vision for Prioritising Human Well-Being with Autonomous and Intelligent Systems* (IEEE, 2nd Ed, 2017), 19 <https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_v2.pdf>.
639 United Nations Human Rights Council, *Resolution 17/4 Human rights and transnational corporations and other business enterprises*, UN Doc A/HRC/RES/17/4 (16 June 2011).
640 See, for example, the work being undertaken by the Department of Industry, Innovation and Science in relation to Artificial Intelligence: <https://www.industry.gov.au/strategies-for-the-future/artificial-intelligence>; and the 'Digital Technology Taskforce', established within the Department of Prime Minister and Cabinet, see Prime Minister of Australia, the Hon Scott Morrison MP, Speech (Business Council of Australia Annual Dinner, Sydney, 21 November 2019).
641 Since 2011, new Bills and certain legislative instruments must be assessed for their compatibility with human rights. A member of the Australian Parliament who introduces a Bill must prepare a Statement of Compatibility, justifying any limitations on individual rights and freedoms. *Human Rights (Parliamentary Scrutiny) Act 2011* (Cth), ss 8, 9(1); *Legislation Act 2003* (Cth), s 15J(2)(f).
642 *Privacy Act 1988* (Cth) s 33D.
643 Under legislation passed in 2018, modern slavery statements—describing the risks of modern slavery in the operations and supply chains, and actions taken to address those risks—must be provided to the Minister by specific entities: *Modern Slavery Act 2018* (Cth), Part 2.
644 See, for example: Microsoft, *Annual Human Rights Report* (2018), 8 <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE2FMZY>; BSR, *Google Celebrity Recognition API Human Rights Assessment*, Executive Summary (October 2019) <https://www.bsr.org/reports/BSR-Google-CR-API-HRIA-Executive-Summary.pdf>; Mark Katonero, *Governing Artificial Intelligence: Upholding Human Rights And Dignity*, Data and Society (2018) 18 <https://datasociety.net/wp-content/uploads/2018/10/DataSociety_Governing_Artificial_Intelligence_Upholding_Human_Rights.pdf>.
645 The Government of Canada has developed an Algorithmic Impact Assessment: Government of Canada *Algorithmic Impact Assessment (AIA)* (May 2019) <https://www.canada.ca/en/government/system/digital-government/modern-emerging-technologies/responsible-use-ai/algorithmic-impact-assessment.html>. A national strategy for meaningful AI produced by a mission assigned by the French Prime Minister Edouard Philippe, recommended a 'Discrimination Impact Assessment' be introduced: C Villani, *For a Meaningful Artificial Intelligence: Towards a French and European Strategy* (A parliamentary mission assigned by the Prime Minister from 8th September 2017 to 8th March 2018), 120 <www.aiforhumanity.fr/en/>.
646 Dillon Reisman, Jason Schultz, Kate Crawford, Meredith Whittaker (AI Now Institute) *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability* (April 2018) <https://ainowinstitute.org/aiareport2018.pdf>.
647 D Dawson, E Schleiger, J Horton, J McLaughlin, C Robinson, G Quezada, J Scowroft and S Hajkowicz, *Artificial Intelligence: Australia's Ethics Framework* (2019) (Data 61 CSIRO, Australia), 8.
648 Department of Industry, Innovation and Science *AI Ethics Principles* (2019) <https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework/ai-ethics-principles>.
649 Financial Conduct Authority (United Kingdom), *Regulatory sandbox lessons learned report* (20 October 2017), 3 <https://www.fca.org.uk/publications/research/regulatory-sandbox-lessons-learned-report>.
650 Jacob Turner, *Robot Rules: Regulating Artificial Intelligence* (Palgrave MacMillan, 2019).
651 Jorge Gabriel Jimenez and Margaret Hagan, *A regulatory sandbox for the industry of Law* (March 2019) (Legal Design Lab and Legal Executive Institute), 2 <http://www.legalexecutiveinstitute.com/wp-content/uploads/2019/03/Regulatory-Sandbox-for-the-Industry-of-Law.pdf>.
652 Infocomm Media Development Authority, Singapore Government *Guide to Data Sharing* (undated, accessed November 2019), 3 <https://www.imda.gov.sg/programme-listing/data-collaborative-programme>.
653 Reserve Bank of India, *Enabling Framework for Regulatory Sandbox* (13 August 2019), 6.1.2.
654 Ministry of Economic Affairs and Employment of Finland, *Leading the way into the age of artificial intelligence, Final report of Finland's Artificial Intelligence Programme 2019 (2019:41)* (2019), 56, <https://julkaisut.valtioneuvosto.fi/handle/10024/161688>.
655 European Securities and Markets Authority, European Banking Authority, European Insurance and Occupational Pensions Authority *Report, FinTech: Regulatory sandboxes and innovation hubs* (2018) (JC 2018 74) <https://www.esma.europa.eu/sites/default/files/library/jc_2018_74_joint_report_on_regulatory_sandboxes_and_innovation_hubs.pdf>.
656 Australian Securities and Investments Commission, *Regulatory Guide 257: Testing fintech products and services without holding an AFS or credit licence* (August 2017) <https://download.asic.gov.au/media/4420907/rg257-published-23-august-2017.pdf>. See also

the recent report of the Australian Energy Market Commission draft recommendation to the Council of Australian Governments Energy Council regarding the introduction of regulatory sandboxes to support innovative energy projects and effectively manage risk: Australian Energy Market Commission, *Draft Report: Regulatory Sandbox Arrangements to Support Proof-of-concept trials* (11 July 2019) <https://www.aemc.gov.au/news-centre/media-releases/aemc-recommends-regulatory-sandbox-toolkit-support-innovative-trials>.

657 Australian Securities and Investment Commission *Regulatory Guide 257, Testing fintech products and services without holding an AFS or credit licence* (August 2017), 15 <https://download.asic.gov.au/media/4420907/rg257-published-23-august-2017.pdf>.

658 See Consultative Group to Assist the Poor and World Bank Group, *Regulatory Sandbox Global Survey* (July 2019) <https://www.findevgateway.org/library/cgap-world-bank-regulatory-sandbox-global-survey-2019>.

659 See, for example, Financial Conduct Authority (United Kingdom), *Regulatory sandbox lessons learned report* (20 October 2017), 12, 13 <https://www.fca.org.uk/publications/research/regulatory-sandbox-lessons-learned-report>.

660 See, for example, Information Commissioner's Office (United Kingdom) *ICO opens Sandbox beta phase to enhance data protection and support innovation* (29 March 2019) <https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/03/ico-opens-sandbox-beta-phase-to-enhance-data-protection-and-support-innovation/>.

661 Centre for Information Policy Leadership, *Regulatory Sandboxes in Data Protection: Constructive Engagement and Innovative Regulation in Practice* (8 March 2019) <https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_on_regulatory_sandboxes_in_data_protection_-_constructive_engagement_and_innovative_regulation_in_practice__8_march_2019_.pdf>.

662 See, for example, Katerina Yordanova 'The shifting sands of regulatory sandboxes for AI', *KU Leuven Blog* <https://www.law.kuleuven.be/citip/blog/the-shifting-sands-of-regulatory-sandboxes-for-ai>.

663 See IP, including: Castan Centre for Human Rights Law, 17; University of Technology Sydney, 60. See also WP submissions: L Young, 1; R Lewis, 4.

664 Law Council of Australia, submission to the IP, 19.

665 CBS News, 'CEO Zuckerberg: Facebook's 5 core values' (17 May 2012) <https://www.cbsnews.com/news/ceo-zuckerberg-facebooks-5-core-values/>.

666 Bharath Ganesh (Researcher, Oxford Internet Institute) 'The Silicon Valley mantra of 'move fast and break things' has fallen miserably short' *New Statesman* (26 April 2018) <https://tech.newstatesman.com/guest-opinion/move-fast-break-things-mantra>.

667 For example, Microsoft's AI chatbot 'Tay' started tweeting discriminatory and hateful speech within 24 hours of being tested. See Elle Hunt 'Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter,' *The Guardian* (24 March 2016) <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter?CMP=twt_a-technology_b-gdntech>.

668 The labelling of African Americans as 'gorillas' by the Google Photo app was identified by users once the app had been launched: Jana Kasperkevic 'Google says sorry for racist auto-tag in photo app', *The Guardian* (2 July 2015), <https://www.theguardian.com/technology/2015/jul/01/google-sorry-racist-auto-tag-photo-app>.

669 Australian Securities and Investments Commission, 'Fintech regulatory sandbox' (Webpage), <https://asic.gov.au/for-business/innovation-hub/fintech-regulatory-sandbox/>. See also, the Regulatory Sandboxes set up by NSW Government: Innovation NSW, 'Regulatory Sandboxes' (Webpage, NSW Government), <https://www.innovation.nsw.gov.au/regulatorysandboxes>.

670 OECD, *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449), [2.2], <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

671 A related issue relevant to education and training is the need to re-train workers who are likely to have their employment disrupted or displaced as a result of the rise of AI. While that issue is beyond the scope of this Project, it was raised by a number of submissions. See, for example, WP submissions: Australian Services Union; Australian Information and Industry Association, 4; and IP submission: Chartered Accountants Australia and New Zealand 23, 2.

672 Noted in IP submissions: Law Council of Australia; University of Technology Sydney, 3.4.5; Royal Australian and New Zealand College of Radiologists, 7; Chartered Accountants Australia and New Zealand, 2; R McLay, 10, 11.

673 Noted in IP submissions: University of Technology Sydney, 9; National Association of Community Legal Centres, 10; WP submissions: NTT Communications Cloud Infrastructure Services Inc, 5; R Chalmers, 2; H Dobson, 4; R Lewis, 5; Australian Information and Industry Association, 1; Izerobzero, 5; G Greenleaf, R Clarke, D Lindsay, 2.

674 See WP submissions: R Newman Knake, 2; NTT Communications Cloud Infrastructure Services Inc, 5.

675 See IP submissions: I Law; E Salazar, J Parker Roberto, L Gelman, A Gajo; See also R Chalmers, submission to the WP, 2.

676 IP submissions: University of Technology Sydney, 61; LexisNexis, 1, 9, 18; Emergency Services Levy Insurance Monitor, 37; Law Council of Australia, 19; University of Technology Sydney, 34; PwC Indigenous Consulting, 12; M Guihot and M Rimmer, submission to the WP, 3.

677 IP submissions: University of Technology Sydney, 32; E Salazar, J Parker Roberto, L Gelman, A Gajo, 18-19; K Mathews-Hunt; Migrant Worker Justice Initiative, UNSW, University of Technology Sydney; Australian Red Cross, 2; WP submissions: K Matthews-Hunt; 156; R Chalmers, 2; Centre for Policy Futures, University of Queensland, 6; Social Innovation Research Institute, Swinburne University, 3; R Lewis, 5; Law Council of Australia, 5; Montreal AI Ethics Institute, 8; R Clarke, 16; University of Melbourne, 7; G Greenleaf, R Clarke, D Lindsay, 2.

678 See, for example, WP submissions: Office of the eSafety Commissioner; R Lewis, 5; Allens Hub for Technology, Law and Innovation, 7.

679 See, for example, The IEEE *Ethically Aligned Design: A vision for prioritizing human well-being with autonomous and intelligent systems* (First Edition, 2019), 56 <https://ethicsinaction.ieee.org>. Matthew Beard and Simon Longstaff, *Ethical by Design: Principles for Good Technology* (The Ethics Centre, 2018), 51.

680 The IEEE Global Initiative Ethics of Autonomous and Intelligence Systems, *Ethically Aligned Design: A Vision for Prioritising Human Well-Being with Autonomous and Intelligent Systems* (IEEE, 2nd Ed, 2017), 31.

681 See, for example, 'Professional regulation of legal practitioners, Professional regulation in Australia at present' (Webpage, Australian Institute of Criminology, 2019) <https://aic.gov.au/publications/rpp/rpp122-part2/professional-regulation-australia-present>; Insurance Council of Australia, *General Insurance Code of Practice 2014*, <http://www.insurancecouncil.com.au/for-consumers/code-of-practice>. Australian Banking Association. Banking Code of Practice,1 July 2019, <https://www.ausbanking.org.au/wp-content/uploads/2019/06/Banking-Code-of-Practice-2019-web.pdf>.

682 See, for example IEEE, *Ethically Aligned Design: A vision for prioritizing human well-being with autonomous and intelligent systems* (First Edition, 2019), 31<https://ethicsinaction.ieee.org>.

683 This issue was raised by a number of stakeholders to the IP; see, for example, WebKeyIT, question 3(b).

684 The first general principles in the IEEE Guidelines is that autonomous and intelligent systems 'shall be created to respect, promote and protect internationally recognised human rights': IEEE, *Ethically Aligned Design: A vision for prioritizing human well-being with autonomous and intelligent systems* (First Edition, 2019), 56 <https://ethicsinaction.ieee.org>.

685 See, for example Michael Veale, Max Van Kleek, Reuben Binns, 'Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making' (CHI 2018 Paper, April 21-26, Montreal), 9.

686 Tiberio Caetano and Bill Simpson-Young, "Artificial Intelligence: Australia's Ethics Framework" CSIRO's Data61 discussion paper' (Gradient Institute, 2019).

687 A current guide for automated decision-making was published in 2004: Administrative Review Council *Automated Assistance in Administrative Decision Making* (Report no. 46, November 2004), <https://www.ag.gov.au/LegalSystem/AdministrativeLaw/Pages/publications/report-46.aspx>.

688 N Suzor, K Weatherall, A Daly, A Vromen, M Mann, submission to the IP, 12.
689 See, for example Justice R French, 'Judicial review rights' *AIAL* Forum, no. 28 (2011), 30, cited in Administrative Review Council *Automated Assistance in Administrative Decision Making, Report to the Attorney-General* (Report no. 46) (November 2004), vii <https://www.arc.ag.gov.au/Documents/AAADMreportPDF.pdf>.
690 Toby Walsh, Neil Levy, Genevieve Bell, Anthony Elliott, James Maclaurin, Iven Mareels and Fiona Wood, *The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve our Wellbeing* (Report for the Australian Council of Learned Academies, 2019), 81, <https://acola.org/hs4-artificial-intelligence-australia/>.
691 See, for example, Justice Melissa Perry '*i*Decide: Administrative Decision-Making in the Digital World' (2017) 91 *Australian Law Journal* 29, 30; Toby Walsh, Neil Levy, Genevieve Bell, Anthony Elliott, James Maclaurin, Iven Mareels and Fiona Wood, *The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve our Wellbeing* (Australian Council of Learned Academies, 2019), 81.
692 The Law Society of New South Wales *The Future of Law and Innovation in the Profession (FLIP) (Commission of Inquiry)* (2017), 5, 7 <https://www.lawsociety.com.au/sites/default/files/2018-03/1272952.pdf>.
693 See, for example, the work being undertaken by the Australian Institute of Company Directors and the World Economic Forum: Australian Institute for Company *Directors, Directors, Is Your Business Ready for AI?* (1 June 2019) <https://aicd.companydirectors.com.au/membership/company-director-magazine/2019-back-editions/june/ai-ethics>.
694 World Economic Forum and Ipsos, 'Evolving Technology: A Global Survey on Data Privacy/Security and Automation/Artificial Intelligence' (2019).
695 IEEE, *Ethically Aligned Design: A Vision for Prioritising Human Well-Being with Autonomous and Intelligent Systems* (IEEE, 2nd Ed, 2017), 31.
696 Castan Centre for Human Rights Law, Monash University, submission to the IP, 35.
697 University of Technology Sydney, submission to the IP, 81.
698 Erica Southgate, Karen Blackmore, Stephanie Pieschl, Susan Grimes, Jessie McGuire, Kate Smithers *Research Report: Artificial Intelligence and Emerging Technologies in Schools* (August 2019) (Commissioned by the Australian Government Department of Education) <https://docs-edu.govcms.gov.au/system/files/doc/other/aiet_final_report_august_2019.pdf>; Toby Walsh, Neil Levy, Genevieve Bell, Anthony Elliott, James Maclaurin, Iven Mareels and Fiona Wood, *The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve our Wellbeing* (Australian Council of Learned Academies, 2019).
699 Australian Human Rights Commission, Submission to the Senate Standing Environment and Communications References Committee, *Inquiry into harm being done to Australian children through access to pornography on the Internet* (April 2016); Australian Human Rights Commission, *Cyberbullying and the Bystander* (Report, 2012). The Australian Curriculum addresses digital technologies, but in a manner that is fairly basic and which has not kept pace with emerging technological developments. Computer technology has not been cohesively integrated into primary-level education, despite the fact that most children in Australia have regular access to digital devices prior to commencing school. Education about AI and its implications for human rights would only be effective if digital literacy was strengthened more broadly in existing primary and secondary school curricula.
700 Lyria Bennett Moses, *Helping future citizens navigate an automated, datafied world (Education: Future Frontiers, Occasional Papers Series)* (A paper commissioned by the Department of Education) (April 2019), 16 <https://education.nsw.gov.au/our-priorities/innovate-for-the-future/education-for-a-changing-world/media/documents/future-frontiers-education-for-an-ai-world/Helping-Future-Citizens-Lyria-Bennett-Moses.pdf>.
701 See, for example, PwC Indigenous Consulting, submission to the IP, 14.
702 Australian National Audit Office *A guide to conducting performance audits* (31 March 2017) <https://www.anao.gov.au/work/corporate/guide-to-conducting-performance-audits>.
703 Australian National Audit Office *Implementation of the MyHealth Record System* (25 November 2019) <https://www.anao.gov.au/work/performance-audit/implementation-the-my-health-record-system>.
704 Administrative Review Council *Automated Assistance in Administrative Decision Making, Report to the Attorney-General* (Report No. 46) (November 2004), Principles 18 and 19, 14, 43 <https://www.arc.ag.gov.au/Documents/AAADMreportPDF.pdf>.
705 Commonwealth Ombudsman, submission the IP, 11.
706 AI Now Institute, *AI Now Report 2018* (December 2018), Recommendations 1 and 3, 4. See also, Petra Molnar and Lex Bill (The Citizen Lab, Munk School of Global Affairs and Public Policy); *Bots at the gate: A human rights analysis of automated decision-making in Canada's immigration and refugee system* (University of Toronto, 2018), 63.
707 The IEEE, *Ethically Aligned Design: A vision for prioritizing human well-being with autonomous and intelligent systems* (First Edition, 2019), 7, <https://ethicsinaction.ieee.org>.
708 Data and Society *Algorithmic Accountability: A Primer* (Prepared for the Congressional Progressive Caucus: 'Tech Algorithm Briefing: How Algorithms Perpetuate Racial Bias and Inequality) (18 April 2018, Washington DC), 2, 7; AI Now Institute *AI Now Report 2018* (December 2018), 4.
709 Data and Society, *Algorithmic Accountability: A Primer* (Prepared for the Congressional Progressive Caucus: 'Tech Algorithm Briefing: How Algorithms Perpetuate Racial Bias and Inequality' (18 April 2018, Washington DC), 8.
710 See IP submissions: University of Technology Sydney, 34; CHOICE, 8; Consumer Policy Research Centre, 46 R Clarke, 12; Castan Centre for Human Rights Law, 18; Ethics Centre, 45.
711 See IP submissions: University of Technology Sydney, 48; LexisNexis, 22; See WP submissions: Remi AI, 6; R Clarke, 6.
712 See, for example, CHOICE, submission to the IP, 20.
713 See, for example, P Butt, submission to the IP, 10.
714 LexisNexis, submission to the IP, 13.
715 University of Technology Sydney, submission to the IP, 48.
716 See, for example *Audit-AI*, noted in Pymetrics, submission to the IP, 2. Technology company Northraine has an auditory tool, see Northraine, submission to the IP, 4. O'Neil Risk Consulting & Algorithmic Auditing, <http://www.oneilrisk.com/>.
717 LexisNexis, submission to the IP, 22.
718 S Murphy, submission to the IP, question 7.
719 See, for example, IP submissions: University of Technology Sydney, 15; LexisNexis, 13; Public Interest Advocacy Centre, 13; N Suzor, K Weatherall, A Daly, A Vromen, M Mann, 18; Law Council of Australia, 23; Data Synergies, 4.
720 See Allens and Linklaters *AI Toolkit: Ethical, Safe, Lawful (Practical guidance for AI projects)* (May 2019), 45 <https://www.allens.com.au/globalassets/pdfs/campaigns/report-ai-toolkit_may19.pdf>.
721 Government of Canada, *Directive on Automated Decision-Making* (February 2019) <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>.
722 See, eg the *Commonwealth Procurement Rules* (20 April 2019), made under s 105B(1) *Public Governance, Performance and Accountability Act 2013* (Cth) which guide Australian Government agencies in their procurement activities and are made by the Minister for Finance.
723 Senator the Hon Mathias Cormann, Minister for Finance and Public Service, 'Foreword' *Commonwealth Procurement Rules* (20 April 2019) 3,<https://www.finance.gov.au/sites/default/files/2019-11/CPRs-20-April-2019_1.pdf>.
724 The *Government Procurement (Judicial Review) Act 2018* (Cth) came into force on 19 October 2018. It provides for procurement processes to be challenged in the Federal Court of Australia or the Federal Circuit Court on the basis of failure to comply with the *Commonwealth Procurement Rules* or proposed conduct that would breach the Rules.
725 Australian Government, Digital Transformation Agency *Digital Sourcing Framework for ICT procurement* (undated, accessed November 2019) <https://www.dta.gov.au/help-and-advice/ict-procurement/digital-sourcing-framework-ict-procurement>.

726 Australian Government, Innovation and Science Australia, 'Strategic Opportunity 3.4' *Prosperity through innovation, A plan for Australia to thrive in the global innovation race* (2017), 63 <https://www.industry.gov.au/sites/g/files/net3906/f/May%202018/document/pdf/australia-2030-prosperity-through-innovation-full-report.pdf>.
727 House of Lords Select Committee on Artificial Intelligence, *AI in the UK: ready, willing and able?* Report of Session 2017-19 (16 April 2018), 6, 66 <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>.
728 World Economic Forum *Unlocking public sector artificial intelligence* (undated), <https://www.weforum.org/projects/unlocking-public-sector-artificial-intelligence>.
729 See IP submissions: University of Technology Sydney, 70; LexisNexis, 13; Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University, 11; University of Melbourne, 10; Castan Centre for Human Rights Law, 34; Global Partners Digital, 11.
730 Global Partners Digital, submission to the IP, 10.
731 See IP submissions: Castan Centre for Human Rights Law, 5.5.1.; Society on Social Impacts of Technology, IEEE and Law Futures Centre, Griffith University, 11.
732 LexisNexis, submission to the IP, 13.
733 University of Technology Sydney, submission to the IP, 41.
734 Castan Centre for Human Rights Law, submission to the IP, 34.
735 See IP submissions: University of Melbourne; 10; Castan Centre for Human Rights Law, 34.
736 In the WP, this potential body was referred to as a 'Responsible Innovation Organisation'. Stakeholders used terms such as 'organisation', 'body' and 'leadership' in submissions to refer to the potential entity.
737 'Governance', as defined in chapter 3, refers to the overarching system that supports regulation and oversight, including regulators, dispute resolution bodies, courts, industry and professional bodies, among others. Governance also refers to the underpinnings of the regulatory and oversight system in law, policy and education.
738 See WP submissions: G Greenleaf, R Clarke, D Lindsay, 4; Chartered Accountants Australia and NZ, 1; Australian Council of Trade Unions, 4; M Smith, 3; Australian Human Rights Institute, 5; NTT Communications Cloud Infrastructure Services, 1; Effective Altruism, 9; Royal Australian and New Zealand College of Radiologists, 4; Swinburne Social Innovation Research Institute, 4; Portable, 3; Izerobzero , 3; H Dobson, 2; Montreal AI Ethics Institute, 4; Centre for Policy Futures, University of Queensland, 2; Confidential submission, 52; Also see IP submissions: Deloitte, 6; University of Technology Sydney, 59; S Carter, N Houssami, W Rogers, 3; Newman Knake, 2; Northraine, 4; Access Now, 6; National Archives of Australia.
739 See for example, Deloitte, submission to the IP, 6.
740 See WP submissions: Centre for Policy Futures, University of Queensland, 3; J Ford; Royal Australian and New Zealand College of Radiologists, 4; Confidential submission 14, 2.
741 See WP submissions: Access Now, 6; Swinburne Social Innovation Research Institute, 3; Remi AI, 5; Portable, 1, 2.
742 Access Now, submission to the WP, 6.
743 NTT Communications Cloud Infrastructure Services, submission to the WP, 6.
744 See for example, Office of the Australian Information Commissioner, submission to the WP, 3.
745 See for example, University of Technology Sydney, submission to the IP, 93; Law Council of Australia, submission to the WP, 3.
746 See, for example, T Metzinger, 'Towards a Global Artificial Intelligence Charter' in *Should we fear artificial intelligence?* (European Parliamentary Research Service, 2018) 27, 31-32.
747 Marchant and Wallach (unpublished) and cited in T Metzinger, 'Towards a Global Artificial Intelligence Charter' *Should we fear artificial intelligence?* (European Parliamentary Research Service, 2018), 27, 32.
748 See WP submissions: Standards Australia, 1; Ethics for AI and ADM, 3; M Guihot and M Rimmer, 3; Australian Information Industry Association Australian Information and Industry Association; Digital Industry Group, 4; Chalmers; 3, 4; Microsoft, 3.
749 See for example Australian Industry Group, submission to the WP, 7.
750 See for example Australian Information Industry Association, submission to the WP, 9.
751 Confidential submission 14 to the WP, 2.
752 Effective Altruism ANZ, submission to the WP, 11.
753 Confidential submission 14 to the WP, 2.
754 Access Now, submission to the WP, 6.
755 See WP submissions: Portable, 3; Consumer Policy Research Centre, 3; Confidential submission 52.
756 J Ford, submission to the WP, 13.
757 Confidential submission 52 to the WP.
758 Confidential submission 52 to the WP.
759 See WP submissions: Access Now, 6; Swinburne Social Innovation Research Institute, 4.
760 Swinburne Social Innovation Research Institute, submission to the WP, 4.
761 FutureLab.Legal, submission to the WP, 3, 4.
762 H Wilkinson, submission to the WP, 6.
763 Confidential submission 52 to the WP, 4.
764 Confidential submission 52 to the WP.
765 See WP submissions: Portable, 3; Northraine, 4; NTT Communications Cloud Infrastructure, WP 1.
766 Consumer Policy Research Centre, submission to the WP, 3.
767 Northraine, submission to the WP, 4.
768 See WP submissions: Australian Human Rights Institute, 6; Confidential submission 52.
769 See WP submissions: Remi AI, 3-4; Portable, 3; Confidential submission 52.
770 See WP submissions: Northraine, 4; Confidential submission 52.
771 Portable, submission to the WP, 3, 4.
772 See WP submissions: Northraine, 4; Data 2X and the UN Foundation, 1; Portable, 3, 4; Confidential submission 52 to the WP.
773 Portable to WP 3, 4; C Nichols 5; Chartered Accountants Australia and New Zealand, 2.
774 See WP submissions: University of Melbourne, 8; Swinburne Social Innovation Research Institute, 4; University of Technology Sydney, submission to the IP, 95.
775 Swinburne Social Innovation Research Institute, submission to the WP, 4.
776 S Carter, N Houssami, W Rogers, submission to the WP, 6.
777 J Ford, submission to the WP, 13-14.
778 Royal Australian and New Zealand College of Radiologists, submission to the WP, 4.
779 See WP submissions: G Greenleaf, R Clarke, D Lindsay, 5; Remi AI, 4, 5; J Ford, 14; S Carter, N Houssami, W Rogers, 6; Queensland Council of Civil Liberties, 6; Remi AI 4, 5.
780 H Wilkinson, submission to the WP, 8.
781 C Nichols, submission to the WP, 6.
782 Australian Human Rights Institute, submission to the WP, 6.
783 Northraine, submission to the IP, 5.
784 Montreal AI Ethics Institute, submission to the WP, 11, 12.
785 Peter Engelke, 'Three ways the Fourth Industrial Revolution is shaping geopolitics,' *World Economic Forum* (24 August 2018) <https://www.weforum.org/agenda/2018/08/three-ways-the-fourth-industrial-revolution-is-shaping-geopolitics/>.
786 Eleonore Pauwels, 'The New Geopolitics of Converging Risks: The UN and Prevention in the Era of AI' (United Nations University Centre for Policy Research, 2019) <https://cpr.unu.edu/the-new-geopolitics-of-converging-risks-the-un-and-prevention-in-the-era-of-ai.html>.
787 Tim Dutton, Brent Barron, and Gaga Boskovic, *Building an AI world: Report on National and Regional AI Strategies*, (CIFAR, Canada, November, 2018), 8 <https://www.cifar.ca/docs/default-source/ai-society/buildinganaiworld_eng.pdf>.
788 Dutton T, B Barron and G Boskovic, *Building an AI world: Report on National and Regional AI Strategies*, (CIFAR, Canada, November, 2018), 9 <https://www.cifar.ca/docs/default-source/ai-society/buildinganaiworld_eng.pdf>.
789 Eleonore Pauwels, *The New Geopolitics of Converging Risks: The UN and Prevention in the Era of AI*, (United Nations University Centre for Policy Research, 29 April 2019), 39 <https://i.unu.edu/media/cpr.unu.edu/attachment/3472/PauwelsAIGeopolitics.pdf>.

790 Eleonore Pauwels, *The New Geopolitics of Converging Risks: The UN and Prevention in the Era of AI*, (United Nations University Centre for Policy Research, 29 April 2019), 39 <https://i.unu.edu/media/cpr.unu.edu/attachment/3472/PauwelsAIGeopolitics.pdf>.
791 OECD, OECD *AI Policy Observatory, A platform for AI information, evidence and policy options* (September 2019) <http://www.oecd.org/going-digital/ai/about-the-oecd-ai-policy-observatory.pdf>.
792 Centre for Data Ethics and Innovation 'About Us' (United Kingdom Government), <https://www.gov.uk/government/organisations/centre-for-data-ethics-and-innovation/about>.
793 Centre for Data Ethics (CDEI), *Independent Report: Centre for Data Ethics (CDEI) 2 Year Strategy* (United Kingdom Government, 21 March 2019) <https://www.gov.uk/government/publications/the-centre-for-data-ethics-and-innovation-cdei-2-year-strategy/centre-for-data-ethics-cdei-2-year-strategy>.
794 Toby Walsh, Neil Levy, Genevieve Bell, Anthony Elliott, James Maclaurin, Iven Mareels and Fiona Wood, *The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve our Wellbeing* (Report for the Australian Council of Learned Academies, 2019), 13, <https://acola.org/hs4-artificial-intelligence-australia/>.
795 Toby Walsh, Neil Levy, Genevieve Bell, Anthony Elliott, James Maclaurin, Iven Mareels and Fiona Wood, *The effective and ethical development of artificial intelligence: An opportunity to improve our wellbeing*, 13. Report for the Australian Council of Learned Academies <www.acola.org>.
796 SA Hajkowicz, S Karimi, T Wark, C Chen, M Evans, N Rens, D Dawson, A Charlton, T Brennan, C Moffatt, S Srikumar, KJ Tong, *Artificial intelligence: Solving problems, growing the economy and improving our quality of life* (CSIRO Data61, 2019), 50.
797 See for example IP submissions: Access Now, 6; University of Technology Sydney, 93. See also WP submissions: Australian Council of Trade Unions, 4; Portable, 3; G Greenleaf, R Clarke, D Lindsay, 4; M Smith, 3; J Ford, 1213; Swinburne Social Innovation Research Institute University, 4; Royal Australian and New Zealand College of Radiologists, 4; Chartered Accountants Australia and NZ to WP 1; Australian Human Rights Institute, 4, 5; Effective Altruism ANZ; S Carter, N Houssami, W Rogers, 3; Confidential submission 14, 2, 3.
798 See WP submissions: Consumer Policy Research Centre, 5; University of Melbourne, 9; Montreal AI Ethics Institute, 15; Northraine, 6; Allens Hub for Technology. Law and Innovation, 8; Law Council of Australia, 5.
799 See, for example WP submissions: University of Melbourne, 4, 5, 6; Australian Council of Trade Unions; NTT Communications Cloud Infrastructure Services, 6; Centre for Policy Futures, University of Queensland, 6; M Guihot and M Rimmer, 3; Law Council of Australia, 6; University of Melbourne, 6; WP Symposium, 6 March 2019.
800 See for example, the following WP submissions: University of Melbourne, 6; Centre for Policy Futures, 6; Law Council of Australia, 6.
801 University of Melbourne, submission to the WP, 6.
802 University of Melbourne, submission to the WP, 4.
803 Ethics for AI and ADM, submission to the WP, 3.
804 M Guihot and M Rimmer, submission to the WP, 3.
805 Centre for Policy Futures, University of Queensland, submission to the WP, 3, 4.
806 For example, see WP submissions: J Powles, M Rizzi, F McGaughey, D Glance, 1, 2; G Greenleaf, R Clarke, D Lindsay, 1; Allens Hub for Technology, Law and Innovation, 7. WP Symposium, 6 March 2019.
807 WP Symposium, 6 March 2019.
808 WP Symposium, 6 March 2019.
809 See for example, WP submissions: Nichols to WP 6; Consumer Policy Research Centre, 3; S Moore, 2; Australian Council of Trade Unions, 4, 5.
810 Office of the eSafety Commissioner, submission to the WP, 3.
811 See WP submissions: Consumer Policy Research Centre, 3; Effective Altruism ANZ, 12.
812 See WP submissions: Swinburne Social Innovation Research Institute, 2; R Lewis, 4; Montreal AI Ethics Institute, 13, 14; NTT Communications Cloud Infrastructure Services Inc, 7; S Carter, N Houssami, W Rogers, 4, 5; H Dobson, 4; H Wilkinson, 1.
813 Australian Council of Trade Unions, submission to the WP, 4, 5.
814 Australian Services Union, submission to the WP, 3.
815 See WP submissions: Swinburne Social Innovation Research Institute, 5; Confidential submission 14, 3; H Dobson to WP 4; Consumer Policy Research Centre, 3; Northraine, 6; Royal Australian and New Zealand College of Radiologists, 2; Australian Human Rights Institute, 8.
816 See WP submissions: H Wilkinson, 2, 3; J Ford, 15; Chartered Accountant Australia and New Zealand, 2; confidential submission 14, 3.
817 University of Technology Sydney, submission to IP, 8.
818 See WP submissions: Centre for Policy Futures, 6; Law Council of Australia, 5, 6; H Dobson, 5; Consumer Policy Research Centre, 4; Australian Human Rights Institute, 8.
819 IP subs: Law Council of Australia, 3; LexisNexis, 22. WP subs: R Lewis, 4, 5; ; Australian Human Rights Institute8; Montreal AI Ethics Institute, 15.
820 See WP submissions: Northraine, 6; Law Council of Australia, 5.
821 Lewis to WP 5; University of Technology Sydney to IP, 95.
822 See WP submissions: Consumer Policy Research Centre, 5; University of Melbourne, 6, 7; Remi AI, 2, 5; Montreal AI Ethics Institute, 14; H Dobson, 5; Swinburne Social Innovation Research Institute, 5; Allens Hub for Technology, Law and Innovation, 8.
823 See WP submissions, Swinburne Social Innovation Research Institute, 5; University of Melbourne to, 6, 7; Australian Human Rights Institute, 8; Remi AI, 5.
824 Consumer Policy Research Centre, submission to WP, 5.
825 LexisNexis, submission to the IP 22; WP submissions: Swinburne Social Innovation Research Institute, 5; University of Melbourne, 6, 7; Australian Human Rights Institute, 8; Remi AI, 5; Consumer Policy Research Centre, 4, 5.
826 WP submissions: Royal Australian and New Zealand College of Radiologists, 4-5; S Carter, N Houssami, W Rogers, 3.
827 WP submissions: Consumer Policy Research Centre, 4; Australian Human Rights Institute, 8; Law Council of Australia, 6; Lewis, 4, 5; H Dobson to WP 5; Submissions to the IP: University of Technology Sydney, 96; LexisNexis, 22.
828 See for example, WP submissions: Montreal AI Ethics Institute, 14; Centre for Policy Futures, University of Queensland, 6; Law Council of Australia, 3; Consumer Policy Research Centre, 3, 4; Queensland Council of Civil Liberties; 6; Swinburne Social Innovation Research Institute, 5; R Lewis, 4, 5; H Dobson, 5.
829 R Lewis, submission to the WP, 6.
830 See WP submissions: Northraine, 6; Swinburne Social Innovation Research Institute, 5; Consumer Policy Research Centre, 4; Law Council of Australia, 3; Montreal AI Ethics Institute, 15.
831 See WP submissions: University of Melbourne, 6, 7; Consumer Policy Research Centre, 4; Queensland Council of Civil Liberties, 6; LexisNexis, 22; Centre for Policy Futures, University of Queensland, 6; Allens Hub for Technology, Law and Innovation, 8.
832 Law Council of Australia, submission to the WP, 6.
833 University of Technology Sydney, submission to the IP, 96.
834 University of Technology Sydney sub to the IP, 98; J Evans to WP, 2.
835 See WP submissions: H Dobson, 4; Effective Altruism ANZ, 12; Confidential submissions 55; Confidential submission 14, 3.
836 See WP submissions: H Dobson 4; Confidential submission 14, 3.
837 Confidential submission 14 to WP, 3.
838 See WP submissions: Effective Altruism ANZ, 12; Confidential submission 55.
839 See WP submissions: Centre for Policy Futures, 6; H Dobson, 5; J Ford, 16; G Greenleaf, R Clarke, D Lindsay, 5.
840 Law Council of Australia, submission to the WP, 4.
841 See WP submissions: H Wilkinson, 8; University of Melbourne, 6-7.
842 H Dobson, submission to the WP, 5.
843 Swinburne Social Innovation Research Institute, submission to the WP, 5.
844 Remi AI, submission to the WP, 5.
845 See WP submissions: C Nichols, 7; G Greenleaf, R Clarke, D Lindsay, 6; Effective Altruism ANZ,13; Law Council of Australia, 3, 4; H Dobson, 6, 9; University of Melbourne, 9; Remi AI, 5; Queensland Council of Civil Liberties, 6-7; Northraine, 6; Swinburne Social Innovation Research Institute University, 6; Australian Human Rights Institute, 9; Montreal Ethics, 16, 17; Lexis Nexis submission to the IP 2.

846 See WP submission: G Greenleaf, R Clarke, D Lindsay, 6; Effective Altruism ANZ, 13; University of Melbourne, 9; Queensland Council of Civil Liberties, 7; Northraine, 6; Swinburne Social Innovation Research Institute, 6; Australian Human Rights Institute, 9.
847 See IP submissions: G Greenleaf, R Clarke, D Lindsay, 6; University of Melbourne, 9.
848 LexisNexis, submission to the IP, 21.
849 See WP submissions: Effective Altruism ANZ,13.
850 See for example, WP submissions: C Nichols, 7; Northraine, 6.
851 C Nichols, submission to the WP, 7.
852 See WP submissions: Swinburne Social Innovation Research Institute, 6; Queensland Council of Civil Liberties, 7.
853 Northraine, submission to the WP, 6.
854 See WP submissions: Montreal AI Ethics Institute, 16,17; G Greenleaf, R Clarke, D Lindsay, 5, 6; Confidential submission 14.
855 Queensland Council of Civil Liberties, submission to the WP, 6.
856 Confidential, submission 14 to the WP, 3.
857 Centre for Policy Futures, University of Queensland, submission to the WP, 6.
858 See WP submissions: H Dobson; Queensland Council of Civil Liberties, 6; NTT Communications Cloud Infrastructure Services, 8.
859 See WP submissions: Australian Human Rights Institute; University of Melbourne, 9; Swinburne Social Innovation Research Institute; J Evans; Portable.
860 Portable, submission to the WP, 1, 2.
861 See WP submissions, University of Melbourne, 6-9; Australian Human Rights Institute, 8; Swinburne Social Innovation Research Institute, 5, 6; Effective Altruism ANZ, 10.
862 Swinburne Social Innovation Research Institute University, 5, 6; Australian Human Rights Institute, 8; Swinburne Social Innovation Research Institute, 5, 6.
863 See WP submissions: H Dobson, 7; Australian Human Rights Institute, 8; Swinburne Social Innovation Research Institute, 5, 6; J Evans, 2.
864 See for example, WP submissions: University of Melbourne, 6, 7; Swinburne Social Innovation Research Institute, 5, 6; Confidential submission 52; University of Technology Sydney, submission to the IP, 98.
865 H Dobson, submission to the WP, 7.
866 Northraine, submission to the WP, 6.
867 Northraine, submission to the WP, 6.
868 Swinburne Social Innovation Research Institute, submission to the IP, 5, 6.
869 Swinburne Social Innovation Research Institute, submission to the IP, 5, 6.
870 See WP submissions: Australian Council of Trade Unions,4; J Evans 2; Swinburne Social Innovation Research Institute, 5,6; Portable, 1, 2.
871 Swinburne Social Innovation Research Institute, submission to the IP, 5, 6.
872 See, WP submissions: Australian Council of Trade Unions, 4; Australian Services Union, 3; Portable, 1, 2.
873 See WP submissions: Montreal AI Ethics Institute, 17; NTT Com Cloud Infrastructure Services, 8; Dobson, 7; G Greenleaf, R Clarke, D Lindsay, 6; University of Melbourne, 9.
874 See WP submissions: Confidential submission 52; Montreal AI Ethics Institute,17; Remi AI,3, 5; Dobson,6; and University of Technology Sydney, submission to the IP, 98.
875 Montreal AI Ethics Institute, submission to the WP, 17.
876 See WP submissions: Chartered Accountants Australia and NZ, 2; Centre for Policy Futures, University of Queensland, 6; H Dobson, 8; Remi AI, 6; Northraine, 6; Confidential submission 52; Australian Human Rights Institute, 9.
877 Chartered Accountants Australia and New Zealand, submission to the WP, 2.
878 H Dobson, submission to the WP, 8.
879 See WP submissions: Law Council of Australia, 6; Remi AI, 6; G Greenleaf, R Clarke, D Lindsay, 6.
880 See WP submissions: G Greenleaf, R Clarke, D Lindsay, 6; Swinburne Social Innovation Research Institute 6; Australian Human Rights Institute 8; Queensland Council of Civil Liberties 7; Northraine, 6; Remi AI, 6.
881 Swinburne Social Innovation Research Institute University, submission to the WP, 6.
882 For example, Remi AI submitted that the expert body could publish a live feed of its activities and findings online, curated by AI Remi AI, submission to the WP, 6.
883 See WP submissions: C Nichols, 7; NTT Communications Cloud Infrastructure Services, 9; Northraine, 6; Remi AI, 6; G Greenleaf, R Clarke, D Lindsay, 6; Queensland Council of Civil Liberties, 7.
884 See Chapter 3, Regulation.
885 See Chapter 6, Accountable AI informed decision making.
886 See Chapter 6, Accountable AI informed decision making for a discussion on the human rights engaged through the use of AI-informed decision making.
887 See Chapter 6: Accountable AI informed decision making.
888 *Enhancing Online Safety Act 2015* (Cth) Part 2.
889 For example, the following IP submissions refer broadly to these two types of access: Digital Gap Initiative, 8; Speech Pathology Australia, 4.
890 ICCPR, arts 2, 26; ICESCR art 2.
891 CRPD, art 1.
892 Excerpts have been adapted from Disability Action, *What is the UNCRPD?* (2018) <https://www.disabilityaction.org/what-is-the-uncrpd>.
893 CRPD, art 24.
894 CRPD, art 25.
895 CRPD, art 27.
896 CRPD, art 29.
897 CRPD art 5, 12.
898 CRPD, art 30.
899 CRPD, art 4.
900 CRPD, art 31.
901 CRPD, art 33.
902 CRPD, art 9(2)(a), 9(2)(h).
903 CRPD, art 3.
904 Committee on the Rights of Persons with Disabilities, *General comment No. 2 (2014) Article 9: Accessibility*, 11th sess, CRPD/C/GC/2 (11 April 2014) 2.
905 Committee on the Rights of Persons with Disabilities, *General comment No. 2 (2014) Article 9: Accessibility*, 11th sess, CRPD/C/GC/2 (11 April 2014) 2.
906 Committee on the Rights of Persons with Disabilities, *General comment No. 2 (2014) Article 9: Accessibility*, 11th sess, CRPD/C/GC/2 (11 April 2014), 2.
907 ICCPR, art 15(1)(b).
908 CRPD, art 9(h).
909 The key state and territory laws are: *Discrimination Act 1991* (ACT); *Anti-Discrimination Act 1977* (NSW); *Anti-Discrimination Act 1996* (NT); *Anti-Discrimination Act 1991* (Qld); *Equal Opportunity Act 1984* (SA); *Anti-Discrimination Act 1998* (Tas); *Equal Opportunity Act 2010* (Vic); *Equal Opportunity Act 1984* (WA).
910 *Disability Discrimination Act 1992* (Cth) s 5(1).
911 *Disability Discrimination Act 1992* (Cth) s 6(1)-(2).
912 *Disability Discrimination Act 1992* (Cth) s 5(2).
913 *Australian Human Rights Commission Act 1986* (Cth) s 46P.
914 *Australian Human Rights Commission Act 1986* (Cth) s 46PO.
915 *Disability Discrimination Act 1992* (Cth) s 31; Disability (Access to Premises – Buildings) Standards 2010; Disability Standards for Accessible Public Transport 2002; Disability Standards for Education 2005.
916 *Broadcasting Services Act 1992* (Cth), s 5(1).
917 *Telecommunications Act 1997* (Cth), s 376, ss 380-381.
918 Australian Communications and Media Authority, *Telecommunications Disability Standard* (Requirements for Customer Equipment for use with the Standard Telephone Service — Features for special needs of persons with disabilities, AS/ACIF S040) (17 February 2015) <https://www.legislation.gov.au/Details/F2015L00191>.
919 Standards Australia, *About us* (20 June 2018) <https://www.standards.org.au/>.
920 Standards Australia, *Standards Catalogue* <https://www.standards.org.au/standards-catalogue/sa-snz/communication/it-040/as--en--301-dot-549-colon-2016>.
921 Department of Finance, *Accessibility ICT Procurement Standard* (9 September 2016) <https://www.finance.gov.au/blog/2016/09/09/Accessibility-ICT-Procurement-Standard/>.

922 W3C Web Accessibility Initiative, *Web Content Accessibility Guidelines (WCAG) Overview* (15 June 2018) <https://www.w3.org/WAI/standards-guidelines/wcag/>.
923 W3C Web Accessibility Initiative, *Web Content Accessibility Guidelines (WCAG) Overview* (15 June 2018).
924 Australian Government, *Accessibility and Inclusivity* (10 May 2018) <https://guides.service.gov.au/content-guide/accessibility-inclusivity/>.
925 International Organization for Standardization, *Standards* (20 June 2018) <https://www.iso.org /standards.html>.
926 International Electrotechnical Commission, *International standards and conformity assessment for all electrical, electronic and related technologies* (10 June 2018) <http://www.iec.ch/index.htm>.
927 Commonwealth of Australia, National Disability Strategy 2010-2020 (2011).
928 *National Disability Insurance Scheme Act 2013* (Cth), s 4(2).
929 Digital Transformation Agency, *Digital Service Standard* (10 May 2017) <https://www.dta.gov.au/standard/>.
930 Digital Transformation Agency, *Accessibility and inclusivity* (10 May 2018) <https://guides.service.gov.au/content-guide/accessibility-inclusivity/>.
931 See, for example, IP submissions: Google, University of Technology Sydney; ANZ.
932 See, for example, IP: Digital Gap Initiative; Australian Communications Consumer Action Network; Public Interest Advocacy Centre; National Disability Services.
933 United Nations Population Fund, *Human Rights Principles* (2005) <https://www.unfpa.org/resources/human-rights-principles>.
934 See, for example, IP submissions: Australian Communications Consumer Action Network, 8, 9; Kingsford Legal Centre, 8, 9; National Association of Community Legal Centres, 10, 11.
935 UNESCO, *Global Education Monitoring Report Education for people and planet: Creating sustainable futures for all* (2016), 8.
936 Committee on the Rights of Persons with Disabilities, *General comment No. 2 (2014) Article 9: Accessibility*, 11th sess, CRPD/C/GC/2 (11 April 2014), 2.
937 Australian Communications Consumer Action Network, submission to the IP, 6.
938 See IP submissions: Speech Pathology Australia, 5; La Trobe LawTech, 9.
939 P Harpur, submission to the IP, 16–19.
940 LexisNexis, submission to the IP, 6.
941 See IP submissions: Digital Gap Initiative, 9, 10; WebKeyIT, 1.
942 See IP submissions: Intopia,10; Digital Gap Initiative, 9.
943 Intopia, submission to the IP, 3.
944 Intellectual Disability Rights Service, submission to the IP, 1.
945 Intellectual Disability Rights Service, submission to the IP, 1.
946 See IP submissions: Intellectual Disability Rights Service, 1; Intopia, 3; P Harpur, 12-15; Web Key IT, 2.
947 See IP submissions: Intopia, 4; Australian Rehabilitation and Assistive Technology Association, 4, 5.
948 Intopia, submission to the IP, 3.
949 See IP submissions: C Devenney and C Mills; M Browning, M Ellis, K Yeoh, with T Leiman.
950 See, for example IP submissions: Intopia,4; Castan Centre for Human Rights Law, 20; Australian Centre for Health Law Research, 7.
951 Australian Rehabilitation and Assistive Technology Association, submission to the IP, 9.
952 See IP submissions: Australian Rehabilitation and Assistive Technology Association, 3; Northraine, 13; Digital Gap Initiative, 4; Intopia, 10.
953 Ability Research Centre, submission to the IP, 2.
954 National Disability Services, submission to the IP, 1.
955 See IP submissions: Intopia, 10; University of Technology Sydney, 65.
956 Intellectual Disability Rights Service, submission to the IP, 1.
957 Carers Australia, submission to the IP, 1.
958 See, for example, IP submissions: Australian Rehabilitation and Assistive Technology Association; Northraine, 13.
959 National Disability Services, submission to the IP, 1.
960 Ability Research Centre, submission to the IP, 2.
961 Public Interest Advocacy Centre, submission to the IP, 16.
962 Australian Rehabilitation and Assistive Technology Association, submission to the IP, 11, 12.
963 See, for example, NDIS Evaluation by National Institute of Labour Studies at Flinders University, SA <https://www.dss.gov.au/disability-and-carers/programmes-services/for-people-with-disability/national-disability-insurance-scheme>; and Inquiries conducted by the Joint Standing Committee on the National Disability Insurance Scheme <https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/National_Disability_Insurance_Scheme>.
964 See, for example, Carers Australia; Consultation Roundtable 14.
965 Australian Human Rights Commission, Human Rights and Technology Conference, 23 July 2018.
966 National Disability Services, submission to the IP, 2.
967 See IP submissions: Australian Rehabilitation and Assistive Technology Association, 12; Australian Communications Consumer Action Network, 8; Confidential submission 2, 2; Confidential Individual Consultation with the Australian Human Rights Commission.
968 Consultation Roundtable 15.
969 National Disability Services, submission to the IP, 2; Confidential submission 2.
970 See IP submissions: R Bryer; S Murphy, 2; National Association of Community Legal Centres, 11.
971 S Murphy, submission to the IP, 6.
972 Australian Human Rights Commission, 'Human Rights and Technology IP' (July 2018), 39 <https://tech.humanrights.gov.au>.
973 S Murphy, submission to the IP, 6.
974 Confidential submission 38 to the IP, 2.
975 Consultation Roundtable 10.
976 ANZ, submission to the IP, 1.
977 ANZ, submission to the IP, 1.
978 Noted in IP submissions: S Murphy, 2, 6; Digital Gap Initiative 8, 9; WebKey IT, 1; Intopia, 4, 10; Australian Rehabilitation and Assistive Technology Association, 9; A Normand, 1; P Harpur, 24; Australian Communications Consumer Action Network, 7; Intellectual Disability Rights Service, 3, 4.
979 Australian Rehabilitation and Assistive Technology Association, submission to the IP, 9.
980 S Murphy, submission to the IP, 7.
981 Noted in IP submissions: Intopia, 10; Digital Gap Initiative, 9; S Murphy, 7.
982 See IP submissions: Public Interest Advocacy Centre, 20. Consultation roundtables 14, 15; Australian Communications Consumer Action Network, 8, 9.
983 Consultation roundtables 14 ,15; Australian Communications Consumer Action Network, submission to the IP, 9.
984 Consultation roundtable 14.
985 A Normand, submission to the IP, 1.
986 P Harpur, submission to the IP, 24.
987 Australian Communications Consumer Action Network, submission to the IP, 7.
988 Australian Communications Consumer Action Network, submission to the IP, 7.
989 See IP submissions: Public Interest Advocacy Centre, 20; Australian Communications Consumer Action Network, 9; Consultation roundtables 14, 15.
990 Australian Communications Consumer Action Network, submission to the IP, 8-9.
991 See IP submissions; Consultation roundtables 14,15.
992 See IP submissions: Confidential submission 2, 3; S Murphy, 9.
993 Portable, submission to the IP, 3.
994 ANZ, submission to the IP, 2.
995 Australian Rehabilitation and Assistive Technology Association, submission to the IP, 10.
996 Australian Communications Consumer Action Network, submission to the IP, 5.
997 Intopia, submission to the IP, 6.
998 See IP submissions: S Murphy, 10; University of Technology Sydney, 72; Digital Gap Initiative, 6.
999 See IP submissions: S Murphy, 6; University of Technology Sydney, 72.
1000 See IP submissions: S Murphy, 6; University of Technology Sydney, 72; Consultation Roundtable 14.

1001 Noted in IP submissions: National Association of Community Legal Centres, 4, 10; Intellectual Disability Rights Service, 3-4; Public Interest Advocacy Centre, 16; Speech Pathology Australia, 4.
1002 National Association of Community Legal Centres, submission to the IP, 10.
1003 Noted in IP submissions: Public Interest Advocacy Centre, 16; S Murphy, 7.
1004 S Murphy, submission to the IP, 7.
1005 Public Interest Advocacy Centre, submission to the IP, 16.
1006 Intellectual Disability Rights Service, submission to the IP, 3, 4.
1007 Speech Pathology Australia, submission to the IP, 4.
1008 See IP submissions: Digital Gap Initiative, 11; Intopia, 6; WebKey IT, 6.
1009 Digital Gap Initiative, submission to the IP,11.
1010 Speech Pathology Australia, submission to the IP, 7.
1011 Intellectual Disability Rights Service, submission to the IP, 3, 4.
1012 Noted in IP submissions: S Murphy, 8; National Association of Community Legal Centres, 11; Public Interest Advocacy Centre, 19, 20.
1013 Intopia, submission to the IP, 11.
1014 Senate Community Affairs References Committee, Commonwealth of Australia, *Delivery of outcomes under the National Disability Strategy 2010-2020 to build inclusive and accessible communities* (2017) <https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/ Community_Affairs/AccessibleCommunities/Report>.
1015 Australian Government, *Shut Out: The Experience of People with Disabilities and their Families in Australia* (2009) <https://www.dss.gov.au/our-responsibilities/disability-and-carers/publications-articles/policy-research/shut-out-the-experience-of-people-with-disabilities-and-their-families-in-australia>.
1016 Australian Bureau of Statistics, *4159.0 – General Social Survey: Summary Results, Australia* (2014).
1017 CRPD art 4.
1018 CRPD art 9.
1019 Committee on the Rights of Persons with Disabilities, *General comment No. 2 (2014) Article 9: Accessibility*, 11th sess, CRPD/C/GC/2 (11 April 2014) 9-10.
1020 Gerard Goggin, Scott Hollier and Wayne Hawkins, 'Internet accessibility and disability policy lessons for digital inclusion and equality from Australia' (2017) 6 (1) *Internet Policy Review* 9.
1021 Australian Government, Commonwealth Indigenous Procurement Policy (2015), 19.
1022 Business Council of Australia, "$3 billion boost for Indigenous businesses" (Media Release, 12 August 2019).
1023 Commonwealth of Australia, *National Disability Strategy 2010-2020* (2011), 35. Universal design is discussed in Chapter 10.
1024 Department of Finance, *Accessibility ICT Procurement Standard* (9 September 2016) <https://www.finance.gov.au/blog/2016/09/09/Accessibility-ICT-Procurement-Standard/>.
1025 Department of Finance and Deregulation, *Web accessibility National Transition Strategy* (2010) <https://www.finance.gov.au/publications/wcag-2-implementation/docs/wcag-transition-strategy.pdf>.
1026 Australian Government Department of Finance, *Web Accessibility National Transition Strategy 2012 Progress Report Data* (2014) <https://data.gov.au/dataset/ds-dga-e59c06a1-86e6-4ac7-8c34-d52f4d2406d8/details>.
1027 Vivienne Conway, 'Website accessibility in Australia and the national transition strategy: Outcomes and findings' (Doctor of Philosophy, Edith Cowan University, 2014).
1028 CRPD, arts 9 and 30(1)(b).
1029 See, for example, *Audio Description in Australia* (25 November 2019) <http://audiodescriptionau.com.au/>; Blind Citizens Australia, *Tv4all* (25 November 2019) <http://www.tv4all.com.au/>.
1030 Katie Ellis, Gwyneth Peaty, Leanne McRae, Mike Kent and Kathryn Locke, *Audio Description in Australia* (Australian Communications Consumer Action Network, 2019), 21.
1031 Katie Ellis, Gwyneth Peaty, Leanne McRae, Mike Kent and Kathryn Locke, *Audio Description in Australia* (Australian Communications Consumer Action Network, 2019), 42.
1032 ACMA, *Captioning obligations: Free-to-air* (17 September 2018) <https://www.acma.gov.au/Citizen/TV-Radio/Television/Captioning/captioning-obligations-on-freetoair-subscription-television-i-acma>.
1033 See, for example, Australian Communications Consumer Action Network, submission to the IP, 9.
1034 See, for example, Australian Communications and Media Authority, 'Report on the review of the captioning obligations in the Broadcasting Services Act 1992', 7.
1035 Broadcasting Services Amendment (Audio Description) Bill 2019 (Cth) Item 14.
1036 Roy Morgan, *Measuring Australia's Digital Divide: Australian Digital Inclusion Index* (2018), 7, 8.
1037 'Digital divide' may be described as the gap between those who are digitally included and those who are not. Digital inclusion is based on the premise that everyone should be able to make full use of digital technologies to manage their health and wellbeing, access education and services, organise their finances, and connect with friends, family and the world beyond. Roy Morgan, *Measuring Australia's Digital Divide: Australian Digital Inclusion Index* (2018).
1038 See IP submissions: La Trobe LawTech, 9; National Association of Community Legal Centres, 4.
1039 Roy Morgan, *Measuring Australia's Digital Divide: Australian Digital Inclusion Index* (2018), 10.
1040 Tony Vinson and Margot Rawsthorne, *Dropping off the Edge Report* (2015), 48.
1041 ACOSS, *Digital Divide Policy Snapshot* (January 2016) 2.
1042 Australian Red Cross, submission to the IP, 8.
1043 National Association of Community Legal Centres, submission to the IP, 8.
1044 See, for example IP submissions: Red Cross, 8; Australian Rehabilitation and Assistive Technology Association, 8-9; CHOICE, 5. See also Roy Morgan, *Measuring Australia's Digital Divide: Australian Digital Inclusion Index* (2018) 17.
1045 See, for example, IP submissions: Law Council of Australia; University of Technology Sydney; University of Melbourne.
1046 See IP submissions: The Allens Hub for Technology, Law and Innovation, 2; Law Council of Australia, 7; Consultation Roundtable 13.
1047 For example, University of Melbourne, submission to the IP, 5.
1048 See IP submissions: Intellectual Disability Rights Service, 2, 5; Northraine, 13; National Association of Community Legal Centres, 11, 12; Australian Communications Consumer Action Network, 7, 8; Confidential submission 102, 2.
1049 National Disability Services, submission to the IP, 2.
1050 Intopia, submission to the IP, 4, 10.
1051 Consultation roundtable 15.
1052 WebKeyIT, submission to the IP, 1.
1053 See IP submissions: Dietitians Association of Australia, 2-4; Speech Pathology Australia, 4.
1054 Intellectual Disability Rights Service, submission to the IP, 1.
1055 Intopia, submission to the IP, 4.
1056 Noted in IP submissions: Northraine, 13; National Association of Community Legal Centres, 11; Intellectual Disability Rights Service, 4; University of Technology Sydney, 67; Consultation Roundtable 15.
1057 Confidential submission 102 to the IP, 2.
1058 Australian Communications Consumer Action Network, submission to the IP, 15.
1059 See IP submissions: Intellectual Disability Rights Service, 1; Australian Rehabilitation and Assistive Technology Association, 12. Consultation Roundtable 18.
1060 Confidential submission 102 to the IP, 1.
1061 Intopia, submission to the IP, 10.
1062 Ability Research Centre, submission to the IP, 2.
1063 University of Melbourne, submission to the IP, 5.
1064 University of Technology Sydney, submission to the IP, 66.
1065 Carers Australia, submission to the IP, 2.
1066 Carers Australia, submission to the IP, 2.
1067 Law Council of Australia, submission to the IP, 6.
1068 WebKeyIT, submission to the IP, 2.
1069 Confidential submission 97 to the IP, 21.
1070 WebKeyIT, submission to the IP, 2.
1071 Speech Pathology Australia, submission to the IP, 4.
1072 University of Melbourne, submission to the IP, 5.
1073 Dietitians Association of Australia, submission to the IP, 3.
1074 Centre for Inclusive Design, PwC Australia, *The Benefit of Designing for Everyone* (May 2019) 1.

1075 Standards Australia, *What is a Standard?* <https://www.standards.org.au/standards-development/what-is-standard>.
1076 See also Australian Communications Consumer Action Network, *Affordability Map* (June 2015) 8 <https://accan.org.au/Affordability%20Map%20June%202016.pdf>.
1077 Australian Communications Consumer Action Network, *Media Release No Australian Left Offline: ACCAN pushes for affordable broadband* (2019) <https://accan.org.au>.
1078 Australian Communications Consumer Action Network, 'No Australian Left Offline: ACCAN pushes for affordable broadband' (Media Release, 2019), 1.
1079 Australian Communications Consumer Action Network, 'No Australian Left Offline: ACCAN pushes for affordable broadband' (Media Release, 2019), 4.
1080 CRPD art 21.
1081 See, for example, IP submissions: S Murphy; Australian Communications Consumer Action Network, 9-10; Digital Gap Initiative, 18, 19; Centre for Inclusive Design, 1, 2; Adobe, 3; Google, 6-8.
1082 Google, submission to the IP, 6.
1083 See, for example, IP submissions: Google; Adobe; Digital Gap Initiative, Consultation Roundtables 10, 19.
1084 Digital Gap Initiative, submission to the IP, 17.
1085 'Human rights by design' is also discussed in Chapter 7 in relation to AI.
1086 BSR, *Artificial Intelligence: A Rights-Based Blueprint for Business. Paper 3: Implementing Human Rights Due Diligence* (Working Paper, August 2018), 13.
1087 CPRD, art 2.
1088 North Carolina State University, The Center for Universal Design, *The Principles of Universal Design* <https://projects.ncsu.edu/design/cud/about_ud/udprinciplestext.htm>.
1089 P Harpur, submission to the IP, 11.
1090 See IP submissions: Australian Rehabilitation and Assistive Technology Association, 10; Northraine, 13; National Association of Community Legal Centres, 11; Intopia, 12; Consultation Roundtable 14.
1091 P Harpur, submission to the IP, 11.
1092 University of Washington, 'What is the difference between accessible, usable, and universal design?' *Disabilities, Opportunities, Internetworking and Technology* (30 April 2019); Matt May, *The same, but different: breaking down accessibility, universality, and inclusion in design* (4 February 2018).
1093 W3C, *Web Content Accessibility Guidelines (WCAG) 2.1* (5 June 2018) <https://www.w3.org/TR/WCAG21/>.
1094 See, for example, S Murphy, submission to the IP, 6; Matt May, *The same, but different: breaking down accessibility, universality, and inclusion in design* (4 February 2018) <https://theblog.adobe.com/different-breaking-accessibility-universality-inclusion-design/>; W3C, *Web Content Accessibility Guidelines (WCAG) 2.1* (5 June 2018)<https://www.w3.org/TR/WCAG21/>.
1095 Adobe, submission to the IP, 2, 3.
1096 S Murphy, submission to the IP, 6.
1097 WebKeyIT, submission to the IP, 6.
1098 OCAD University, Inclusive Design Research Centre, *What is Inclusive Design* <https://idrc.ocadu.ca/about-the-idrc/49-resources/online-resources/articles-and-papers/443-whatisinclusivedesign>.
1099 OCAD University, Inclusive Design Research Centre, *What is Inclusive Design* <https://idrc.ocadu.ca/about-the-idrc/49-resources/online-resources/articles-and-papers/443-whatisinclusivedesign>.
1100 Centre for Inclusive Design, submission to the IP, 5.
1101 Adobe, submission to the IP, 3.
1102 Australian Banking Association, *Every customer counts* <https://www.ausbanking.org.au/images/uploads/Accessibility_Principles_for_Banking_web.pdf>.
1103 Australian Banking Association, *Every customer counts*, 5.
1104 W3C, *Accessibility, Usability, and Inclusion* (6 May 2016) <http://www.w3.org/WAI/intro/usable>.
1105 Jutta Treviranus, 'The Three Dimensions of Inclusive Design, Part Two' on *Medium* (30 March 2018).
1106 Jutta Treviranus, 'The Three Dimensions of Inclusive Design, Part Two' on *Medium* (30 March 2018).
1107 National Disability Services, submission to the IP, 4.
1108 University of Melbourne, submission to the IP, 11; M Johnson, submission to the WP.
1109 Centre for Inclusive Design, submission to the IP, 11.
1110 University of Melbourne, submission to the IP, 12; M Johnson, submission to the WP.
1111 See IP submissions: LexisNexis, 24; University of Technology Sydney, 72.
1112 Consultation Roundtable 15.
1113 University of Technology Sydney, submission to the IP, 72; Consultation Roundtable 15.
1114 Consultation Roundtable 15.
1115 Australian Centre for Health Law Research, submission to the IP, 1.
1116 Noted in IP submissions: WebKeyIT,6; University of Technology Sydney, 70; Australian Communications Consumer Action Network, 11; Digital Gap Initiative, 18; Public Interest Advocacy Centre, 21; Intellectual Disability Rights Service, 6; ANZ.
1117 Noted in IP submissions: University of Technology Sydney, 70; Digital Gap Initiative, 18, 19.
1118 S Murphy, submission to the IP, 9.
1119 ANZ, submission to the IP, 2.
1120 Intopia, submission to the IP, 12.
1121 Australian Rehabilitation and Assistive Technology Association, submission to the IP, 13.
1122 See IP submissions: LexisNexis, 24; University of Technology Sydney, 72.
1123 Intopia, submission to the IP, 12.
1124 Noted in IP submissions: WebKeyIT, 3; Digital Gap Initiative, 19; University of Technology Sydney, 72.
1125 S Murphy, submission to the IP, 10; Confidential Individual Consultation.
1126 University of Technology Sydney, submission to the IP, 72.
1127 Intopia, submission to the IP, 11.
1128 See IP submissions: ANZ, 2; Intopia, 6; Digital Gap Initiative, 18.
1129 Speech Pathology Australia, submission to the IP, 8.
1130 LexisNexis, submission to the IP, 24.
1131 See IP submissions: LexisNexis, 24; University of Melbourne, 12; Intopia, 12; Confidential submission 102; Consultation Roundtable 18.
1132 University of Technology Sydney, submission to the IP, 71.
1133 Centre for Inclusive Design, PwC Australia, *The Benefit of Designing for Everyone* (May 2019), 1.
1134 Forbes, *How to harness the power of network effects* 2 January 2018 <https://www.forbes.com/sites/forbescoachescouncil/2018/01/02/how-to-harness-the-power-of-network-effects/#5d28163962e8>.
1135 University of Melbourne, submission to the IP, 12.
1136 Sound Advice, *Deafness Led To The Phone, Internet & SMS Texts* (26 November 2019) <https://sound-advice.ie/deaf-phone-internet-sms-text/>.
1137 LexisNexis, submission to the IP, 24.
1138 Accenture, *Getting to equal: the disability inclusion advantage* (2018), 7; Intopia, submission to the IP, 12.
1139 See Digital Gap Initiative, submission to the IP, 17; Consultation Roundtables 14, 18.
1140 Noted in IP submissions: Public Interest Advocacy Centre; LexisNexis, 6; Consultation Roundtables 10, 19.
1141 Consultation Roundtable 10.
1142 WebKeyIT, submission to the IP, 1.
1143 Public Interest Advocacy Centre, submission to the IP, 15.
1144 Consultation Roundtable 14.
1145 Centre for Inclusive Design, PwC Australia, *The Benefit of Designing for Everyone* (May 2019), 20; Consultation roundtables 14 and 18.
1146 Digital Gap Initiative, submission to the IP, 17.
1147 P Harpur, submission to the IP, 11.
1148 CRPD, arts 9(2)(h) and 4(f).
1149 Commonwealth of Australia, *National Disability Strategy 2010-2020* (2011), 35.
1150 For example, W3C, *The Business Case for Digital Accessibility* <http://www.w3.org/WAI/business-case/>; Centre for Inclusive Design, PwC Australia, *The Benefit of Designing for Everyone* (May 2019).
1151 See, for example, Matt May, *The same, but different: breaking down accessibility, universality, and inclusion in design* (4 February 2018) <https://theblog.adobe.com/different-breaking-accessibility-universality-inclusion-design/>.

1152 BSR, *Artificial Intelligence: A Rights-Based Blueprint for Business. Paper 3: Implementing Human Rights Due Diligence* (Working Paper, August 2018), 13.
1153 Office of Information Commissioner, *Privacy management framework: enabling compliance and encouraging good practice* (4 May 2015).
1154 Filippo Raso, Hannah Hilligoss, Vivek Krishnamurthy, Christopher Bavitz and Levin K, 'Artificial Intelligence & Human Rights: Opportunities & Risks' *Berkman Klein Center Research Publication No 2018-6* (25 October 2018), 15.
1155 CRPD art 2.
1156 CRPD art 3.
1157 Australian Government Department of Social Services (2019) Council of Australian Governments Disability Reform Council <https://www.dss.gov.au/our-responsibilities/disability-and-carers/programmes-services/government-international/disability-reform-council>.
1158 Department of Social Services 'Shaping the future of disability policy for 2020 and beyond' (2019) <https://www.dss.gov.au/our-responsibilities/disability-and-carers/shaping-the-future-of-disability-policy-for-2020-and-beyond>.
1159 Australian Council of Learned Academies, *About us*, <https://acola.org/about-us/>.
1160 Engineers Australia, 'Continuing Professional Development,' *Training and Development* <https://www.engineersaustralia.org.au/Training-And-Development/Continuing-Professional-Development>.
1161 Engineers Australia, *Continuing Professional Development – Types and conditions* (2016) <https://www.engineersaustralia.org.au/sites/default/files/content-files/2016-12/cpd_types_and_conditions_march_2014.pdf>.
1162 See, for example, IP submissions: Digital Gap Initiative; Australian Communications Consumer Action Network.
1163 See, for example, IP submissions: Public Interest Advocacy Centre; Intopia; Digital Gap Initiative; Australian Communications Consumer Action Network.
1164 See, for example, IP submissions: Digital Gap Initiative; Australian Communications Consumer Action Network.
1165 Digital Gap Initiative, submission to the IP, 5.
1166 Consultation Roundtable 14, 15.
1167 See, for example IP submissions: University of Melbourne, Digital Gap Initiative.
1168 *Disability Discrimination Act 1992* (Cth) s 24.
1169 Australian Human Rights Commission, *Complaints under the Disability Discrimination Act* (14 December 2012) <https://www.humanrights.gov.au/complaints/complaint-guides/complaints-under-disability-discrimination-act>.
1170 For example, *Anti-Discrimination At 1977* (NSW), *Human Rights Act 2004* (ACT).
1171 As noted in the following IP submissions: Digital Gap Initiative, 5; Australian Communications Consumer Action Network, 11; Public Interest Advocacy Centre, 16; Consultation Roundtable 15.
1172 Australian Communications Consumer Action Network, submission to the IP, 11.
1173 P Harpur, submission to the IP, 41.
1174 University of Technology Sydney, submission to the IP, 69.
1175 As noted in the following IP submissions: Kingsford Legal Centre,15; Intopia, 5; Public Interest Advocacy Centre, 19; WebKeyIT, 2; Digital Gap Initiative, 16; Consultation roundtables 13, 14, 15.
1176 Standards Australia, *About us* (20 June 2018) <https://www.standards.org.au/>.
1177 Digital Gap Initiative, submission to the IP, 13.
1178 See IP submissions: Public Interest Advocacy Centre, 19; Digital Gap Initiative, 16; A Normand, 2; S Murphy, 8, 9; Kingsford Legal Centre, 15; Consultation Roundtable 14.
1179 Public Interest Advocacy Centre, submission to the IP, 19.
1180 Kingsford Legal Centre, submission to the IP, 15.
1181 A Normand, submission to the IP, 3.
1182 Public Interest Advocacy Centre, submission to the IP, 19.
1183 Digital Gap Initiative, submission to the IP, 16; Consultation roundtable 13.
1184 P Harpur, submission to the IP, 5.
1185 Consultation roundtables 10, 13, 14.
1186 See, IP submissions: Digital Gap Initiative; S Murphy; Consultation roundtable 15.
1187 See IP submissions: A Normand, 2; Digital Gap Initiative, 81; S Murphy; Consultation roundtable 15.
1188 See, for example, IP submissions: Digital Gap Initiative, 5, 16; Australian Communications Consumer Action Network, 5.
1189 See IP submissions: Intopia,11; Digital Gap Initiative, 14.
1190 Kingsford Legal Centre, submission to the IP, 15.
1191 Digital Gap Initiative, submission to the IP, 17.
1192 LexisNexis, submission to the IP, 23, 24.
1193 LexisNexis, submission to the IP, 23, 24.
1194 CRPD art 4(a).
1195 CRPD art 33.
1196 Committee on the Rights of Persons with Disabilities, *General comment No. 2 (2014) Article 9: Accessibility*, 11th sess, CRPD/C/GC/2 (11 April 2014), 2.
1197 Australian Human Rights Commission, *Discussion paper: Priorities for federal discrimination law reform* (Discussion Paper, 2019) <https://www.humanrights.gov.au/our-work/rights-and-freedoms/publications/discussion-paper-priorities-federal-discrimination-law>.
1198 For example: Senate Community Affairs References Committee, Commonwealth of Australia, *Delivery of outcomes under the National Disability Strategy 2010-2020 to build inclusive and accessible communities* (2017); Australian Government, *Shut Out: The Experience of People with Disabilities and their Families in Australia* (2009) <https://www.dss.gov.au/sites/default/files/documents/05_2012/nds_report.pdf>.
1199 Australian Human Rights Commission, submission No 30 to the Department of Infrastructure, Regional Development and Cities' Review of the Disability Standards for Accessible Public Transport 2002 (December 2018), 15.
1200 Australian Human Rights Commission, submission No 30 to the Department of Infrastructure, Regional Development and Cities' Review of the Disability Standards for Accessible Public Transport 2002 (December 2018), 15.
1201 *Disability Discrimination Act 1992* (Cth) s 31.
1202 *Disability Discrimination Act 1992* (Cth) s 31.
1203 Australian Human Rights Commission, *Disability Standards* <https://www.humanrights.gov.au/our-work/disability-rights/disability-standards>.
1204 *Disability Discrimination Act 1992* (Cth) s 32.
1205 *Disability Discrimination Act 1992* (Cth) s 34.
1206 For example, Disability Standards for Accessible Public Transport 2002, Part 34 Review, Schedule 1 Target dates for compliance.
1207 Australian Human Rights Commission, submission No 30 to the Department of Infrastructure, Regional Development and Cities' Review of the Disability Standards for Accessible Public Transport 2002 (December 2018), 15.
1208 Australian Human Rights Commission. 'Information concerning Australia's compliance with the Convention on the Rights of Persons with Disabilities', submission to the UN Committee on the Rights of Persons with Disabilities (2019), 15, 16.
1209 Australian Human Rights Commission, 'Information concerning Australia's compliance with the Convention on the Rights of Persons with Disabilities', submission to the UN Committee on the Rights of Persons with Disabilities (2019).
1210 Australian Human Rights Commission, *Free and Equal: An Australian conversation on human rights* (31 May 2019) <https://www.humanrights.gov.au/our-work/rights-and-freedoms/projects/free-and-equal-australian-conversation-human-rights>.
1211 Australian Human Rights Commission, *Discussion paper: Priorities for federal discrimination law reform* (Discussion Paper, 2019).
1212 Australian Human Rights Commission, *Discussion paper: Priorities for federal discrimination law reform* (Discussion Paper, 2019), 13-17.
1213 Australian Human Rights Commission, *Discussion paper: Priorities for federal discrimination law reform* (Discussion Paper, 2019), 14.
1214 Australian Human Rights Commission, Discussion paper: Priorities for federal discrimination law reform (Discussion Paper, 2019), 16.
1215 Australian Human Rights Commission (2019) *Discussion paper: Priorities for federal discrimination law reform* 13.
1216 Australian Human Rights Commission (2019) *Discussion paper: Priorities for federal discrimination law reform* 13.

Australian Human Rights Commission
www.humanrights.gov.au