James Martin Institute for Public Policy

Collaborative Project Report

Leadership for Responsible AI

A Constructive Agenda for NSW

DECEMBER 2023

About JMI

A formal partner of the NSW Government, JMI is a unique joint venture between government and leading Australian universities. Launched in 2021, we work closely with government ministers, departments, and other decision-makers to help address their most pressing policy priorities, enabling them to harness a wide range of expert advice. JMI is an independent, non-partisan policy institute with charitable status.

Acknowledgement of Country

The James Martin Institute for Public Policy acknowledges the Gadigal people of the Eora Nation upon whose ancestral lands our Institute stands. We pay respect to Elders both past and present, acknowledging them as the traditional custodians of knowledge for these lands.
We celebrate the diversity of Aboriginal peoples and their ongoing cultures and connections to the lands and waters of NSW.

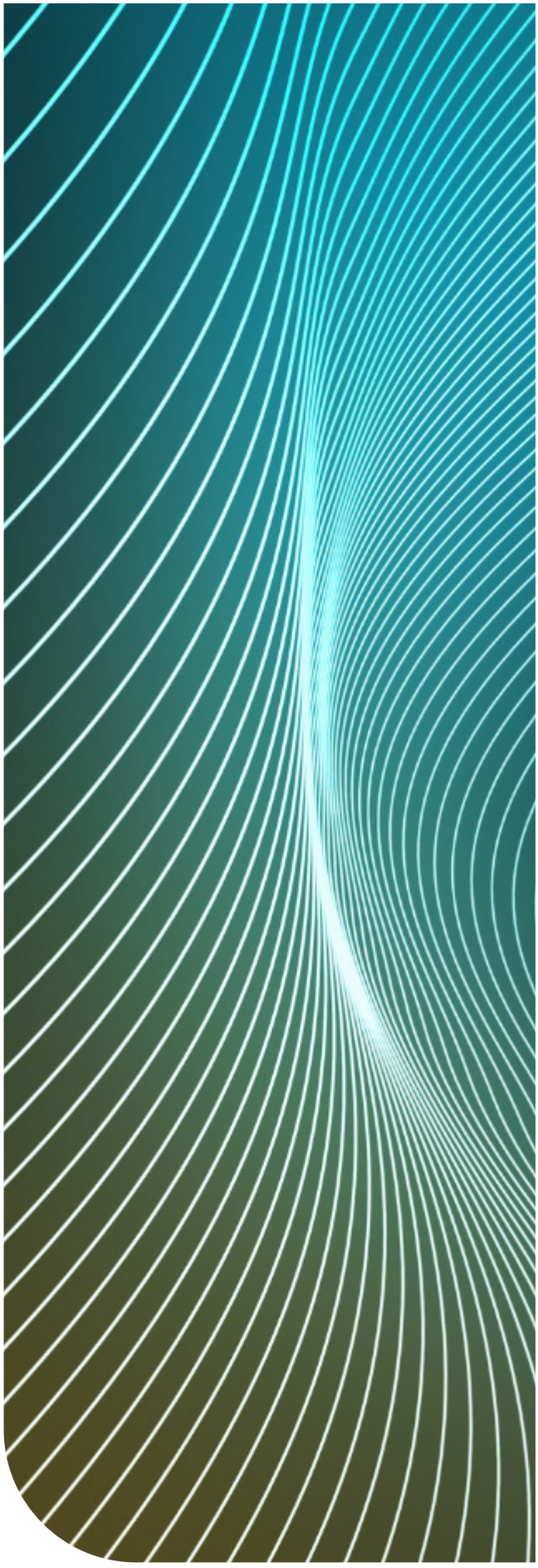
JMI collaborative projects

JMI collaborative projects are unique. They are delivered by joint project teams, comprising of Institute staff, policymakers, and academic experts.

An Expert Advisory Group (EAG) provides advice and direction to the blended project team.

JMI collaborative projects:

- tackle the big, multi-dimensional policy challenges facing governments;
- leverage expertise from academics, policymakers, practitioners and the wider community to deliver innovative and pragmatic solutions;
- forge dynamic collaborations between academic experts and government policymakers that enable ongoing relationships.



Contents

Cover: Getty Images

Executive summary

Key points

- The NSW Government is well placed to take a leading role in addressing the impact of AI activity, for the benefit of our economy and local communities.
- The increasing development and use of artificial intelligence (AI) technologies by private, non-government actors creates significant opportunities for NSW, but also has the potential to significantly affect people's rights and interests.
- The NSW Government is widely recognised as a global digital leader. This provides a strong foundation for the public sector to confidently pursue a constructive agenda that supports AI technologies in the market to have wide social and economic benefit for people in the state.
- This agenda builds on clear strengths in the private sector. With established clusters of advanced computing businesses, the state benefits from 41 per cent of all Australian software and application programmers and 45 per cent of all Australian AI businesses.
- Positioning NSW as home for effectively regulated, trustworthy AI will support the state to be a more attractive site for AI investment and innovation, while ensuring people and communities benefit fairly from emerging technologies.
- Addressing this challenge will involve leveraging and improving existing laws and policies, and considering new regulations to specifically address AI.
- This report outlines 15 policy opportunities where NSW policymakers could make a step-change on AI, including the establishment of an Emerging Technology Commissioner.

This report, the result of a collaborative project between the James Martin Institute for Public Policy and the NSW Government, outlines the forward-looking role that NSW policymakers could take to respond to the unprecedented development and use of AI.

The NSW Government has already made substantial progress in assuring the government's own use of AI by implementing the AI Assurance Framework for NSW Government projects. Now is a critical moment for the NSW Government to go further and enact a coordinated, considered response to the regulation, market shaping, and public sector capability uplift needed to manage this technology into the future.

AI impacts many policy areas where responsibility is distributed to state governments in Australia.
In addition to **regulation**, this report explores the ways in which the NSW Government might constructively **shape** the future direction of AI technologies. Progress in both aspects – "regulating" and "shaping" – depends on further building and maintaining effective public sector capability in AI, including technical, strategic, and sociotechnical dimensions.

The report suggests opportunities for the NSW Government to address these challenges in two parts:

1. **Regulation and shaping:** The report explores how NSW could effectively adapt and use existing laws and regulation, as well as exploring the need for new measures, to more fully account for the impacts of AI technologies. It also explores how the NSW Government could more actively work to shape AI technologies in the market for the benefit of people and communities.
2. **Enabling the public sector to respond effectively**: The report outlines cost-effective opportunities to support the public sector to respond effectively and meet public needs in respect of this urgent challenge.

The report suggests a range of explicit policy opportunities for the NSW Government to consider. Fifteen detailed policy opportunities are outlined in the report. These include:

- Creating an Emerging Technology Commissioner to facilitate responsible development and use of AI across NSW, and support the development of government and regulator capability in an efficient and cost-effective way;
- Embedding the intellectual leadership of Aboriginal and Torres Strait Islander people in the ongoing response to AI;
- Selectively using new regulatory tools for the private sector, including auditing and reporting requirements, and developing the NSW AI assurance industry to support this effort;
- Utilising government's existing levers, such as procurement, more effectively to drive responsible technology in the private sector;
- Resourcing and explicitly encouraging the NSW AI Review Committee to publish educative case studies and assurance reviews communicating lessons from its work;
- Establishing a regulatory strategy, including reviewing gaps in existing legislation, and supporting regulators and government agencies to respond effectively to new AI-related challenges in their area of focus; and
- Developing lived experience advisory mechanisms for groups at risk of harm.

An effective, strategic regulatory and shaping agenda on AI is a necessary element of ensuring NSW is a digital leader – capable of attracting talent and investment in support of inclusive economic growth and ensuring technological advances deliver better outcomes for diverse communities.

Opportunities for the NSW Government

The NSW Government can consider a range of policy opportunities to address challenges related to non-government development and use of AI. These include the following potential actions:

Law, guidance and oversight to manage risks and reduce harm

1. Establish a regulatory strategy for AI development and use in the market, reflecting the AI regulation principles of notification, explainability, disputes and appeals, and liability.
2. Undertake a review of existing legislation to identify points requiring clarification in the form of new legislative guidance, or legislative amendment, to reflect the AI regulation principles of notification, explainability, disputes and appeals, and liability.
3. Resource regulators appropriately to develop the capability and capacity to respond to emerging risks presented by AI.
4. Implement auditing and public reporting requirements with respect to AI systems in response to the greatest risks of harm to the community arising from AI.
5. Stimulate and support the development of an effective local AI assurance industry to enable its assurance activities across firms in the NSW economy, including through stewardship of assurance standards.

Shaping the development and use of AI to increase social benefit

1. Explicitly refer to or integrate the AI Assurance Framework – and the obligations in confers in respect to procurement - into the NSW Government Procurement Policy Framework.
2. Amend the NSW Supplier Code of Conduct to include elements of the AI Ethics Policy to clearly outline suppliers' responsibilities in respect of responsible AI.
3. Resource and explicitly encourage the NSW AI Review Committee to publish educative case studies and assurance reviews communicating lessons from its work.
4. Adopt and/or endorse codes of practice, to guide compliance with existing legal obligations, to integrate into procurement requirements, and for voluntary adoption more broadly across the private sector.
5. Facilitate peer learning and knowledge-transfer among firms to grow the private sector's understanding of ethical and responsible AI practices, and provide opportunities for firms to showcase their responsible practices.

Supporting the public sector to respond effectively

1. Create an independent Emerging Technology Commissioner to support and facilitate responsible AI design, development, and deployment across the market. The Commissioner's functions could include:
    - Engaging both internally, to build NSW Government's capacity, and externally, with the public and to provide general advice and information across the private sector;
    - Issuing and endorsing codes of practice;
    - Representing NSW in interjurisdictional forums.
2. Embed the leadership of Aboriginal and Torres Strait Islander experts in NSW Government's response to AI on an ongoing basis, in staffing and advisory functions.
3. Establish ongoing advisory mechanisms to engage people with diverse lived experience to advise government on how to best protect the rights and interests of people in NSW in respect of AI and emerging technologies. This could include:
    - Embedding lived experience in relevant government organisations, including the proposed Office of the Emerging Technology Commissioner;
    - Establishing a civil society advisory committee; and
    - Convening a citizens' assembly on AI.
4. Develop a map of AI regulatory capabilities required within the public sector.
5. Develop and implement a strategic plan for the acquisition and development of the AI regulatory capabilities identified through mapping.

Introduction

AI presents transformative opportunities, and unprecedented challenges for NSW

This report proposes policy opportunities for the NSW Government as it continues the important task of responding to the development and use of AI in NSW. It focuses on the role that the NSW Government could play in regulating and shaping the development and use of AI to serve human and community interests, and how the NSW Government might coordinate its response and leadership on this critical challenge.

The emergence of AI, as the Organisation for Economic Cooperation and Development (OECD) has emphasised, "has pervasive, far-reaching and global implications that are transforming societies, economic sectors and the world of work, and are likely to increasingly do so in the future."[1] NSW is a global digital leader, and benefits from a thriving and competitive advanced computing and software industry. The responsible development and use of AI systems can generate significant economic and social opportunities that have the potential to positively transform many aspects of public and private life. However, the increasing ubiquity of AI across society also poses risks to fundamental human rights, democratic values, and the welfare and wellbeing of people in NSW.[2] Therefore, in order to create the conditions for AI technologies and industries to thrive and be of greatest benefit in NSW, government should craft a strategic response.

The current pace of developments in AI is unprecedented, including because of the advancement of generative AI. The risks arising from the development and use of AI technologies are not merely hypothetical but have been realised in instances of actual harm to citizens and communities around the world.

Significant energy is being invested in this challenge nationally and internationally. Within Australia, the Commonwealth Government is progressing its approach, including through the recent Department of Industry, Science and Resources consultation on supporting responsible AI, the AI Ethics Framework, and membership of the Global Partnership on Artificial Intelligence. Globally, Australia has recently joined the European Union (EU) and 27 countries, including the United States (US), the United Kingdom (UK) and China, in signing the Bletchley Declaration, affirming that AI should be developed and used in a manner that is safe, human-centric, trustworthy, and responsible. However, the role of subnational governments in shaping and regulating the development and use of AI has traditionally received less attention and analysis from governments and researchers.[3]

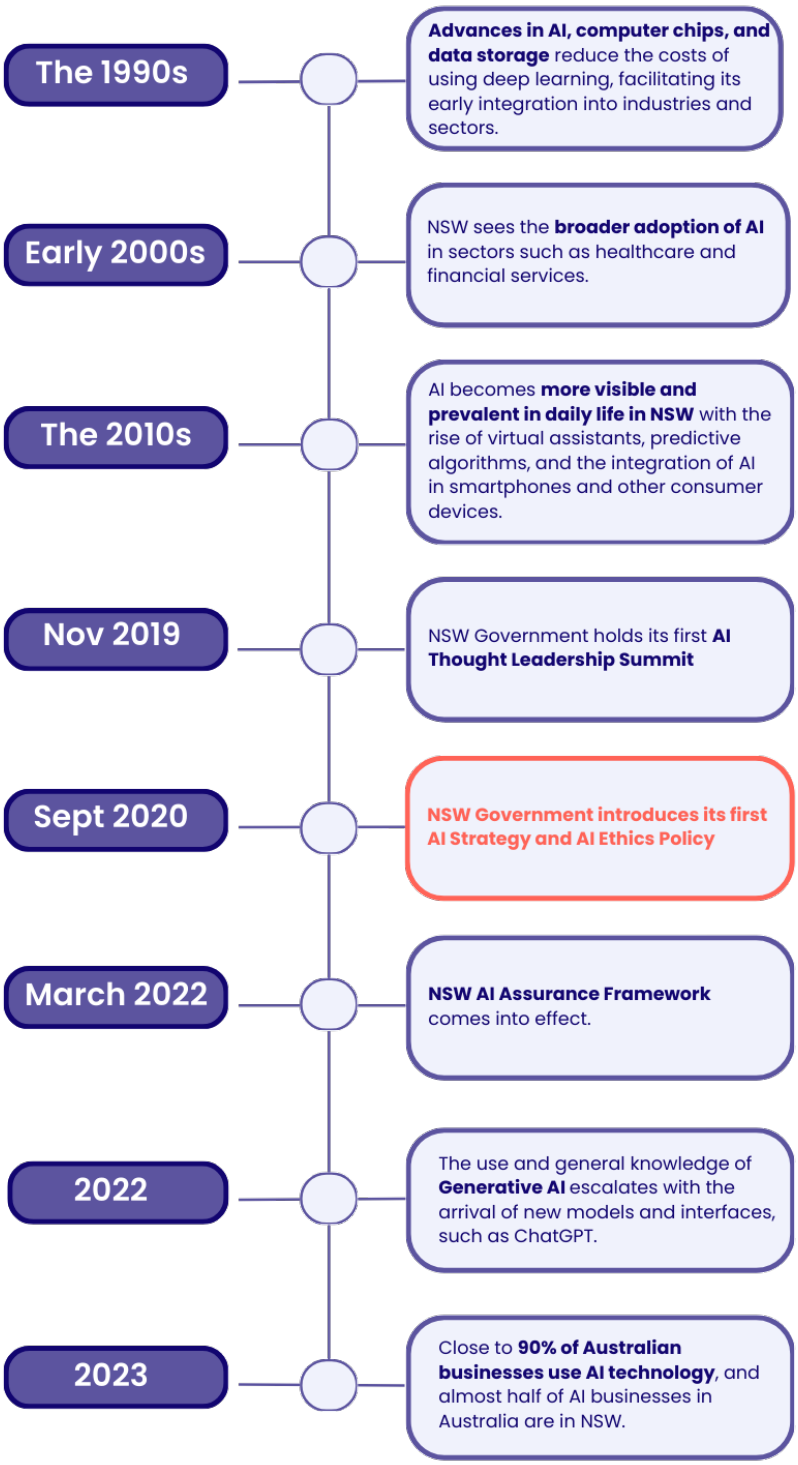
NSW Government can act now to protect the rights and interests of its people and communities

The NSW Government has demonstrated considerable leadership in enabling responsible use of AI across the public sector. To ensure that its own use of AI has been ethical and responsible, NSW Government has implemented an AI Strategy, AI Ethics Policy, and the nationally recognised NSW AI Assurance Framework. This investment forms the foundation of a policy response to protect and benefit people in NSW. The NSW Government is now well-positioned to extend this work and consider the role it should play with respect to all development and use of AI technologies and the impact of AI activity on people and communities. The development and use of AI by private, non-government actors impacts widely on private and public life and has the potential to significantly affect the rights and interests of people in NSW.

AI is connected to many policy areas where responsibility is distributed to state governments in Australia. Therefore, this report focuses on responses that are uniquely, or most meaningfully, instituted, facilitated, or catalysed by the NSW Government, rather than the Commonwealth Government or global governance bodies. The report suggests pathways that the NSW Government might follow to progress its response to promoting responsible AI.

A timely response from the NSW Government is important to ensure that AI technologies serve the wellbeing of people who live in NSW. Positioning NSW as a digital leader and home for effectively regulated, trustworthy AI will ultimately help make the state a more attractive site for AI investment and innovation. Moreover, as governments explore new ways to shape – and not merely remediate – markets, including through industry policy and other levers, AI innovation stands out as an area where government can play a constructive role in setting a direction for markets so that they better serve people and communities.

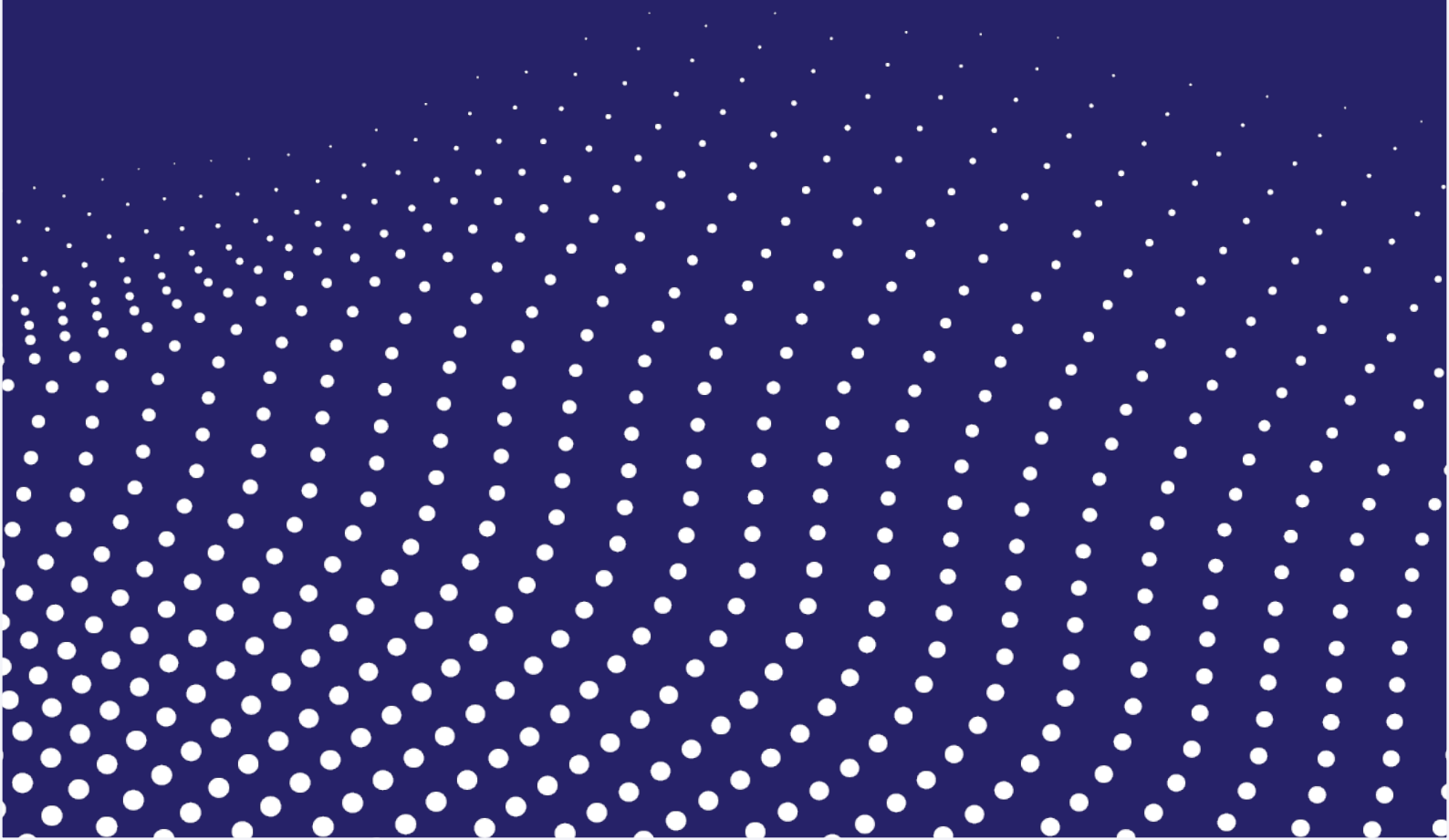
Figure 1 - AI in NSW: A brief timeline



Sources: Digital NSW, Investment NSW, Reserve Bank of Australia, University of Adelaide, ABC Science

Steering responsible AI beyond government

The NSW Government can implement practical policy measures to protect NSW residents' rights and interests with respect to the development and use of AI technologies across society. This first part of the report identifies several of these measures. rights and interests of people in NSW in the market and society more widely.



Law, guidance and oversight to manage risks and reduce harm

Subnational governments in general, including the NSW Government, have existing responsibility for a range of legislative and policy areas that intersect with AI. The impetus for regulation arises not because of the technology itself, but to respond to the "new sociotechnical landscape", including new "negative features" that technology may present, such as harms, risks, market failures and inequality.[4]

Laws of general application may relevantly cover some of these new negative features across the AI life cycle, while others may require additional regulatory intervention to provide sufficient protection to people and communities. In particular, the assumptions that underly existing laws may be inappropriate in the context of automated content generation.[5] This context is critical to underscore the importance of crafting a state government regulatory strategy that corresponds to the dimensions of the state's existing responsibilities.

The following sections on regulation cover:

- Common challenges for governments seeking to regulate the development and use of AI;
- Overarching principles that may be used to guide a regulatory strategy in NSW;
- An in-depth case study of the *Anti-Discrimination Act 1977* (NSW) as an example of an existing NSW regulatory framework that could be leveraged and adapted to respond to AI-related risks more appropriately; and
- Opportunities to selectively deploy audit and reporting requirements on the private sector to respond to high-risk uses of AI.

Regulating AI is a challenge for all governments

AI regulation is a broad, challenging field for all governments to navigate. Challenges in crafting an effective response include:

- **The pace of technological change:** AI technology is being developed, deployed and transformed at a very rapid pace. As the Australian Department of Industry, Science and Resources has acknowledged, "the speed and scale at which AI can be deployed (to generate benefits as well as cause potential harm) is one of the most significant policy challenges prompting calls for greater regulatory action."[6] Policymakers and regulators must be aware of the "orders of magnitude difference between the pace of technological change and that of regulatory adaptation"[7] and adjust regulatory approaches to enable faster adjustment. The mainstream emergence of generative AI, and particularly large language models over the course of the last year, has shown powerfully how developments in AI can rapidly disrupt existing ways of life.

- **The complexity of regulatory change:** AI is an emerging area of regulatory concern that has attracted attention from major international organisations and states including the United Nations[8], OECD,[9] the US[10] and the EU,[11] international standard-setting organisations,[12] as well as the Australian federal government.[13] In the US, President Biden recently signed an Executive Order on AI that requires private companies to share AI safety test results and other critical information during the product development phase in sensitive policy domains such as national security and public health.[14] The Executive Order also supports the development of new safety standards in the context of critical infrastructure and enhances assistance on privacy-preserving data training techniques across the private sector.[15] AI governance is an "unorganised field" with diverse stakeholders, occupying various types of initiatives and implementations.[16] From 2016 to 2022, countries passed 123 AI-related bills, at an increasing pace over recent years.[17] This means that, while there are valuable resources and opportunities for mutual learning across the global community, the regulatory environment concerning AI is becoming increasingly multi-dimensional. Moreover, AI systems are technically complex, often relying on very large amounts of data, and interacting in complicated ways with existing legal, economic and social institutions. The success of any regulatory intervention thus relies on ensuring that policymakers, regulators and dispute resolution bodies have the relevant sociotechnical knowledge to ensure that any regulation achieves its purposes.

- **New permutations of risk arising from AI:** The OECD Council on Artificial Intelligence has emphasised that the transformative effects of AI "may have disparate effects within, and between societies and economies, notably regarding economic shifts, competition, transitions in the labour market, inequalities, and implications for democracy and human rights, privacy and data protection, and digital security."[18] Risks include the use of AI to generate "deepfakes" for deceitful purposes, creating misinformation and disinformation, and causing harm to persons, inaccuracies from AI models, algorithmic bias, effects on privacy accountability and transparency, validity and reliability of data and the exacerbating of market inequalities and restriction of competition.[19]
- **The 'black box' effect:** The 'black box' effect is a key example, whereby it is difficult or impossible for people to understand how an AI system generates its outputs.[20] This can arise due to the complexity of algorithms, their use of large amounts of data to generate outputs, and the lack of transparency or 'reasons' as to how a particular result is generated. In fact, large technology companies may deliberately design "algorithmic opacity" into their products to maintain their competitive advantage and to preserve their trade secrets.[21] While some risks are entirely novel, and others are an amplification of existing technology risks, both cases require creative responses from regulators.
- **Lack of public trust and understanding:** Although recent research has identified that many Australian organisations already rely on AI-driven systems,[22] international and Australian research has found persistently low levels of public trust in AI.[23] Indeed, fear and worry were identified as the dominant emotions that people in Australia have towards AI,[24] and only 44 per cent of Australians perceive that the benefits of AI applications outweigh the associated risks.[25] Moreover, when compared internationally, Australians are amongst those least comfortable specifically with the deployment of AI in the workplace, particularly in relation to human resources and the monitoring, evaluation and recruitment of employees.[26]
- **Information asymmetries:** Information asymmetries exist between those who develop and deploy AI systems, and those who are affected by, or seek to regulate, the use of these systems. This can perpetuate existing power imbalances between groups (such as landlords and tenants, employers and employees, companies, and consumers) and prevent effective regulatory interventions.
- **Allocating responsibility:** It can be difficult to allocate responsibility for AI-powered decision-making to a person. As the Australian Human Rights Commission (AHRC) has noted, with respect to questions of liability, "some complexities can arise – either where an AI-informed decision-making system operates largely autonomously, or where numerous parties are involved in designing, developing and using the system."[27] This poses challenges for regulators, tribunals and courts who seek to respond to AI-generated decisions or conduct, as there is no clear line of responsibility.

NSW regulation should drive accountability in the use of AI

Where AI is used in a manner that has the potential to affect a person's rights or interests, accountability should be embedded across the entire cycle of its use. There are several high-level principles the NSW Government should consider adopting as part of its strategy for regulating AI, covering notification, explainability, dispute and appeal pathways, and liability.

Figure 2 - High-level principles to drive accountability in the market

Notification

Firstly, where AI is used as a critical or determinative part of processes that affect people's rights and interests, people should be notified of its use. Information asymmetry between those who deploy AI systems and those who might be affected by them prevents people from raising concerns or seeking remedies. Therefore, notification is critical to allow consumers to make informed decisions as to whether they will engage with a company or person and to facilitate a dialogue about concerns or disputes with respect to AI use.[28]

AHRC previously recommended that the Australian Government introduce legislation requiring notification to individuals of the material use of AI in decision-making processes affecting the rights of an individual.[29]

The NSW Government could consider implementing a broad-based requirement for notification. Ideally, any such intervention would align to the greatest extent possible with the requirements of other jurisdictions in Australia and globally. Notification requirements could apply to particular types of AI, such as generative AI that is being used to create text, images, audio or video content that resembles authentic content,[30] or to particular uses of AI, such as in the use of AI to screen and hire employees.[31]

Explainability

The use of AI should not prevent people, regulatory authorities, dispute resolution bodies and courts from obtaining information, reasons, or an explanation of AI-informed decisions or professional practices. However, there is a risk that AI systems may not give reasons for their decisions, and any reasons that are given can be opaque and difficult to understand,[32] or be an inauthentic summary of the decision-making process.

It is important to note that there is no general right of individuals to obtain reasons for decisions that affect them. However, the rule of law demands that a court be able to determine whether the law has been followed. Therefore, even where no explicit legal right to reasons exists, a court may investigate an impugned or disputed decision. However, as noted by the AHRC, in circumstances where private sector actors are already required to provide reasons, or otherwise provide information, policymakers should ensure that the use of AI does not undermine the operation of these provisions. Therefore, when a company or industry is subject to an obligation to provide reasons, whether by legal requirement or by a voluntary code, the use of AI should not prevent the discharge of these obligations.[33]

It is also of fundamental importance that existing powers of regulatory, oversight, or dispute resolution bodies to obtain information or material are not undermined by the use of AI. As the AHRC recommended in its 2021 ***Human Rights and Technology*** report:

> "[W]here a court, or regulatory, oversight or dispute resolution body, has the power to order the production of information or other material from a corporation or other legal person:
>
> (a) for the avoidance of doubt, the person must comply with this order even where the person uses a form of technology, such as AI, that makes it difficult to comply with the order;
>
> (b) if the person fails to comply with the order because of the technology the person uses, the body may draw an adverse inference about the decision-making process or other related matters."[34]

Adopting reforms in NSW with respect to relevant legislation would ensure that the use of AI does not prevent regulatory authorities, tribunals and courts from obtaining information. For example, the NSW Civil and Administrative Tribunal (NCAT), to which anti-discrimination complaints can be referred by the President of the Anti-Discrimination Board, has the power under section 46 of the ***Civil and Administrative Tribunal Act 2013*** (NSW) to call and examine witnesses and compel them to answer any questions that the tribunal considers relevant, and under section 48 to issue a summons requiring a person to attend and produce documents or other things to the tribunal. These sections could be amended to ensure that the use of AI does not allow persons to avoid complying with the obligation to provide information or materials to the Anti-Discrimination Board or NCAT.

Dispute and appeal pathways

People in NSW should be able to raise complaints, disputes and appeals of decisions through clear pathways, that are appropriately resourced to meet their needs in respect of AI and technology-related matters. This requires that people are provided with sufficient information to recognise that they have been subject to harm and that pathways exist for remedy.

Existing NSW regulatory schemes provide pathways for people to make complaints, appeal against decisions, and seek dispute resolution. For example, tenants, residents, landlords and agents can lodge complaints with Fair Trading NSW with respect to real estate and property-related disputes. This process involves a Fair Trading officer providing the parties with relevant information to assist them in coming to a mutual agreement.[35] Similarly, anti-discrimination law allows persons who believe they have been victims of discrimination to lodge a complaint with Anti-Discrimination NSW.[36] The NCAT can also hear referred matters and appeals against decisions in areas including anti-discrimination law, tenancy issues, guardianship decisions, and professional disciplinary matters.[37]

These complaint, dispute resolution and appeal mechanisms should remain available to NSW citizens in circumstances where AI has been used to make a decision or provide a service. However, according to researchers from the ARC Centre of Excellence for Automated Decision-Making and Society, existing enforcement regimes may be unfit to respond to breaches of rights arising from the use of AI due to a lack of appropriate legal remedies, a lack of a clearly navigable and simple pathway for remedy, and insufficiently resourced regulators.[38] It is, therefore, critical that all complaints and dispute resolution bodies are sufficiently resourced to manage the complexity that may arise where AI is relevant to a dispute.

Liability

The use of AI should not be allowed to undermine existing legal mechanisms for allocating responsibility for conduct that causes wrongful harm to the rights and interests of persons.

Allocating legal liability often requires identifying a legal person who is responsible for the relevant conduct. In circumstances where an AI system, which is not a "legal person", acts autonomously, or where multiple parties (such as the developer of an AI system, the procurer and the end user) are involved, it can be difficult to identify who, if anyone, is legally responsible for conduct that might harm legal rights or interests.[39]

AI can complicate other fundamental legal concepts, such as causation, reasonableness, and professional conduct, which are central to allocating responsibility in various areas of law including anti-discrimination law, professional negligence and other tortious liability.[40]

Allocation of legal liability could be enacted through a rebuttable presumption, as previously recommended to the Commonwealth by the AHRC, to provide that "where a corporation or other legal person is responsible for making a decision, that entity is legally liable for the decision regardless of how it is made, including whether the decision is automated or is made using AI".[41] Liability could also be enacted through a stricter presumption that assumes the liability of the corporation or legal person, thereby providing greater protection to claimants, and associated avenues for those liable to pursue actions against suppliers as relevant. Where applicable, the NSW Government should consider implementing reforms to ensure clarity with respect to legal responsibility for conduct when AI is used.

The application of these regulatory principles to the *Anti-Discrimination Act 1977* (NSW) is considered in an extended case study, contained in Appendix A.

Policy opportunities

1. Establish a regulatory strategy for AI development and use in the market, reflecting the AI regulation principles of notification, explainability, disputes and appeals, and liability.
2. Undertake a review of existing legislation to identify points requiring clarification in the form of new legislative guidance, or legislative amendment, to reflect the AI regulation principles of notification, explainability, disputes and appeals, and liability.
3. Resource regulators appropriately to develop the capability and capacity to respond to emerging risks presented by AI.

Audit and reporting requirements can be selectively deployed

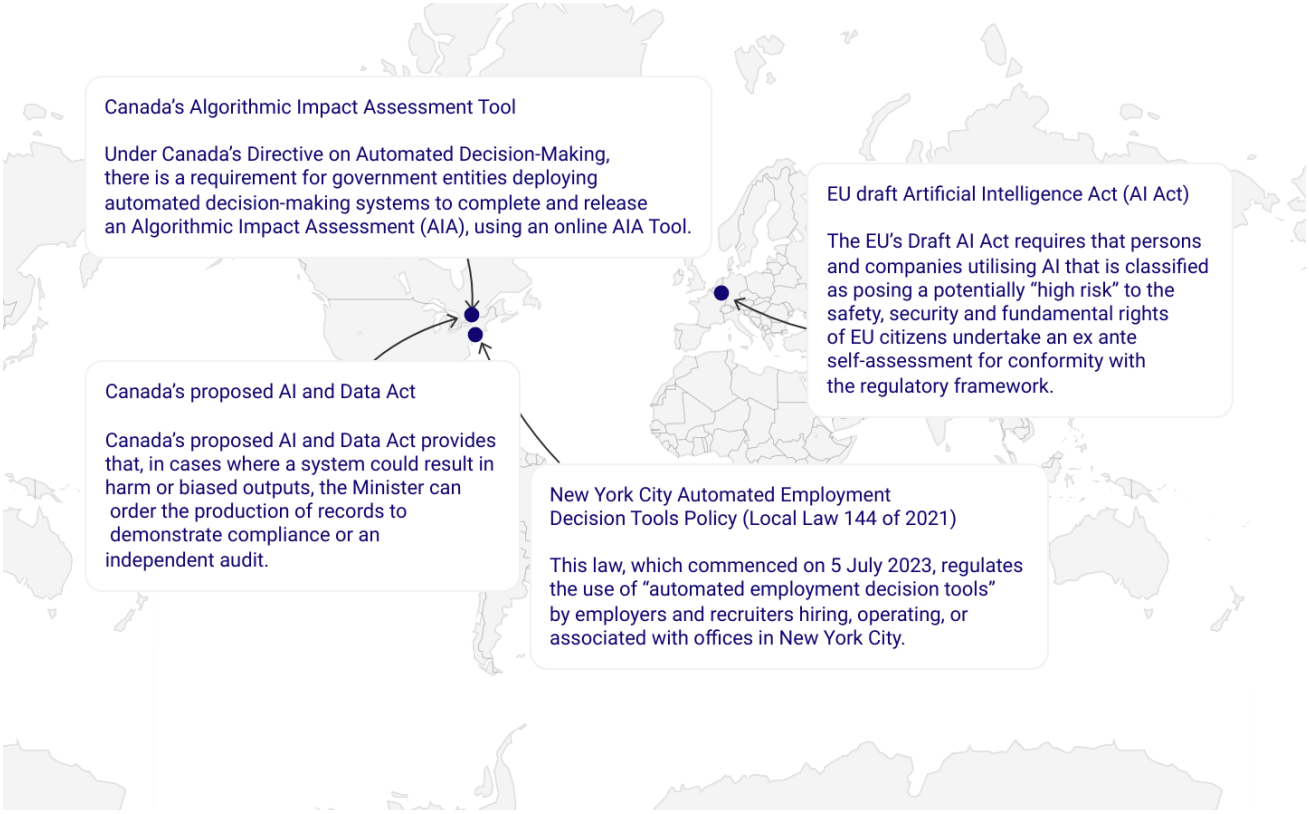
Jurisdictions around the world have implemented, or are considering requirements, that entities using AI be subject to regular "audits" or reporting requirements. These requirements can be implemented with respect to the technology's particular purpose, ways of operating, or the risks involved, such as potential discrimination. The NSW Government could consider using similar regulatory tools to drive transparency and compliance with legal obligations.

In this context, auditing refers to a "structured process by which an organisation's present or past behaviour is assessed for consistency with relevant principles, regulations, or norms."[42] To be relevant and effective in this area, auditing must occur regularly, if not continuously. This might involve requirements for organisations to be assessed by an independent third party to determine whether the use of AI complies with relevant laws, ethical principles, or other standards. In other jurisdictions, assessment and reporting requirements such as "algorithmic impact assessments" require organisations to consider and document the potential impact of their use of AI .[43]

Audits, impact assessments or other documentation and reporting requirements have been proposed to ensure that technology providers, users and policymakers can identify and mitigate risks associated with AI systems by:

- Ensuring procedural regularity and transparency and contributing to good governance;
- Encouraging proactivity in the design of AI systems that helps identify risks and prevent harm before it occurs; and
- Facilitating objective and professional evaluations of the use of AI systems.[44]

Figure 3: Audits to manage risks of AI – Examples from overseas



Auditing is an increasingly common tool used to manage the risks of AI

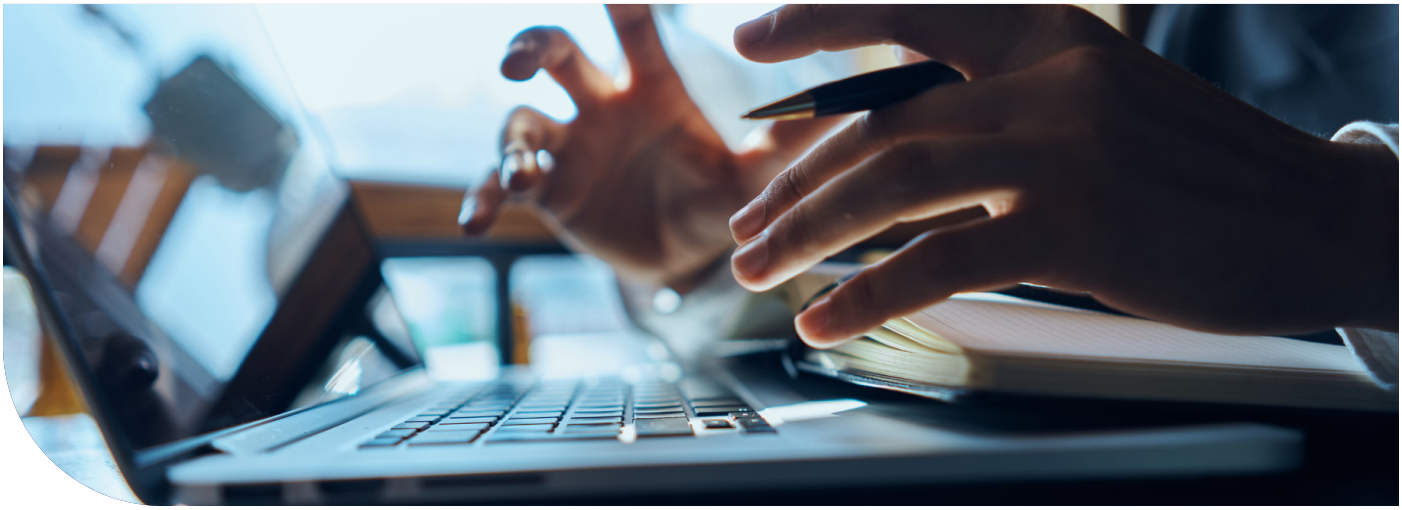
Audits have been employed and proposed in various jurisdictions with respect to AI, concerning discrimination as well as other risks. Examples include the following:

- **New York City Automated Employment Decision Tools Policy (Local Law 144 of 2021):** This law, which commenced on 5 July 2023, regulates the use of "automated employment decision tools" by employers and recruiters hiring, operating, or associated with offices in New York City. The law prohibits employers and employment agencies from using AI tools in recruitment or promotion processes and establishes fines for non-compliance. The exception is if such a tool has been subject to an independent external bias audit in the previous year, the results of the independent audit are publicly available, and notice requirements to employees or job candidates are satisfied. Some experts, as well as media commentators, have nevertheless expressed scepticism as to whether the fines are sufficient to deter non-compliance, and whether other aspects of the law have been "watered down", reducing its effectiveness.[45]

- **EU draft Artificial Intelligence Act (AI Act):** The EU's Draft AI Act requires that persons and companies utilising AI – that is classified as posing a potentially "high risk" to the safety, security and fundamental rights of EU citizens – undertake an self-assessment prior to implementation (or assessment by a regulatory body in the case of publicly deployed biometric surveillance systems) for conformity with the regulatory framework. These self-assessments are to be registered in an EU-wide database. There has been some criticism of the EU's reporting requirements, due to its reliance on self-assessment.[46]

- **Canada's Algorithmic Impact Assessment Tool:** Under Canada's Directive on Automated Decision-Making, currently in force, there is a requirement for government entities deploying automated decision-making systems to complete and release an Algorithmic Impact Assessment (AIA), using an online AIA Tool.[47] The questionnaire asks organisations around 80 questions concerning business process, data, system design, and algorithm. The results display an impact level for an automation project and provide information about applicable requirements under the Directive. The Government of Canada does not have access to the results of AIAs, which are only stored on the local hard drive of the person who filled it out.

- **Canada's proposed AI and Data Act:** Canada's proposed AI and Data Act provides that, in cases where a system could result in harm or biased outputs, the relevant minister can order the production of records to demonstrate compliance or an independent audit. These can only be conducted by appropriately qualified external auditors and is fully payable by those being audited. Those being audited must give all assistance that is reasonably required to enable the audit to be undertaken, including by providing any records or other information that is reasonably required including records or other information. The relevant minister is also empowered to order cessation of the use of a system or disclose publicly information regarding contraventions of the Act, in circumstances where there is a risk of imminent harm.[48]



Audit and reporting requirements should be carefully designed

Policymakers should consider when auditing or reporting are required

Audits and reporting requirements can be imposed universally or only in specific contexts. At the broadest end of the spectrum, all bodies designing or deploying AI systems could be required to carry out or submit to an assessment of their use of AI to ensure that it complies with relevant human rights and other obligations. This is the case in Canada with respect to the AIA.[49] At the narrower end of the spectrum, audits or reporting might only be required in circumstances where a particular use of AI has been determined to pose a particular risk to safety, privacy or fundamental rights, as in the EU,[50] or in a particular use-case, as in New York for employment.[51]

In the NSW context, consideration might be given to using a risk-based approach to deploying this requirement towards either processes, such as hiring and recruitment, or industries, such as housing, where this regulatory intervention will have the greatest impact on people and communities. In determining the risk rating of applications of technology, it is critical to consider multiple dimensions of risk and harm (i.e. brief vs. sustained risk, reversible and irreversible risks, and the potentially lasting impacts of harm occurring even if reversed).[52] Policymakers should also consider harnessing a range of perspectives in rating risk. Developers and parties may have incentives to underestimate risk, while people with lived experience of relevant harms may be well-placed to contribute perspectives about these harms.[53] Governments may consider whether accreditation of specialist auditors might be an effective means of ensuring quality, given this is a new area of practice.

Policymakers should consider the target for audit and reporting requirements

Auditing or reporting requirements can potentially apply to a broad range of organisations and persons involved. These include the developers of AI systems, those procuring AI systems, and the end users of AI systems. Further, audits and reporting can address three levels of activity: governance audits of the companies developing or deploying AI, model audits of AI systems before their public release, and application audits once an AI system is deployed in the real world.[54] Audit and reporting requirements should be deployed strategically, at all three levels, and in concert with other measures, to ensure the effective translation of principles and standards into practice.[55]

Policymakers should consider options for oversight and enforcement

As the examples explored above indicate, there are several different models for auditing or reporting requirements. One important distinction between them concerns who carries out the relevant assessment. This can range from self-assessments and internal audits to auditing by private expert organisations and, ultimately, to audits by public authorities or government departments.[56] Further, the degree of public disclosure can range from private assessments that are not communicated to any public register, to the mandatory provision of audit or assessment results to a publicly available register.[57] Finally, there can be different degrees of oversight and enforcement of auditing requirements and a variety of sanctions for non-compliance. This could range from voluntary reporting or certification regimes that might award a form of "trust mark" to compliant organisations to hard enforcement by a regulatory with the capacity to issue fines or penalties for non-compliance.[58] The NSW Government can use its levers in education and skills to contribute to developing and maturing a professional workforce to meet the demand for these increasingly important services.

Policymakers should strive for effectiveness but reduce burdens where possible

More stringent auditing requirements, with independent audits, public disclosure, and rigorous enforcement can encourage higher levels of compliance, secure public trust, and ensure that contraventions can be addressed in a timely fashion. However, this can also be more costly, both for government and in terms of the regulatory burden imposed on the private sector. It also requires significant institutional capacity on the part of auditors, auditees and regulators.

In terms of the burden on regulators and government, the NSW Government could explore opportunities to develop and leverage an AI assurance industry in NSW. Engaging supportively with the private sector to develop this capability could promote access and uptake of AI assurance processes across firms. Governments can actively work with this industry to advance its own democratic goals, including potentially issuing guidance that improves the goals and standards of assurance processes provided by private firms.[59] Other jurisdictions are working to shape and develop this industry in line with their regulatory strategy, such as the UK which developed and released a roadmap to an effective AI assurance ecosystem.

Additionally, NSW Government can also consider implementing less stringent requirements proportionate to risk, which might include self-assessments and disclosure requirements. These can be easier to implement broadly and at a lower cost, and may be considered sufficient if the level of risk to be addressed is lower. Assessments of proportionality might also contemplate the size and market share of firms: larger and/or more established firms might be subjected to more stringent obligations, compliance with which might unduly burden the innovative capacity of a smaller firm. Policymakers should consider factors including the desired effect, capability and capacity among relevant actors, and the regulatory impact when designing any such mechanism.

In terms of the burden on the private sector, working towards harmonisation of regulation and standards is a critical means to reduce the burden of any intervention in this space. Harmonisation allows organisations to streamline their processes across a range of jurisdictions, thereby enhancing the prospects of competitiveness across a range of markets. Bolstering the assurance industry in NSW might also support firms to take both proactive and reactive measures to ensure the responsible use of AI, proportionate to their size and resources.

Policy opportunities

4. Implement auditing and public reporting requirements with respect to AI systems in response to the greatest risks of harm to the community arising from AI.
5. Stimulate and support the development of an effective local AI assurance industry to enable its assurance activities across firms in the NSW economy, including through stewardship of assurance standards.

Shaping non-government development and use of AI in NSW

Governments can consider their role as a "shaper" of markets with respect to AI, by moving beyond merely "fixing" market failures.[60] In doing so, governments can consider how they can help create and shape markets for better AI innovation that creates real value for people and communities.

The following sections on shaping cover:

- Opportunities for government to use procurement as a lever of indirect influence over the market.
- Ways in which a state government can act to directly support and encourage responsible practice across the market.

Procurement can be used as a lever of indirect influence

Government procurement can be used as a lever of horizontal influence over the market

AI technologies create unique challenges for strategic procurement in light of their complexity, rapid advancement, and uncertain development trajectories. This necessitates strategic approaches to influencing the development and purchase of responsible and ethical AI products via NSW Government procurement policy and processes. Procurement can play an important role in the process of shaping AI in line with social benefits.[61]

The existing market for AI technologies in NSW is strong, with a cluster of advanced computing businesses broadly comparable in scale to international competitors such as Berlin, Singapore and Toronto.[62] Another indicator is the size of the existing digital technology workforce in NSW, with 41 per cent of all Australian software and application programmers based in the state, and 45 per cent of all Australian AI businesses.[63] In this context, procurement can be a lever for government to influence private sector activities towards the creation of social and public value by steering the early development of emerging technologies towards public benefit.[64] Within the NSW government, approximately 15,000 public servants are involved in procurement, with an annual spend of $35-40 billion.[65] Governments, including the NSW Government, are also likely to gradually increase the frequency and scale of procurement of AI-enabled products and services in line with their advancement and ubiquity.

In addition, government procurement can drive acceptance of standards across the market. The flow-on impact of standards adopted by governments may be greater where public spending has a significant influence on market activities. Factors that will influence the degree of market influence that can be wielded by the NSW Government through its procurement activities include:

- **Market coverage:** Due to its substantial responsibility for service delivery, the NSW Government is likely to hold a significant market share of overall procurement of specialist AI-enabled products within certain sectors, but a relatively small share in other sectors. In areas where the NSW Government has greater market share, it will have more shaping power with respect to AI technologies, which may inform prioritisation of effort in this area.
- **Effectiveness of "off-the-shelf" solutions:** Where "off-the-shelf" solutions are considered fit-for-purpose and good value for money, the opportunities to influence the wider market by working to implement responsible practices in development may be limited.

From this perspective, procurement requirements can articulate a government's expectations of minimal assurances for private sector service delivery and investment behaviours, helping to drive socially responsible outcomes in alignment with government policy priorities and public expectations.

Steering AI technologies towards socially beneficial outcomes

Procurement can also be leveraged to create horizontal, market-shaping effects, influencing the direction of innovation across the economy to tackle complex policy problems.[66] Governments often act as a "lead user" for emerging technologies, "enabling the formation of embryonic markets with potential for further diffusion."[67] Procurement can directly incentivise the development of new technologies, legitimise new product standards and create new markets, shaping the overall composition of entire industries.[68] Accordingly, procurement is not merely an administrative task, but a growing strategic priority for governments to address complex social, economic and environmental challenges.[69]

Education and edtech

Education is an example of an area where AI applications have significant impact and potential, opening new opportunities for personalised learning and other data-driven student interventions.[70] The NSW Department of Education is the largest provider of education in Australia and is a significant procurer of technology solutions for students and staff. Accordingly, strategic government procurement in education can be leveraged to achieve wider social benefit and to address entrenched policy challenges.

AI technologies could be a powerful tool to help tackle persistent divides in educational outcomes, particularly for students from vulnerable backgrounds. On the other hand, the introduction of these new technologies could also heighten educational inequalities by reinforcing existing resourcing inequities and by entrenching a new digital divide. This is a significant concern, given that technology companies often have monopoly power in certain markets, and their products are increasingly embedded across all levels of the education system.[71] Professor Kalervo Gulson has cautioned that that AI technologies can influence the purposes of the education system more widely, such as by impacting objectives like promoting democratic equality, social efficiency, and social mobility.

Nevertheless, procurement policy remains a powerful lever to shape the direction of these technologies to improve equity and student learning outcomes. Professor Leslie Loble argues that procurement policy can empower educators to shape edtech policy and product design more directly.[72] This can be achieved through integrating co-design methods into procurement contracts, thereby encouraging a greater degree of educator-led innovation. This strategy could be further augmented through the use of open technology standards that can encourage greater product innovation, reduce ICT costs, and prevent the problem of "provider capture".

Loble also sees procurement policy as an opportunity to direct greater private sector edtech investment towards tackling educational disadvantage. For instance, it could lead to the creation of new targeted products for students from disadvantaged backgrounds or students with special needs.

Finally, there Is potential to steer edtech development via procurement so that it better guards against potential discrimination and bias. This could be achieved by setting specific thresholds for AI explainability, integrating improved privacy measures, and strengthening personal data protections as part of procurement contracts.

Loble stresses that the effectiveness of AI and technological solutions in education relies heavily on the quality of products, their effective integration into teacher-led instruction, and a well-aligned edtech market with appropriate governance and institutions.

Loble's recommendations are supported by Bello and Gulson, who argue that governments can strengthen edtech procurement by aligning international human rights frameworks with relevant procurement policies and AI specific legislation or guidelines.[73] This process can help policymakers to navigate potential trade-offs between protecting individual rights and using AI to improve wider education system performance. Ultimately, procurement policy can play a pivotal role in shaping the ethical and effective development and usage of AI in education, serving as a catalyst for greater innovation while fostering inclusion and improving student outcomes.

The NSW Government should use its significant procurement footprint strategically

It is important that the NSW Government procurement frameworks are positioned to effectively shape this emerging market towards ethical and responsible products and socially beneficial outcomes.

The NSW Government Procurement Policy Framework articulates the requirements and operational structure for NSW Government procurement, serving as a definitive guide on procurement rules. NSW Government procurers are required to periodically test their compliance with the mandatory requirements of the framework, and other NSW Government Procurement Board policies and directions.[74] The framework also articulates "best practice" principles to be implemented where mandatory procurement requirements do not apply.

The NSW Government AI Assurance Framework must be used when NSW is procuring custom AI systems, customisable AI systems and for all development projects using generic AI platforms.[75] It is also used as part of periodic review processes of AI systems. For small-scale AI projects, NSW Government agencies are only obliged to undertake a self-assessment against the standards of the framework. The Assurance Framework does not apply to widely commercially available AI systems, or solutions that are not being customised. The framework poses questions that project teams should address at each phase of AI system deployment and operation. This assurance process is augmented by requiring the submission of qualifying assessments to the NSW Government AI Review Committee where appropriate, which can make non-binding recommendations to mitigate any risks identified. The NSW AI Review Committee is composed of industry, academic and government experts. Projects must be submitted for review if they are valued over $5 million and/or are funded from the Digital Restart Fund. However, the Committee can be convened for a range of discretionary advisory purposes. The AI Assurance Framework includes the NSW Government AI Ethics Principles, which outline explicit principles to ensure best practice use of AI with a focus on trust, transparency, community benefit, fairness, privacy and accountability.[76] These principles are directly translated into the structure and operational procedures of the framework.

AI procurement in Estonia

The Estonian Government leverages its procurement processes to improve public services via AI technologies. The program successfully launched an AI voice and text-based virtual assistant tool in 2022, enabling citizens to complete tasks that previously required making phone calls or visits to government offices, such as completing official documentation forms, making payments and registering life events.[77] While reducing government costs, it is also designed to help improve equity access to government services via uplifting user-friendliness and accessibility. The Estonian Government announced in mid-2022 that they were going to embark on another procurement round to further develop the platform, utilising an open-source approach to build public trust, and to improve product design. Such an approach was followed because the procurement contract was global, allowing development partners to work on the source code collaboratively across the world.[78] The procurement strategy was also centred on collaboration with the private sector, including a key partnership with Microsoft.[79]

Public servants must be supported to enact socially beneficial and strategic AI procurement

Previous research by the NSW Government, as part of the AI Strategy development process, noted that the complexity of the regulatory landscape presents challenges for AI procurement. Procurement processes are often too slow and make it difficult for AI start-ups and innovators to collaborate with the NSW Government.[80] The strategy recommended that the NSW Government should identify strategies to "allow for more timely adoption of emerging technologies" and to improve opportunities to innovate and experiment with AI solutions through the procurement process.[81]

In addition, government stakeholders noted their difficulties in identifying appropriate AI solution providers with the right experience and capabilities and with a suitable track record of delivery while representing value for money.[82] NSW Government stakeholders also expressed uncertainty as to how to determine whether "off-the-shelf" AI solutions were suitable for their circumstances. In addition, they also reported a greater need for suppliers that are capable of committing to longer-term testing and refinement of AI solutions.[83]

Finally, NSW Government procurers need better information as to how to mitigate AI risks as part of procurement processes, and require greater guidance as to best practices for effective AI procurement. There is a specific need to improve the capacities of NSW Government procurers to address critical AI design issues, including data and intellectual property management and privacy protection.[84]

Guidance for NSW Government procurement can be enhanced

Currently, the NSW Government Procurement Policy Framework does not refer explicitly to the NSW Government AI Assurance Framework and its guidelines with respect to ensuring transparency, privacy and security, fairness and accountability principles. As the key reference policy for staff across the public sector who are undertaking procurement activities, referencing the AI Assurance Framework would encourage compliance with the AI Assurance Framework's mandatory requirements.

In addition, integrating the AI Assurance Framework within the Procurement Policy Framework would assist in addressing concerns raised by NSW Government procurers in relation to the need for improved guidance on AI risk mitigation, adopting best practices for AI procurement, and gaining greater knowledge about managing key AI procurement issues, such as data and privacy protections. The integration of these policy documents could also be supported by initiatives to build the awareness and capability of public servants about their obligations.

Policy opportunity

6. Explicitly refer to or integrate the AI Assurance Framework – and the obligations in confers in respect to procurement - into the NSW Government Procurement Policy Framework.

Government can send market signals through procurement

Communicating more strategically about the expectations and requirements of the NSW Government with respect to ethical and responsible AI could drive responsible practices across the market, for the benefit of NSW people and communities. The NSW Government's Supplier's Code of Conduct, which outlines key principles for providing goods and services to the NSW public sector, does not currently refer to the mandatory AI Ethics Policy under the AI Assurance Framework. For example, there are no specific ethical principles in relation to data quality assurance specified within the Supplier Code of Conduct, or specific ethical principles to steer the design of AI products or services in support of diversity concerns. Given that the Supplier Code of Conduct outlines a "minimum set of expectations and behaviours for doing business with NSW Government", integrating elements of the AI Ethics Policy may help the private sector to better understand government procurement requirements with respect to ethical and responsible AI.[85]

Another opportunity to shape the private sector is utilising the expertise of the NSW Government AI Review Committee. The AI Review Committee exists to support government to ensure that AI solutions align with ethical and transparent standards, helping to safeguard public trust. It also plays a significant role in procurement processes by ensuring that mandatory policy requirements have been considered under the auspices of the AI Assurance Framework.

At present, there is limited publicly available information about its decisions, potentially reducing its ability to influence the actions of the private sector. The AI Review Committee could be resourced and encouraged to publish, where possible, case studies and assurance reviews. This would provide guidance to the private sector by developing an understanding of the NSW Government's expectations with respect to ethical and responsible AI, as well as common issues identified during the review process. This understanding, in turn, may help the private sector to improve its offerings of ethical and responsible AI solutions across the wider market.

Policy opportunities

7. **Amend the NSW Supplier Code of Conduct to include elements of the AI Ethics Policy to clearly outline suppliers' responsibilities in respect of responsible AI.**
8. **Resource and explicitly encourage the NSW AI Review Committee to publish educative case studies and assurance reviews communicating lessons from its work.**

Government can directly support and encourage responsible practice

Using standards and codes of practice to influence non-government actors

Codes of practice can be used to explain how non-government actors can meet their binding legal obligations, and outline aspirational or ideal standards of conduct for a particular industry. Codes of practice can originate from a variety of sources and have broad or narrow targets. Codes might be issued to guide adherence to legislative obligations (for example, under the *Anti-Discrimination Act 1977* (NSW)), to articulate standards for a particular industry or type of actor (for example, a code of practice for lawyers or the financial services industry), or to articulate standards for carrying out a particular function (for example, recruitment activities). A range of legislation in NSW already provides for the development and implementation of codes of practice in specific areas or by reference to specific legislation.

Adopting existing standards can be efficient for governments by avoiding the duplication of highly technical work, and drive efficiency for businesses that currently operate, or intend to expand operations, across multiple jurisdictions. Any NSW Government project issuing codes of practice or guidelines should be careful not to duplicate existing codes or standards.

As discussed later in this report, an Emerging Technology Commissioner or other body could be empowered to adopt existing standards as their own standards, facilitating harmonisation of NSW with international best practice. Legislation may also adopt or refer to established Australian or international standards.

Some NSW legislation provides for the issuing of codes of practice to guide compliance with existing legislative obligations. For example, the NSW Anti-Discrimination Board has the power under section 120A of the ***Anti-Discrimination Act 1977*** (NSW) to develop and promote codes of practice to provide guidance to persons in specific areas of conduct as to what kinds of activity might constitute unlawful discrimination and how to limit, avoid, or restrict the risk of contravening the Act. These codes of practice are not legally binding and cannot, of themselves, render conduct lawful or unlawful. However, evidence of compliance with, or contravention of, a code may be considered by the President of Anti-Discrimination NSW and the Tribunal in the exercise of functions under the Act, the ***Administrative Decisions Review Act 1997*** (NSW) and the ***Civil and Administrative Tribunal Act 2013*** (NSW). Codes of practice issued as a form of legislative guidance are influential insofar as adherence to them, or contravention of them, may form a material part of consideration as to whether legislation has been adhered to.

Other legislation provides for the development of codes of practice to influence behaviour beyond adherence to existing legal obligations. For example, under section 27 of the ***Modern Slavery Act 2018*** (NSW), the NSW Anti-slavery Commissioner has the power to "develop, and make publicly available, codes of practice for the purpose of providing guidance in identifying modern slavery taking place within the supply chains of organisations and steps that can be taken to remediate or monitor identified risks."[86] The Commissioner may refer to or incorporate any other published standard as a part of this power to issue codes, allowing the Commissioner to signal the relevance of existing standards.[87] Using this power, the Clean Energy Council and the Anti-Slavery Commissioner are developing a code of practice in partnership to help manage modern slavery risks in renewable energy supply chains in NSW. This voluntary code, alongside the development of implementation guides, will set a formal benchmark across the renewable energy supply chain, and will be integrated as part of the NSW Government's procurement of renewable energy solutions.[88]

Codes of practice should complement, rather than replace, other interventions

Codes of practice may be a useful tool in certain contexts, and there are levers available to government that can create pressure to adhere to codes across the market. For example, government can incorporate relevant codes of practice into its procurement of goods and services. It can require adherence to codes as a requirement of becoming a pre-qualified supplier or joining a panel of pre-qualified suppliers in relevant categories, which is an effective way of requiring longer-term change within firms that are existing or prospective suppliers.

Voluntary codes, nevertheless, have weaker legal influence than codes that explicitly guide the discharge of existing legal obligations. In addition, a wide range of standards, guidelines and codes of practice in Australia and internationally already address aspects of AI, some of which are summarised in Appendix B.

Government can partner with relevant non-government organisations to encourage the adoption of codes of practice at an industry or sector level, or by encouraging the uptake of voluntary codes by large organisations across the market. Firms and industry bodies might be incentivised to partake for several reasons, including:

- positive publicity associated with taking a responsible stance in respect of a high-profile social risk;
- visible adherence to and discharge of ESG obligations;
- increased opportunities to participate as a supplier to government and other organisations that have adopted a given code as a requirement; and
- the free availability of expert advice from an independent expert, such as the Emerging Technology Commissioner.

Codes should be developed sparingly and adopted selectively

Questions to consider when determining whether to develop or adopt a code of practice include:

- Is there an existing regulatory body in NSW with the power to issue codes of practice or guidance that has persuasive force in determining adherence to or contraventions of legal obligations? (i.e., the Anti-Discrimination Board's power under section 120A). In these circumstances, the NSW Government should consider resourcing the relevant regulatory body, or another body such as an Emerging Technology Commissioner as discussed subsequently in this report, that could work with that body, to develop relevant guidelines.
- If there is no existing legislative power to develop guidelines, is there an existing standard, guideline or code of practice that covers the field? If so, could this existing standard be adopted and enforced as a NSW standard?
- If there is no existing standard that covers the field, and the issue is one where specific NSW guidance would be of value, the body developing the code of practice should consider:
  - How can experts, persons with lived experience of potential harms, and representatives of relevant industries or professions be involved in developing relevant guidance?
  - What levers are available to create pressure on non-government actors to adhere to the code?

Appendix B includes a range of examples of guidelines drawn from other Australian and international jurisdictions.

Policy opportunity

9. **Adopt and/or endorse codes of practice, to guide compliance with existing legal obligations, to integrate into procurement requirements, and for voluntary adoption more broadly across the private sector.**

Supporting responsible development through "regulatory sandboxes" and "safe harbours"

Regulatory "sandboxes" are an experimentation tool that can be used to test new approaches and legal provisions, outside of existing regulatory structures.[89]

Defining a regulatory sandbox

According to the OECD, "A regulatory sandbox is a limited form of regulatory waiver or flexibility for firms, enabling them to test new business models with reduced requirements. It often includes mechanisms to ensure overarching objectives such as consumer protection. Regulatory sandboxes are typically organised and administered on a case-by-case basis by the relevant authorities (Attrey, 2020[10]). Their main characteristics are that they are: (1) temporary; (2) use a trial-and-error approach; and (3) involve collaboration and iteration between stakeholders. Regulatory sandboxes require thorough design and testing with robust methodological and assessment frameworks, evaluating feasibility, demand, potential outcomes, and collateral effects."[90]

Using sandboxes, governments engage non-government organisations, particularly from the private sector, to test innovative products and services within a controlled environment, where standing legal and compliance requirements are waived.[91] The OECD recommends the use of such experimental tools to governments seeking innovative ways to respond to rapid technology innovation cycles, suggesting that in this context regulators are likely to require experimentation to inform decisions.[92]

Sandboxes present mutual opportunities for government, firms, and consumers. For governments, sandboxes can support fast but evidence-based AI regulatory responses, as governments are increasingly called upon to respond to new technologies at a faster rate than traditional legislative processes allow.[93]

In addition, the industry collaboration facilitated through sandboxes can help regulators better understand key market issues in relation to AI technologies, and the costs and benefits of different regulatory responses. This is because sandboxes can help governments signal to the private sector their commitment to innovation while also improving engagement and communication ties between government and industry.[94] Given that sandboxes enable regulators to effectively "operate in start-up mode", they can also enable the rapid up/down scaling or abandonment of different AI policy options in response to their outcomes while facilitating experimentation with respect to balancing different regulatory concerns.[95] Sandboxes can help policymakers with "regulatory discovery", such as "evaluating the suitability of a legal framework" or improving decisions as to whether or how to amend legislation, enabling quicker development of effective AI regulation.[96] This time responsiveness is critical in relation to rapidly changing technologies like AI.

For industry, sandboxes are intended to encourage competition and innovation by opening market entry opportunities and by improving speed to market. They can achieve these goals by enabling waivers from specific legal provisions, streamlining the market authorisation process, reducing uncertainty, and improving market knowledge about the legal context around AI products and services. These considerations, in turn, support local research, development and deployment of AI products that embed trustworthiness and adherence to desirable standards throughout their lifecycle while also reducing barriers to finance, enabling the market expansion of responsible and ethical AI solutions and experimentation with different business models. [97] Consumers also benefit from this process as sandboxes facilitate the market entry of more innovative, ethical, and responsible AI solutions over time, in alignment with their needs and expectations.

There are precedents for the use of regulatory sandboxes for AI around the world, including in Australia, such as the "Enhanced regulatory sandbox" (ERS) administered by ASIC for the fintech industry.

ASIC's Enhanced Regulatory Sandbox

The ERS is made available to eligible entities to test innovative financial services or credit products. The ERS allows entities to test products without holding an Australian financial services or credit licence for a period of up to 24 months. Entities wishing to use this service must meet entry requirements and comply with ongoing conditions limiting the exposure of clients and must be accepted through an application process.

UK Information Commissioner's Regulatory Sandbox

The UK Information Commissioner's Regulatory Sandbox is designed to support organisations creating products and services that use personal data. The sandbox aims to facilitate responsible innovation by providing free access to the Information Commissioner's expertise in mitigating data risks and embedding "data protection by design" principles into product and service design. While the sandbox service supports firms of diverse types and sizes, and across multiple sectors, it is clearly bounded. The sandbox, for instance, does not help firms procure data, does not provide hosting services, and does not offer support in engaging with other regulators.[98]

Regulatory sandboxes carry inherent risks and may be resource-intensive to deliver. Therefore, they should be deployed strategically in areas where they would most usefully augment and support a novel regulatory response and priority industry.

Regulatory "safe harbours" are used as a tool to encourage adherence to standards and recognised best practices. They may incentivise firms to voluntarily adopt standards and assurance tools in return for limited, waived liability. In effect, safe harbours operate as a regulatory 'carrot', to provide protection to firms who comply thoroughly and in good faith with standards. Any protection provided would need to be strictly limited, and proportionate to the nature of the non-compliance. By way of example, government could choose to limit the liability of any firm that can demonstrate its compliance with auditing and assurance requirements in respect of preventing data misuse to the standards.[99] In order to be effective, it is important that the required standards are cheaper to implement than the cost of risking legal exposure, to ensure that it is in firms' rational economic interests to comply with the requirements.[100]

Benefits of regulatory safe harbours

Yaniskey-Ravid and Hallisey[101] articulate the potential benefits of this tool for AI data transparency, including:

- Rewarding desirable behaviour by encouraging those responsible for datasets to make diligent inquiries and meaningfully attempt violating any laws;
- Incentivising the internalisation of risks by firms, rather than leaving consumers to accept the costs;
- Maintaining penalties for intentional infringement or negligence; and
- Reducing harm to consumers overall.

Safe harbours serve a distinct purpose to regulatory sandboxes, and are likely to be substantially less resource-intensive to establish and administer. Unlike sandboxes, they do not offer substantial opportunities for government to test regulation and learn from collaborative experiences, but they may have the effect of substantially enhancing the safety of AI-enabled products and services for consumers. As with other measures described in this report, safe harbours could also be deployed as a limited response to particularly high-risk applications or industries.

Facilitating peer-to-peer learning

NSW Government can facilitate knowledge-transfer across the private sector by directly fostering engagement. While knowledge-transfer between government and individual firms typically occurs as part of procurement processes, there are opportunities to encourage knowledge sharing about ethical and responsible AI development more widely across the market. Research conducted by the Human Technology Institute at UTS indicates that company directors and executives in the Australian private sector believe that "examples of effective AI governance in peer organisations" would support them to implement effective AI Governance.[102] With the assistance of an active Emerging Technology Commissioner as proposed in this report, such engagement could aid and guide the market by deepening its understanding of ethical and responsible AI practices, appreciating their social significance and how they can help enhance AI innovation at the firm level. Although it is not the focus of this report, such an approach may also generate knowledge-transfer benefits among government agencies.

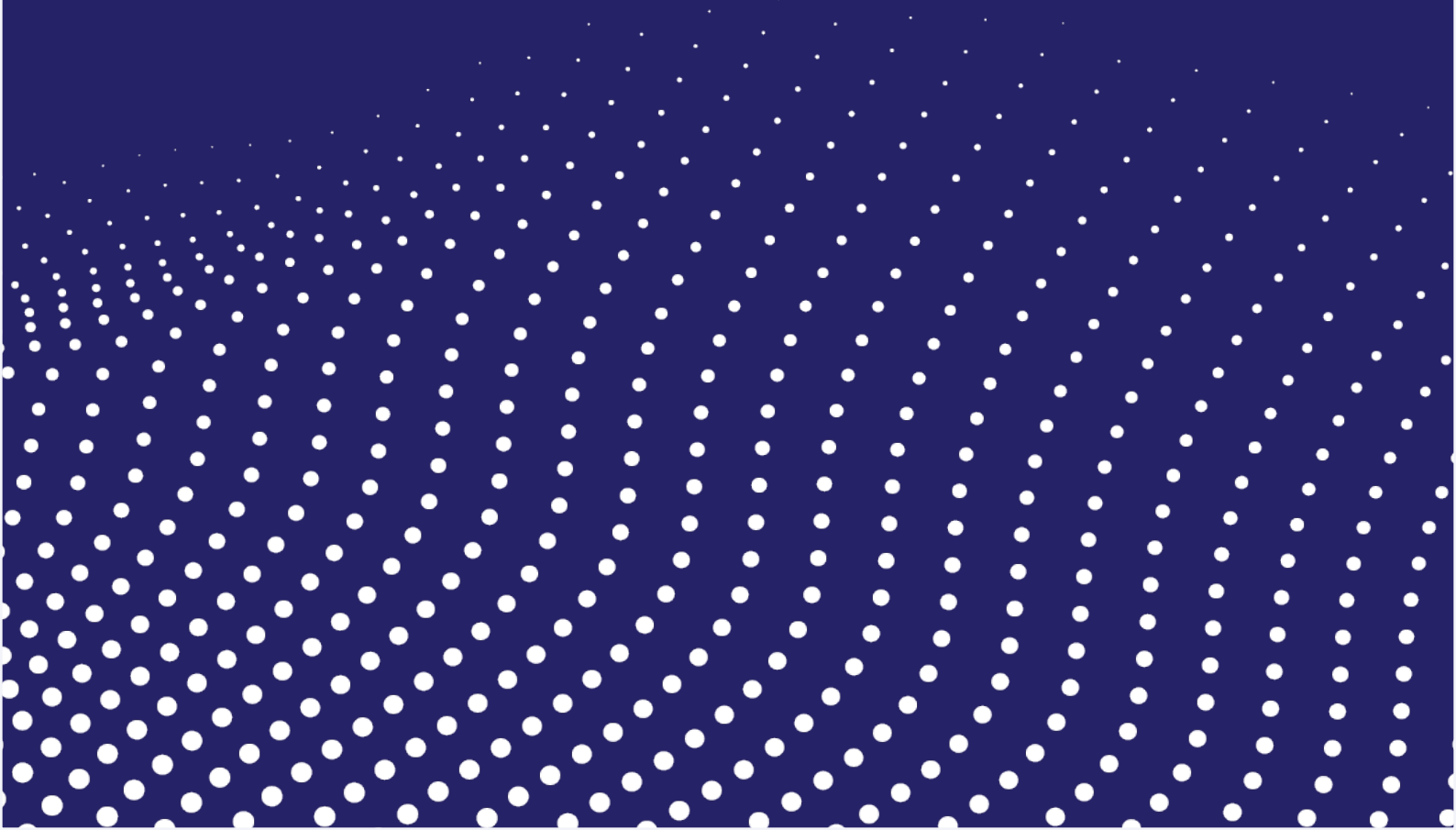
In addition, enabling peer-to-peer learning could support the NSW market by offering firms an opportunity to showcase their ethical and responsible AI solutions, and the lessons they have learned from incidents and "near misses" to other market actors. This is an example of a "demonstration effect", which could raise the profile of ethical and responsible AI solutions across the market. NSW Government could play a role in ensuring these initiatives circulate knowledge between companies of different sizes and business models, enabling exchange between large, established firms, small and medium-sized enterprises (SMEs) and start-ups.

Policy opportunity

10. **Facilitate peer learning and knowledge-transfer among firms to grow the private sector's understanding of ethical and responsible AI practices, and provide opportunities for firms to showcase their responsible practices.**

Supporting the public sector to respond effectively

This second part of the report focuses on the need for the NSW Government to craft institutional arrangements and develop capability that will allow it to more effectively respond to emerging technological changes, including AI. While such capability and arrangements will overlap with measures that improve the NSW Government's own use of AI, the focus here is on outward-facing measures to ultimately protect the rights and interests of people in NSW in the market and society more widely.



Embedding critical leadership capabilities in government's response

The following sections on critical leadership capabilities cover:

- The rationale for establishing an Emerging Technology Commissioner to centralise leadership of regulating and shaping emerging technologies such as AI;
- The specific objectives for which a Commissioner could be responsible; and
- The opportunity to embed Aboriginal and Torres Strait Islander leadership and cultural knowledge on responsible AI into the NSW Government's response.

An Emerging Technology Commissioner would strengthen government capacity

Establishing an Emerging Technology Commissioner in NSW could build on the leadership NSW has already demonstrated through the implementation of the AI Assurance Framework, to ensure that NSW remains a place where people benefit from a thriving emerging technology industry, and are appropriately protected from potential harms of AI technologies.

Agility and speed of response are crucial as the lag between technological developments and policy responses has represented a challenge for legislators and regulators around the world. A generalised remit of "emerging technology", rather than AI alone, is intended to create sufficient flexibility for the Commissioner and their office to respond flexibly to the fast pace of a range of interconnected technological developments. This flexibility also avoids the need for precise definitional clarity in a contested space, ensuring that benefits extend to those users who may be unaware of the nature of technology within products and services. While an Emerging Technology Commissioner might currently focus on AI, their role could adapt to address novel technological developments and challenges, such as implantable technology, autonomous systems, nanobots or quantum computing.

This report considers the role an Emerging Technology Commissioner might play in shaping initiatives and driving regulation of AI outside of government. Further work is required to carefully consider the relationship of this role to current AI oversight and assurance functions within the NSW Government, noting that these are critical and must continue to be strengthened. The NSW Government has made substantial progress on establishing and implementing oversight mechanisms to assure its own use of AI, through the AI Assurance Framework and AI Review Committee. It has also created valuable capabilities inside of government, including the leadership of the NSW Chief Data Scientist in coordinating this agenda.

Centralising leadership in this area is efficient and cost-effective

Government requires a cohesive approach to regulation and shaping

The permeation of AI technology throughout all aspects of life means that all parts of government, and particularly regulators, will require sufficient technical and sociotechnical capability and expertise to respond to emerging issues. A Commissioner could work cooperatively across the public sector to grow the capacity and capabilities of NSW's existing regulators and complaints bodies to respond to AI issues arising for members of the public.

Centralising this function would avoid the need to create new regulatory or complaints bodies, and would instead leverage and support existing entities to adapt to new challenges, without substantially increasing the resourcing required by individual agencies or regulators to meet these new demands. A centralised approach is especially efficient and cost-effective during this crucial period of time where sufficient capability is not yet embedded across government. NSW also requires leadership in developing a medium and long-term regulatory strategy for AI, as well as responding to urgent issues. This regulatory approach should be coordinated and consistent with international and national good practices.

Technology is moving fast and continuously evolving

The speed of change in AI and other technology is challenging for governments, where the pace of technological developments significantly outstrips legislative and regulatory responses. Moreover, subject matter expertise relating to AI may not be currently widely available in NSW government agencies, and these scarce resources should be leveraged and deployed in a strategic manner that maximises impact across the public sector. To best manage the dynamism of technological development, a strategic and integrated approach is necessary, particularly to avoid the risks and costs associated with duplication.

Similarly, subject matter expertise on the responsible use of AI may not be accessible to all firms across the mark despite the rapid uptake of such technologies. In their provision of information and general advice across the market, a Commissioner may meet a particular need for SMEs that may use AI in ways that pose risks to people. In this way, the provision of support to navigate these emerging technologies may be of real benefit to end users across NSW.

A Commissioner could be responsible for specific objectives

A Commissioner could have specific functions and responsibilities related to driving responsible practices in the development and use of AI throughout NSW.

Growing public sector capability

A Commissioner could work to build greater understanding, confidence and consistency across government and support peer-to-peer learning by public servants, and particularly those directly interacting with communities and industry, and those applying the AI Assurance Framework in procurement. A Commissioner could work collaboratively to build capacity and understanding within existing NSW government agencies, regulators and complaints bodies, and assist them in addressing complex legal and ethical issues.

A Commissioner could also oversee the development of strategic workforce development initiatives, such as those recommended throughout in this report. The efficacy of this approach has been demonstrated through the appointment of NSW's first Anti-slavery Commissioner, tasked with the goal of addressing modern slavery within government supply chains.

Coordinating and enhancing government's regulatory response

A Commissioner could be tasked with assisting and working cooperatively and collaboratively with other parts of government, rather than undertaking regulatory functions or responding directly to individual complaints themselves. This would include growing the capacity and capabilities of existing NSW agencies, regulators and complaints bodies to effectively and appropriately respond to AI.

The commissioner could act as a "digital clearinghouse" for the public's submission of complaints or inquiries that are rooted in issues with emerging technologies, supporting members of the public in their understanding of these issues, and allocating them to the appropriate regulatory and/or complaints entities in NSW and the Commonwealth. The Commissioner could also be empowered to cooperate with existing regulators and complaints bodies to monitor AI issues, including by gathering data for ongoing strategic planning or law reform.

The Commissioner could also be responsible for helping to educate and inform NSW residents about responsible AI and assisting to improve the public's perceptions of trust in AI. This could include encouraging the reporting of instances of concerning AI use to appropriate NSW government entities, such as police, the Anti-Discrimination Board, NSW Privacy Commissioner and other relevant complaints bodies.

Engaging with communities and industry

A Commissioner could play a key role in shaping understanding and trust in AI amongst the people of NSW. The role could leverage their technical expertise to engage with diverse stakeholders and lead participatory policymaking initiatives, as discussed in more detail below, in order to foster a better understanding of AI and improve perceptions of trust in its responsible use. A key manner of facilitating the continuous input of people with lived experience might be the inclusion of ongoing lived experience advisors within the Commissioner's office, representing communities with greater exposure to AI-related harms.

A Commissioner could also play a role in industry leadership to shape industry attitudes and practices towards responsible development and use of AI, including through a power to issue or endorse codes of practice. In this function, the Commissioner would benefit from the perception of sufficient influence in their appointment and role, as well as impactful reporting mechanisms, such as the capacity to report directly to the NSW Parliament. Such measures would enhance their capacity to engage with NSW markets and communities, as well as represent NSW interests nationally and globally.

Proposed features of an Emerging Technology Commissioner

Mandate to engage internally and externally

The Commissioner could have a mandate to engage internally across government, and externally beyond government, with the private sector, civil society, and the public.

Independent office holder

It is important that the Commissioner has independence and autonomy in their role. The Commissioner's impartiality could be safeguarded by secure funding from the NSW Government. Depending on the final scope of the role, it may be beneficial for the Commissioner to be established as an independent statutory officeholder with guaranteed tenure, reporting directly to the NSW Parliament.

Mandate to issue or endorse codes of practice

The Commissioner could be empowered to create or mandate standards, codes of practice, or guidelines, both independently and in partnership with other entities, and provide guidance in specific sectors or risk areas. With respect to regulation and shaping, this would enable the Commissioner to work with other government entities, as well as with external stakeholders to develop or endorse codes that are responsive to particular issues, sectors or risks.

Authority for information sharing and to request document production and information from different government entities

A Commissioner could be empowered to share information and request the generation and production of documents and information from other NSW government entities. Requesting information from other NSW government agencies, particularly complaints bodies and regulators, would enable the Commissioner to identify emerging trends and issues regarding AI. This information may be of specific relevance when reviewing existing legislative and regulatory regimes and could inform strategic planning by identifying potential areas for future reform.

Empowered to represent NSW in inter-jurisdictional forums

NSW has demonstrated its leadership through the introduction of the AI Assurance Framework, and has the opportunity to leverage this further to contribute to national discussions regarding responsible AI. NSW's representation and participation at interjurisdictional events is important because the significant buying power of the NSW government has the potential to influence broader practices via procurement processes, particularly in key areas including health, education and transport.

Holistic expertise in emerging technologies

The Commissioner's office should possess strong technical, sociotechnical, and governance expertise. In providing support to navigate AI-related issues across government and the market, the office must be capable of bridging the known gap between high-level principles and frameworks for responsible AI, and ensuring that these intentions are implemented in technical practice. For example, the Gradient Institute has identified approaches including impact assessment, data curation, contextualising fairness, pilot studies and organisational training, as critical features of translating a governance response into technical practice.[103]

Leadership of participatory policymaking in respect of AI

The Commissioner could lead engagement with civil society and groups with lived experience who are at greater risk of experiencing, or have experienced, AI-related harms across the whole of government regulatory response.

Policy opportunity

11. Create an independent Emerging Technology Commissioner to support and facilitate responsible AI design, development, and deployment across the market. The Commissioner's functions could include:
    - Engaging both internally, to build NSW Government's capacity, and externally, with the public and to provide general advice and information across the private sector;
    - Issuing and endorsing codes of practice;
    - Representing NSW in interjurisdictional forums.

The NSW Government should embed Aboriginal and Torres Strait Islander leadership and cultural knowledge on responsible AI systems

The NSW Government has an important opportunity to embed the perspectives and leadership of Aboriginal and Torres Strait Islander people in its future approach to AI ethics and regulation. A consortium of Australian Aboriginal and Torres Strait Islander academics, practitioners and experts in the AI space have worked together towards materialising a practical and tangible reflection of "Indigenous people's future dreaming of what AI could become, presented to the broader global AI community."[104] Through a process of developing an incubator bringing together a "diverse group of Australian Aboriginal and Torres Strait Islander researchers, professionals and practitioners", these experts have developed Indigenous protocols for AI.[105]

> "Indigenous leadership offers opportunities to govern technology developments through ancient practices of non-centralised authority, cooperative dynamics, complex knowledge systems and relational incentive structures. This promotes lawful behaviours that limit negative externalities, ensuring wellbeing not just for the team performing a task, but for all our relations, human and non-human, in the present and for generations into the future."[106]
>
> - Angie Abdilla, Megan Kelleher, Rick Shaw and Tyson Yunkaporta

A key outcome of this process has been the articulation of "an Aboriginal perspective to the architecture of AI systems, to data as a derivative of embodied knowledges and to cultural protocols which govern the intention, affect and effect of AI systems," reflecting "an underlying belief that in complex systems the 'meanings' or ethics of the system are not separable from the system itself."[107] Aboriginal and Torres Strait Islander leadership and Indigenous-led processes have also invented new methodologies that challenge and diverge from traditional Western approaches. An example of this is Country Centred Design, developed as an alternative to human-centred design, which privileges human needs over our ecosystem or environment.

Policy opportunity

12. **Embed the leadership of Aboriginal and Torres Strait Islander experts in NSW Government's response to AI on an ongoing basis, in staffing and advisory functions.**

Building capability to regulate and shape

Public servants across government will require new and adapted capabilities to respond to and shape the responsible use of AI beyond government, especially for staff who have direct engagement with the public or market. All parts of government will require a degree of technical and sociotechnical capability, including broad foundational capability, and must be able to access more detailed technical and sociotechnical expertise. This need for broad-based capability is recognised in the NSW AI Strategy, which notes that "there is an overall need to raise digital capability across the NSW public sector, including in the understanding of and use of AI."[108]

The strategy identified the following goals for public sector capability in respect of AI:[109]

- The public sector needs to know enough about AI solutions to make informed decisions on how to build, maintain and best use AI systems;
- Government needs strong competency to implement and manage AI in the longer term and collect and analyse AI-informed data. Once government has those competencies, it needs to create the right environment to retain that talent; and
- Government must have confidence in its ability to understand what the right AI solution is (or even if AI is the best solution) and how the technology works.

As an extension of digital uplift activities under the AI Strategy, the NSW Government should consider the additional and diverse capabilities required in the public sector to develop and deliver its regulatory response to AI beyond government.

> "The standards and contracting models for government AI and IT need to expand and establish a broader AI for public value innovation system, bringing together a wider cross-section of actors. However, the focus must be on internal talent and training improvement programmes, creating a long-term government learning and development process for generating new technical capabilities. Without this investment, the outsourcing and deployment of models may lead to slower service provision, underdeveloped public service offerings, and reduced efficiency due to mismatched expectations and working approaches."[110]
>
> - Mariana Mazzucato, et al.

The following sections building capability to regulate and shape cover:

- How NSW Government can embed participation of groups with lived experience of relevant harms in their ongoing regulatory approach.
- Features of capability required across NSW Government to respond to the challenge of regulating and shaping the development and use of AI across the market in NSW.

The NSW Government can lead the way on participatory policymaking

The NSW Government should establish systems to engage diverse groups of citizens in participatory policymaking processes concerning AI, especially those who belong to groups at greater risk of experiencing AI-related harms. These groups include, for example, those who have historically experienced various forms of discrimination and/or who are experiencing economic marginalisation. Government leaders should seek to employ innovative methods of engagement and facilitate holistic consideration of "what risks count, what harms matter, and to which values" responses to AI should be aligned in the design, implementation, and revision of policy.[111]

The participation of citizens in policymaking has a number of possible benefits.[112] These include the intrinsic benefit of treating citizens with respect as free and equal participants in a cooperative system of democratic self-rule, and allowing them, including those from at-risk groups, to more effectively protect their basic rights and advance their interests.[113] Participatory policymaking can also have epistemic and educative benefits that are of particular relevance with respect to AI.

- **Improved policy:** Participatory policymaking can bring together diverse groups of citizens to deliberate, shape policy, and provide feedback concerning their experiences of the harms and benefits of AI, the effects of existing policies, and ways to improve government responses to this emerging technology. Policymakers can take advantage of the insights of diverse groups with different lived experiences, and continuously revise policies with reference to their practical effects. This can allow for more effective and efficient problem-solving and policy design with respect to AI.[114]
- **Education and trust:** Participatory policymaking brings citizens into deliberative forums where they can work together with other citizens to solve problems and exercise control over the institutions and policies that govern their lives. This can allow people to develop a sense of shared civic purpose and engagement, as well as trust in the legitimacy and effectiveness of democratic institutions and policies.[115] This could be particularly valuable with respect to emerging technology such as AI, which has generated significant public concern about its potential consequences and its possible misuse.[116]

An Emerging Technology Commissioner would be well-placed to facilitate the inclusion of these voices in the policymaking process. It is important that this engagement occur continuously through an embedded mechanism. Options to facilitate this include:

- **Embedding lived experience in the staffing of relevant government organisations**: Lived experience advisors may be employed or engaged by the government on an ongoing basis, including in the office of the proposed Emerging Technology Commissioner, to better represent the views of particular groups within policy development and implementation.
- **Lived experience advisory functions:** There are a range of forms in which civil society or lived experience advisory bodies might be established. Lived experience advisory bodies offer a mechanism to test policy responses through peer engagement with people with lived experience on an ongoing basis to ensure authentic consultation and advice to decision-makers on how to meet the needs of diverse groups. A civil society advisory committee could also bring together civil society groups in areas such as human rights, consumer advocacy, and representatives of diverse and marginalised communities to provide advice and facilitate the taking into account of these diverse perspectives in any policy reforms.
- **Citizens' assembly:** A citizens' assembly or "mini-public" refers to a body constituted of a randomly selected, representative group of citizens who, often with access to expert advice, deliberate concerning a particular issue of public concern, and design non-binding policy recommendations for implementation.[117] The NSW Government could convene a permanent or recurrent Citizens' Assembly on AI, comprised of representative and randomly selected samples of the NSW population to meet, deliberate and provide recommendations with respect to AI Policy.

These bodies could advise the Commissioner on an ongoing basis, and should have the capacity to raise issues and direct the focus of their own inquiries. The NSW Government should also continue to utilise existing avenues for citizen participation in policymaking, such as consultation processes with respect to legislative reforms, including through the online "Have Your Say" tool,[118] and remain open to other forms of citizen engagement with AI policy.

The **National Disability Data Asset** (NDDA) is an Australian initiative bringing together de-identified information from various government agencies about Australians with and without disability. Research conducted by the Sydney Policy Lab identified the key role of lived experience leadership and management of the asset:

"Disability-led processes and systems of governance were seen as the best way to prevent misinterpretation and misuse and to ensure data is used to benefit the community in the short and long term. This includes people with disability playing leading roles in the NDDA's governance bodies to shape how it develops, and in understanding and interpreting the data and its applications."[119]

As an enactment of this principle, the NDDA enacted a Disability Advisory Council for the asset pilot, who engaged with people with disability, their carers, family members and the broader disability community throughout the pilot, and provided guidance and advice on how to design and operate the asset.[120] Following the pilot, three of the 12 members of the ongoing NDDA Council will be members of the disability community.

The **Canadian Citizens' Assembly on Democratic Expression**, an initiative by the Public Policy Forum (a Canadian non-governmental organisation), provides one example of participatory policymaking in response to emerging technologies. This representative citizens' assembly, which met for the first time in 2020, engaged in deliberation concerning policy responses to online harms and ways to buttress the public good. The Assembly has released reports in January 2021 and January 2022, which set out policy recommendations for regulating online platforms[121] and responding to the spread of online disinformation.[122]

Policy opportunity

13. Establish ongoing advisory mechanisms to engage people with diverse lived experience to advise government on how to best protect the rights and interests of people in NSW in respect of AI and emerging technologies. This could include:
    - Embedding lived experience in relevant government organisations, including the proposed Office of the Emerging Technology Commissioner;
    - Establishing a civil society advisory committee; and
    - Convening a citizens' assembly on AI.

A systematic effort is needed to build public sector capability

Capability is needed across diverse domains of engagement

The NSW Government has a wide range of touchpoints with citizens, businesses and the non-government sector. Ensuring that there is sufficient AI capability across relevant agencies is fundamental to ensure that agencies fulfil their varied obligations to protect the rights and interests of people in NSW.

Key points of engagement include:

- **Regulators and complaints bodies:** Regulators and complaints bodies such as Fair Trading NSW, Anti-Discrimination NSW, and the NSW Ombudsman have a direct role in taking complaints, resolving disputes, and enforcing regulatory schemes. These bodies also provide information to the public about their rights and obligations, such as fact sheets published by Anti-Discrimination NSW.

Therefore, these agencies have a vital role to play both as conduits for people to challenge decisions that adversely affect their rights or interests and in educating the public as to their rights and potential violations thereof. Both of these roles will be vital, and will be transformed, by the emergence of AI.

- **Commissioners interacting with different communities:** Existing Commissioners and their offices interface with particular communities and particular industries, such as the Small Business Commissioner and the Modern Manufacturing Commissioner. These Commissioners have diverse roles, including developing and implementing policy strategies, advocating for particular communities, and fostering communication and collaboration within and between communities, industries and government. These bodies will also play an important role in shaping the response of the private sector in their particular areas to the emergence of AI.
- **Government bodies procuring AI technologies:** Beyond regulatory and complaints bodies and Commissioners, any government entity that procures AI technologies or systems will also play an important role in shaping the development of AI in NSW.

Capability uplift must encompass technical, sociotechnical and governance expertise

To facilitate a sound regulatory response, capability will be required at the intersection of AI and particular subjects (such as healthcare and education) and at the intersection of particular capabilities (such as evaluation and auditing).[123] Public sector agencies must also encompass technical, sociotechnical, and legal and governance expertise in order to respond effectively to the emergence of AI.

- **Technical expertise:** AI systems can be technically complex. In order to effectively regulate and shape the adoption of AI in NSW, it is of fundamental importance that public sector actors have a sufficient understanding of the complex technology underpinning AI systems.[124]
- **Sociotechnical expertise:** Beyond their technical complexity, AI systems are inherently "sociotechnical" in nature, meaning that they are influenced by societal dynamics and human behaviour. Moreover, AI has effects that "can emerge from the interplay of technical aspects combined with societal factors related to how a system is used, its interactions with other AI systems, who operates it, and the social context in which it is deployed."[125] It is crucial to embed an understanding of these complex interactions between the technical, human, organisational and social aspects of AI systems within the NSW public sector.
- **Legal and governance expertise:** It is also important that there are sufficient resources and tools for the NSW public sector to effectively understand the implications of AI for existing legal and regulatory schemes, the role of relevant governance and ethics principles, and the operation of any new or amended legal responses to the emergence of AI.

Capability development efforts should respond directly to areas of need

Building a taxonomy of AI regulatory capabilities

Beyond the foundational capabilities already progressing across the NSW public sector, different agencies and roles within the NSW public sector will require differing skills and degrees of AI capability.

As UNESCO notes in its framework of digital competencies for public servants with respect to AI and transformation, AI competency can be understood as ranging from basic to advanced:[126]

1. **Basic:** Broad understanding and knowledge of a subject and theme and the ability to carry out certain basic tasks related to the subject. For example, a person with a basic understanding of AI might be aware of relevant AI hiring systems and their potential implications for employment discrimination.[127]
2. **Intermediate:** Good understanding of a subject and theme and the ability to carry out more advanced tasks related to the subject. Capability to deal with and provide guidance to others on different tasks related to the subject.
   Building on the above example, a person with an intermediate understanding of AI fundamentals would have a more nuanced understanding of AI hiring systems. They would be able to effectively respond to and propose solutions in circumstances where such an AI system has generated discriminatory outcomes.[128]

3. **Advanced:** Advanced understanding and knowledge of a subject and theme. Demonstration of applied approaches, tools and methods related to the subject and ability to coach other people. Also, the ability to embed the specific skill and related practices across the organisation and to coach others to do so.
With respect to a potentially discriminatory AI hiring system, a person with an advanced understanding (such as a data scientist), would be able to interpret, evaluate and understand the decisions made during the development and design of the relevant hiring algorithm. This person would then be able to identify ways in which flaws leading to discrimination might be remedied. They would also be able to recruit and assign additional people and skillsets to AI projects, such as the development of detailed technical guidance to prevent algorithmic discrimination.[129]

In this report, we have considered the way in which the public sector workforce will be required to respond in a public-facing regulatory strategy. This is just one part of a complex web of workforce adaptation initiatives that will be required to respond to AI. Other dimensions of this challenge include government's internal capacity to develop and deploy AI responsibly, and the challenge of building Australia's AI workforce and broader industry in a coordinated, strategic way to maintain international competitiveness.

Designing and implementing a regulatory workforce development plan

With regard to the broader needs of the NSW Government and the national context, the NSW Government should design and implement a sector-wide AI regulatory capability framework to maximise the efficiency and effectiveness of its regulatory and market-shaping response to AI. This should complement any existing or proposed work previously completed under the AI Strategy, or currently underway to develop NSW Government's capability in using AI within government.

In designing this framework, the NSW Government should carefully consider:

1. Where expertise is already available across government;
2. What kind of capability employees need to perform their function and within their organisational context;
3. Which areas and institutions should be prioritised for capability building and resourcing, considering the urgency of developing AI capabilities and the seriousness of the consequences of a failure to build these capabilities in different agencies and functions;
4. How skills and expertise might be distributed across government to ensure the most effective utilisation of its AI capabilities; and
5. What methods of capability acquisition should be prioritised in different contexts.

The NSW Government should work to embed universal capabilities within the public sector, while also targeting resources and capacity-building in those areas and institutions that are likely to be on the frontline of supporting the public in responding to emerging challenges from AI. The approach to developing public sector capability must be progressive, sustainable, and ongoing to allow rapid adaptation and flexibility to respond to these fast-changing technologies.

Policy opportunities

14. **Develop a map of AI regulatory capabilities required within the public sector.**
15. **Develop and implement a strategic plan for the acquisition and development of the AI regulatory capabilities identified through mapping.**

About this report

Research approach

This report is the result of a James Martin Institute for Public Policy Collaborative Project. This model brings together government, academia, and other experts and stakeholders to work cooperatively on challenging policy issues. The core project team worked in close collaboration with the NSW Government's Chief Data Scientist within the Department of Customer Service, and was advised by an Expert Advisory Group of which the Chief Data Scientist was also a member. The team worked collaboratively throughout a three-month period in mid-2023. It conducted consultations with more than 20 individuals, analysed the evidence collected, and developed policy options for the NSW Government. Leading AI researchers and experts from universities and civil society, policymakers and stakeholders were consulted throughout. While the Institute managed the process around the project and took leadership over the report's final design, its content is a product of genuine collaboration between those involved.

Utilising the Institute's applied policy research approach, this project's primary information sources are:

- **Desktop research** drawing together academic sources, "grey" literature and other public information.
- **Consultations** with expert advisory group members, researchers and experts on AI from academia and civil society, policymakers and stakeholders.
- **Industry discussion** held on 19 July 2023 with representatives of relevant industry organisations.
- The scope of this project meant that it was not possible to conduct a comprehensive survey of the risks, opportunities and challenges for NSW arising from the emergence of AI. Nor was it within its scope to meaningfully consider and engage with every existing program, initiative or other piece of work that might already be occurring in NSW – whether led by government, the community, or others.

The primary audience for this report is policymakers within the NSW Government, with the insights and policy pathways calibrated to focus on system-level reform. At the same time, this report is also intended to be a useful resource for other state and territory governments, the Commonwealth Government, as well as the broader policy community, community leaders, researchers, and other stakeholders.

Contributors

Expert Advisory Group (EAG)

- Professor Lyria Bennett Moses (Director, UNSW Allens Hub for Technology, Law and Innovation; Professor, UNSW Law & Justice)
- Professor Nicole Gillespie (Lead, Trust Ethics and Governance Alliance Research and KPMG Chair in Trust, University of Queensland)
- Professor Heather Horst FAHA (Professor and Director, Institute Culture and Society, Western Sydney University and Chief Investigator, Centre of Excellence for Automated Decision-Making and Society)
- Professor Peter Leonard (Director, Data Synergies, Professor of Practice, UNSW Sydney Business School and Member, NSW AI Review Committee)
- Dr Ian Oppermann (Chief Data Scientist, NSW Government, and Industry Professor, UTS)
- Professor Edward Santow (Director, Policy and Governance, Human Technology Institute, Industry Professor, Responsible Technology, UTS, and Member, NSW AI Review Committee)
- Myfanwy Wallwork (Fellow, Human Technology Institute, UTS)
- Professor Kimberlee Weatherall (Professor of Law, University of Sydney and Chief Investigator, ARC Centre of Excellence for Automated Decision-Making and Society)

Consultations

JMI conducted consultations with a wide range of experts, including:

- Academic researchers
- Experts from civil society groups
- Industry stakeholders
- Government policy staff
- Professional practitioners, including lawyers and actuaries

Explanatory statement on authorship

The James Martin Institute for Public Policy (JMI) is a nonpartisan, independent policy institute that does not adopt an institutional view on specific policy issues. This report reflects the calibrated view of the project team, which operated under a highly collaborative model. Its view was formed on the basis of an assessment of relevant academic research, stakeholder consultations, and engagement with relevant experts, including an expert advisory group (EAG).

The findings and recommendations of any JMI publication do not necessarily reflect the views of the Institute, its Board, funders, advisers, or other partners.

Appendices

Appendix A | Case study on AI and the Anti-Discrimination Act 1977 (NSW)

NSW has existing jurisdiction in respect of preventing and prohibiting discrimination, through the *Anti-Discrimination Act 1977* (NSW). Discrimination and bias are understood as key risks arising from the use of algorithmic decision-making in different aspects of life. The NSW Government might explore opportunities to leverage and reform the existing policy and institutional settings designed to respond to discrimination. Given the clear potential for this to have a demonstrable public benefit, this case study serves as a compelling example of how existing legal and regulatory frameworks in NSW can be leveraged and potentially adapted to respond to new risks created by AI systems.

What is unlawful discrimination?

In NSW, the *Anti-Discrimination Act 1977* (NSW) prohibits discrimination in work, education, the provision of goods and services, accommodation, and in registered clubs on the basis of the following protected attributes:

- Race
- Sex
- Transgender identity
- Sexual orientation
- Marital or domestic status
- Disability
- Responsibilities as a carer
- Age

The NSW Act prohibits both direct and indirect discrimination.[130] Direct discrimination occurs when a person is treated less favourably on the ground of a protected attribute than a person not possessing that attribute would have been. For example, if a company does not offer a woman employment when it would have offered a man the same job in the same or materially similar circumstances, that would constitute direct discrimination.[131] The characteristics extension extends direct discrimination by making it unlawful to discriminate against a person because of a characteristic that is generally possessed by, or imputed to, people who have the relevant protected attribute. This is intended to ensure that these characteristics are not used as "proxies" for discriminating on the grounds covered in the legislation.[132] Indirect discrimination is when a rule or policy applies to everyone but has the effect of disadvantaging some people because of a personal characteristic that is possessed or generally imputed to persons possessing a protected attribute.

Australia's federal anti-discrimination laws also prohibit discrimination on the basis of similar protected attributes.[133] There are some differences in scope and application between federal and NSW anti-discrimination laws.[134] The federal anti-discrimination law is comprised of six relevant acts that address different protected attributes.[135] The areas of public life covered by the federal legislation are also more expansive than those covered in NSW but are inconsistent across the various acts.[136] While NSW should act as a leader and address the emerging risks of algorithmic discrimination under NSW anti-discrimination law, a comprehensive response will require coordinated state, territory and federal reforms.

Discrimination and bias can arise in AI-powered systems

The use of AI technologies can result in unfair or discriminatory decisions. Discrimination and unfair bias can affect all forms of decision-making and are not unique to AI-powered systems. In certain circumstances, the use of AI can even mitigate the risk of discrimination that flows from the biases of human decision-makers and enable better quality, data-driven decision-making.[137] However, when algorithmic discrimination does occur, it can obscure and entrench unfair bias and discrimination and contribute to the deepening of current societal inequalities.[138] Both direct and indirect discrimination can occur when AI-powered systems are used to make or inform decisions.

This is often referred to as the risk of "algorithmic bias" or "algorithmic discrimination" and can arise from problems with the data being used by an AI-powered system, or from problems within the AI system itself.[139] These terms are used generally to mean situations where AI is used to produce outputs that treat one group less favourably than another, without justification.[140] This can occur in two key ways:

- **Flawed AI Model:** Discrimination can arise from the design, implementation and use of an AI system. This can be the result of an algorithm being expressly designed to exclude a particular group or to give extra weight to a protected attribute. It can also arise as the result of "aggregation bias" when an AI model, that is reasonable when applied to one population, is inappropriately applied to another population for which it does not make sense. For instance, the use of credit scores to underwrite loans is reasonable when applied to the general population. However, when applied to younger populations, it can lead to erroneous outcomes.
- **Flawed data:** When the data used to train an AI system is flawed, it can produce results that are unreliable, unfair, or discriminatory. This is often referred to as the "garbage in-garbage out" problem in data science: if the data inputs to an AI system are discriminatory or biased, then its outputs will also be discriminatory. This can perpetuate historical inequalities and injustices against persons on the basis of their race, gender, and other protected attributes.[141]

There are many examples of algorithmic discrimination that have emerged with respect to areas protected under anti-discrimination law, including employment, accommodation, and provision of essential goods and services:

- **Rental screening:** AI tools used to screen tenants applying for rental properties in the US have allegedly discriminated against racial and ethnic minorities by considering matters such as credit history and non-tenancy debt, failing to account for governmental assistance for housing and taken other factors into account in such a way that disproportionately denies applicants from historically marginalised backgrounds.[142]
- **Employment:** An AI-powered recruitment tool was trained to identify patterns in job applications received by the company over a ten-year period. Most of the job applicants in that period were male. The system thus "learned" that male applicants were preferable and favoured male over female candidates in its generation of recommendations for the future workforce.[143]
- **Insurance underwriting:** Insurance underwriting is an area with significant scope for algorithmic discrimination. The AHRC and Actuaries Institute, in a guidance document for avoiding algorithmic discrimination in this context, have provided valuable hypothetical case studies of the risks arising with respect to car, travel, and life insurance.[144]An example of this problem was highlighted recently by a piece of research conducted by Citizens Advice (UK), which identified an "ethnicity penalty" in the UK car insurance market. They observed a correlation between insurance costs and ethnicity, with people of colour paying significantly more for their car insurance, even after controlling for relevant demographic and geographical factors.[145] Citizens Advice expressed concern that opacity in pricing algorithms might be reflecting or perpetuating wider inequalities and human biases in light of the unusual correlations they identified.[146]
- **Medical treatment:** Healthcare tools using AI can produce worse outcomes for persons possessing particular attributes. For example, an algorithm used to distinguish malignant and benign moles that is trained on fair-skinned patients might fail to properly diagnose moles on people of colour,[147] or an algorithm deployed to detect cardiovascular diseases might underperform on women because most of the medical training data concerns men.[148] Further, an algorithm designed to identify patients with high needs for healthcare in the US systematically assigned lower scores (indicating less need) to black patients than to white patients, even when those patients had similar numbers of chronic conditions and other markers of health.[149]

The NSW Government can adapt existing levers to respond to the challenge of algorithmic discrimination

The NSW Government is well-positioned to respond quickly to the risks arising from algorithmic bias. Government can consider proactive interventions, such as regulating the private sector through audit or reporting requirements, and remedial interventions, including targeted reforms to the *Anti-Discrimination Act 1977* (NSW) to improve the protections and remedies available to people in NSW where discrimination does occur.

Reforms could ensure that individuals and relevant authorities are aware that AI has been used and are able to obtain information concerning its use in decision-making and the provision of products and services. In this context, transparency refers to disclosure that AI is being used, and "enabling people to understand how an AI system is developed, trained, operates, and deployed in the relevant application domain, so that consumers, for example, can make more informed choices."[150] A provision could be inserted requiring those using AI to make decisions in the areas covered under the Act to notify those affected by these decisions that AI is being used to support consumers to make informed choices and raise well-informed disputes or complaints if a person suspects a decision has been discriminatory.

Anti-Discrimination NSW could be given the power to obtain information and materials from decision-makers when AI is being used. Under section 90B of the *Anti-Discrimination Act 1977* (NSW), the President of the Anti-Discrimination Board may, by notice in writing, require a complainant or person against whom a complaint is made to supply information or documents. That person must provide the information or documents unless they have a reasonable excuse. This section should be amended to avoid any doubt about its applicability when AI is used, including an exploration of whether the power to compel the creation of documents that explain reasons produced by AI is an appropriate inclusion.

The Act should allow for clear allocation of legal responsibility for decisions informed by AI systems

The Act could usefully be amended to ensure the clear allocation of legal responsibility for AI-informed discriminatory decisions. This could provide a model for other areas where AI might undermine existing tests for legal liability.

Allocating responsibility to a person

The *Anti-Discrimination Act 1977* (NSW) presently applies to the conduct of a "person (the perpetrator)" who must be proven to have engaged in either direct or indirect discrimination. An AI system that makes decisions is not a "person". In some circumstances, there may be no person involved in a particular decision made using an AI system at all. Unlike corporations, there is no provision in the Act that deems an AI system to be a person for the purposes of the Act.

The *Anti-Discrimination Act 1977* (NSW) could usefully be amended to avoid ambiguity in allocating legal liability for AI-powered decisions to the legal person who instigated its use.

Causation

In order to establish direct discrimination, a person must prove that they have been treated less favourably "on the ground of" a protected attribute. This is often referred to as the "causation" component of proving direct discrimination. This requires a court or tribunal to undertake an objective assessment of the respondent's reasons for treating a person less favourably in order to determine whether one of the "grounds" of this differential treatment was a protected attribute.[151] Under the NSW Act, if something is done for multiple reasons, and one of those reasons amounts to of unlawful discrimination (whether or not it is the dominant or a substantial reason), it is taken to have been done for that discriminatory reason.[152]

The use of AI could generate significant difficulties for persons trying to prove causation with respect to discriminatory conduct because of:

1. **Difficulties in determining which factors are considered by an AI-powered system:** It can be very difficult to understand which of the multitude of factors contained in a particular dataset are the reasons or "grounds" for decisions of an AI-powered system. If a court or tribunal is unable to identify the grounds of a decision, it will not be able to find that one of those grounds was discriminatory.

2. **Difficulties in attributing a state of mind to a human being:** Causation requires an objective assessment of the state of mind of the respondent, and whether they engaged in differential treatment of a person "on the ground of" a protected attribute. When AI is used, however, it can be very difficult to attribute knowledge of algorithmic bias to a human perpetrator. The requisite state of mind might, therefore, become even more difficult to establish.[153]

Amendments to the *Anti-Discrimination Act 1977* (NSW) should account for AI-related risks, including:

- Requiring that those deploying AI systems notify affected parties that AI is being used.
- Ensuring that the use of AI or other technologies does not allow persons to avoid complying with the obligation to provide information or materials to the Anti-Discrimination Board under section 90B.
- Clarifying that a person who uses an automated system or tool to make a decision or carry out a function is taken for the purposes of the Act to have made that decision or carried out that function.

Appendix B | Examples of guidelines in the federal and international context

1. **AHRC and Actuaries Institute, Guidance Resource concerning AI and discrimination in insurance and underwriting (2022):** In 2021, the AHRC recommended that the Federal Government should resource the AHRC to produce guidelines for government and non-governmental bodies concerning how to comply with federal anti-discrimination laws in the use of AI-informed decision-making. In December 2022, the Commission and the Actuaries Institute of Australia published a guidance resource concerning AI and discrimination in insurance and underwriting.[154] The Guidance Resource provides:
    - Information about the operation of the federal discrimination laws in relation to the use of AI in insurance pricing and underwriting decisions;
    - Information about the risks of algorithmic bias and discrimination arising from the use of AI; and
    - Practical guidance for insurers for avoiding unlawful discrimination when using AI.
2. **AHRC Algorithmic Bias Technical Paper (2020):** In 2020, the AHRC issued a technical paper titled *Using artificial intelligence to make decisions*: *Addressing the problem of algorithmic bias* in collaboration with the Gradient Institute, Consumer Policy Research Centre, CHOICE and CSIRO's Data61.[155] This paper uses simulations to demonstrate how the risk of algorithmic bias can arise in a decision-making process, how it can be identified, and steps that can be taken to address or mitigate the problem. It provides a valuable model for any Anti-Discrimination NSW codes of practice that might provide practical examples and mitigation measures for those employing AI in decision-making.
3. **CSIRO and Data61 Responsible AI Pattern Catalogue:** CSIRO and Data61 are developing a "Responsible AI Pattern Catalogue", comprised of reusable solutions to commonly occurring problems within a given software development context, for the purposes of operationalising responsible AI from a system perspective.[156]
4. **NSW Bar Association Issues Document:** The NSW Bar Association released a guidance document on 12 July 2023 entitled: "Issues Arising from the Use of AI Language Models (including ChatGPT) in Legal Practice".[157] This document guides and assists NSW barristers to understand their duties under the *Legal Profession Uniform Conduct (Barristers) Rules 2015* (NSW) with respect to the use of AI language models in legal practice. This guidance explores some of the existing professional obligations, identifies some of the risks arising from the use of AI with respect to these obligations, and sets out actions for barristers to take in order to avoid breaching the Barristers Rules. The guidelines also note that they "may be considered by the Bar Council in relation to any complaint received regarding a barrister's use of AI language tools for legal research, advice, and analysis." Finally, the Bar Association observes the speed of developments in the field of AI and notes that: "Barristers retain the responsibility of ensuring that any tool they use to aid them in practice is appropriate for such use and for ensuring at all times adherence to the Barristers Rules."
5. **Singapore's Examples-Based Guidance:** Singapore has issued examples-based guidance for the private sector in the form of a two-volume *Compendium of Use Cases*. This guidance resource complements Singapore's Model AI Governance Framework that converts AI ethics principles into implementable measures for organisations, and its Implementation and Self-Assessment Guide for organisations by showcasing the successful implementation of measures in organisations of different sizes, and in different sectors both in Singapore and internationally.[158] An expert consulted by JMI referred to the Singaporean example of providing a useful example of how AI ethics principles can be converted from "theory to practice".
6. **US Equal Employment Opportunity Commission Guidance on the Americans with Disabilities Act and AI:** The American Equal Employment Opportunity Commission has issued a non-binding guidance that sets out how existing disability discrimination law might apply to the use of AI in employment-related decision-making. It provides practical advice for employers on how to avoid disability discrimination in the use of AI, as well as advice for employees and job applicants whose rights have been violated.[159]

7. **UK Information Commissioner's Office Guidance on AI and Data Protection:** The UK Information Commissioner's Office has issued detailed guidance concerning AI and Data Protection. This guidance is legally non-binding but contains advice on how to interpret relevant data protection law as it applies to AI, and recommendations for good organisational practice for organisational and technical measures to mitigate the risks of AI. While the focus is on data protection law, it also provides advice concerning algorithmic bias with respect to fairness provisions contained in this law.[160]

Endnotes

1 OECD Council on Artificial Intelligence, *Recommendation of the Council on Artificial Intelligence* (Paris: OECD, 2023), https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449.

2 Ibid; Australian Government, Safe and Responsible AI in Australia: Discussion Paper (Canberra: Department of Industry, Science and Resources, 2023), 7-9, https://storage.googleapis.com/converlens-au-industry/industry/p/prj2452c8e24d7a400c72429/public_assets/Safe-and-responsible-AI-in-Australia-discussion-paper.pdf; United States Government, Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People (Washington DC: The White House, 2022), 3, https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf.

3 Laura Liebig, Licinia Güttel, Anna Jobin and Christian Katzenbach, "Subnational AI policy: shaping AI in a multi-level governance system", *AI & Society* (2022), 1-14.

4 Lyria Bennett Moses, "Regulating in the Face of Sociotechnical Change" *The Oxford Handbook of Law, Regulation and Technology,* Roger Brownsword, Eloise Scotford and Karen Yeung (eds.) (Oxford: Oxford University Press, 2016), 573–596.

5 Kimberlee Weatherall et al., *ADM+S submission to the Safe and responsible AI in Australia discussion paper* (Sydney: ARC Centre of Excellence for Automated Decision-Making and Society, 2023), 12, https://www.admscentre.org.au/submissions/.

6 Australian Government, *Safe and Responsible AI in Australia: Discussion Paper* (Canberra: Department of Industry, Science and Resources, 2023), 8, https://storage.googleapis.com/converlens-au-industry/industry/p/prj2452c8e24d7a400c72429/public_assets/Safe-and-responsible-AI-in-Australia-discussion-paper.pdf.

7 Ian Opperman, "The AI genie is out of the bottle – here's how to regulate it", *The Policymaker*, 15 August 2023, https://thepolicymaker.jmi.org.au/the-ai-genie-is-out-of-the-bottle-heres-how-to-regulate-it.

8 United Nations Office of the Secretary-General's Envoy on Technology, "High-Level Advisory Body on Artificial Intelligence", https://www.un.org/techenvoy/ai-advisory-body (accessed 6 December 2023).

9 OECD Council on Artificial Intelligence, *Recommendation of the Council on Artificial Intelligence* (Paris: OECD, 2023), https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449.

10 United States Government, 2022 *Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People* (Washington, DC: The White House, 2022), https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf.

11 European Parliament, "MEPs ready to negotiate first-ever rules for safe and transparent AI", *European Parliament News,* 14 June 2023, https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai.

12 International Organization for Standardization, "ISO/IEC 22989:2022 – Information technology – Artificial intelligence — Artificial intelligence concepts and terminology" (Geneva: International Organization for Standardization, 2023), https://www.iso.org/standard/74296.html.

13 Australian Government, *Safe and Responsible AI in Australia: Discussion Paper* (Canberra: Department of Industry, Science and Resources, 2023), 7-9, https://storage.googleapis.com/converlens-au-industry/industry/p/prj2452c8e24d7a400c72429/public_assets/Safe-and-responsible-AI-in-Australia-discussion-paper.pdf.

14 United States Government, *Fact Sheet: President Biden Issues Executive Order on Safe, Secure and Trustworthy Artificial Intelligence* (Washington DC: The White House, 2023), https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/.

15 Ibid.

16 Laura Liebig, Licinia Güttel, Anna Jobin and Christian Katzenbach, "Subnational AI policy: shaping AI in a multi-level governance system", *AI & Society* (2022), 1-14.

17 Stanford University, 2023, *Artificial Intelligence Index Report* (Palo Alto: Stanford University, 2023), 267, https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf.

18 OECD Council on Artificial Intelligence, *Recommendation of the Council on Artificial Intelligence* (Paris: OECD, 2023), https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449.

19 Australian Government, *Safe and Responsible AI in Australia: Discussion Paper* (Canberra: Department of Industry, Science and Resources, 2023), 7-9, https://storage.googleapis.com/converlens-au-industry/industry/p/prj2452c8e24d7a400c72429/public_assets/Safe-and-responsible-AI-in-Australia-discussion-paper.pdf.

20 Genevieve Bell, Jean Burgess, Julian Thomas, and Shazia Sadiq, *Rapid Response Information Report: Generative AI - language models (LLMs) and multimodal foundation models (MFMs)* (Canberra: Australian Council of Learned Academies, 2023), 12, https://www.chiefscientist.gov.au/sites/default/files/2023-06/Rapid%20Response%20Information%20Report%20-%20Generative%20AI%20v1_1.pdf.

21 Jenna Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms", *Big Data & Society*, 3:1 (2016), https://journals.sagepub.com/doi/full/10.1177/2053951715622512.

22 Lauren Solomon and Nicholas Davis, *The State of AI Governance in Australia* (Sydney: Human Technology Institute, University of Technology Sydney, 2023), https://www.uts.edu.au/sites/default/files/2023-05/HTI%20The%20State%20of%20AI%20Governance%20in%20Australia%20-%2031%20May%202023.pdf.

23 Nicholas Davis, Sophie Farthing and Edward Santow, *Discussion Paper, Safe and responsible AI in Australia Submission to Department of Industry, Science and Resources,* (Sydney: Human Technology Institute, University of Technology Sydney, 2023), 2, https://www.uts.edu.au/sites/default/files/2023-08/23.08.09%20HTI%20Submission%20to%20DISR%20Safe%20and%20Responsible%20AI%20in%20Australia%20DP.pdf.

24 Nicole Gillespie, Stephen Lockey, Caitlin Curtis, Javad Pool and Ali Akbari, *Trust in Artificial Intelligence: A Global Study* (Brisbane: The University of Queensland and KPMG Australia, 2023), 3, https://ai.uq.edu.au/files/6161/Trust in AI Global Report_WEB.pdf.

25 Ibid, 28.

26 Ibid, 4-5.

27 Sophie Farthing, John Howell, Katerina Lecchi, Zoe Paleologos, Phoebe Saintilan and Edward Santow, 2021 *Human Rights and Technology Final Report* (Sydney: Australian Human Rights Commission, 2021), 79, https://humanrights.gov.au/our-work/technology-and-human-rights/publications/final-report-human-rights-and-technology; See also: Karen Yeung, *A Study of the Implications of Advanced Digital Technologies (Including AI Systems) for the Concept of Responsibility Within a Human Rights Framework* (Strasbourg: The Council of Europe, 2019), https://rm.coe.int/a-study-of-the-implications-of-advanced-digital-technologies-including/168096bdab; The Commission to the European Parliament, The Council and The European Economic and Social Committee, *Report on the Safety and Liability Implications of Artificial Intelligence, the Internet of Things and Robotics*, COM/2020/64, (Brussels: The European Commission, 2020), https://eur-lex.europa.eu/legal-content/en/TXT/?qid=1593079180383&uri=CELEX%3A52020DC0064.

28 Sophie Farthing, John Howell, Katerina Lecchi, Zoe Paleologos, Phoebe Saintilan and Edward Santow, 2021 Human Rights and Technology Final Report (Sydney: Australian Human Rights Commission, 2021), 79.

29 Ibid, 77-78.

30 European Commission, *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative Acts* (Brussels: European Commission, 2021), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206.

31 *Automated Employment Decision Tools*, New York City Council Local Law 144 of 2021 (2021), https://www.nyc.gov/site/dca/about/automated-employment-decision-tools.page.

32 Sophie Farthing, John Howell, Katerina Lecchi, Zoe Paleologos, Phoebe Saintilan and Edward Santow, *Human Rights and Technology Final Report* (Sydney: Australian Human Rights Commission, 2021), 80-83, https://humanrights.gov.au/our-work/technology-and-human-rights/publications/final-report-human-rights-and-technology.

33 Ibid, 82-83.

34 Ibid, 80.

35 NSW Government, "Residential Tenancy Complaints" (Sydney: NSW Fair Trading, 2023), https://www.fairtrading.nsw.gov.au/about-fair-trading/our-services/resolving-issues/residential-tenancy-complaints.

36 Anti-Discrimination NSW, "How to Make a Complaint" (Sydney: Anti-Discrimination NSW, 2023), https://antidiscrimination.nsw.gov.au/anti-discrimination-nsw/complaints/how-to-make-a-complaint.html

37 *Civil and Administrative Tribunal Act 2013* (NSW) schs. 3-6.

38 Kimberlee Weatherall et al, *ADM+S submission to the Safe and responsible AI in Australia discussion paper* (Sydney: ARC Centre of Excellence for Automated Decision-Making and Society, 2023), 12, https://www.admscentre.org.au/submissions/.

39 Sophie Farthing, John Howell, Katerina Lecchi, Zoe Paleologos, Phoebe Saintilan and Edward Santow, *Human Rights and Technology Final Report* (Sydney: Australian Human Rights Commission, 2021), 78-79, https://humanrights.gov.au/our-work/technology-and-human-rights/publications/final-report-human-rights-and-technology.

40 Natalie Sheard, "Employment Discrimination by Algorithm: Can anyone be held responsible?", *UNSW Law Journal,* 45:2 (2022), 617-648, 637. See also Jack O'Brien, "Artificial Intelligence as a Diagnostic Tool: What does it mean for the Field of Medical Negligence", *ANU Jolt*, 14 April 2022, https://anujolt.org/post/1475-artificial-intelligence-as-a-diagnostic-tool-what-does-it-mean-for-the-field-of-medical-negligence.

41 Sophie Farthing, John Howell, Katerina Lecchi, Zoe Paleologos, Phoebe Saintilan and Edward Santow, *Human Rights and Technology Final Report* (Sydney: Australian Human Rights Commission, 2021), 78, https://humanrights.gov.au/our-work/technology-and-human-rights/publications/final-report-human-rights-and-technology.

42 Miles Brundage et al., "Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims", *Arxiv: Computer Science-Computers and Society* (2020), 11, https://arxiv.org/abs/2004.07213.

43 Andrew D. Selbst, "An Institutional View of Algorithmic Impact Assessments", *Harvard Journal of Law and Technology*, 35:1 (2021), 117-191.

44 Jakob Mökander, Jonas Schuett, Hannah Rose Kirk and Luciano Floridi, "Auditing large language models: a three-layered approach", *AI and Ethics* (2023), 1-31, https://link.springer.com/article/10.1007/s43681-023-00289-2.

45 Matthew Scherer and Ridhi Shetty, "New York City's Law on Using Tech to Make Hiring Decisions Keeps Getting Weaker", *Slate*, 6 February 2023, https://slate.com/technology/2023/02/aedts-new-york-city-ll-144-automated-employment-decision-tools.html.

46 European Commission, *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative Acts* (Brussels: European Commission, 2021), Title III, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206.

47 Canadian Government, *2023 Algorithmic Impact Assessment* (Ottawa: Government of Canada, 2023), https://open.canada.ca/aia-eia-js/?lang=en.

48 Canadian Government, *The Artificial Intelligence and Data Act (AIDA) – Companion Document* (Ottawa: Government of Canada, 2023), https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document#s9.

49 Canadian Government, *Algorithmic Impact Assessment* (Ottawa: Government of Canada, 2023), https://open.canada.ca/aia-eia-js/?lang=en.

50 European Commission, *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts* (Brussels: European Commission, 2021), Title III, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206.

51 *Automated Employment Decision Tools*, New York City Council Local Law 144 of 2021 (2021), https://www.nyc.gov/site/dca/about/automated-employment-decision-tools.page.

52 Kimberlee Weatherall et al, *ADM+S submission to the Safe and responsible AI in Australia discussion paper*, (Sydney: ARC Centre of Excellence for Automated Decision-Making and Society, 2023), 4, https://www.admscentre.org.au/submissions/.

53 Ibid, 3.

54 Jakob Mökander, Jonas Schuett, Hannah Rose Kirk and Luciano Floridi, "Auditing large language models: a three-layered approach", *AI and Ethics* (2023), https://link.springer.com/article/10.1007/s43681-023-00289-2; Sophie Farthing, John Howell, Katerina Lecchi, Zoe Paleologos, Phoebe Saintilan and Edward Santow, *Human Rights and Technology Final Report* (Sydney: Australian Human Rights Commission, 2021), 103, https://humanrights.gov.au/our-work/technology-and-human-rights/publications/final-report-human-rights-and-technology.

55 Ibid.

56 Miles Brundage et.al, "Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims", *Arxiv: Computer Science-Computers and Society* (2020), 11, https://arxiv.org/abs/2004.07213.

57 Ibid, 11.

58 Sophie Farthing, John Howell, Katerina Lecchi, Zoe Paleologos, Phoebe Saintilan and Edward Santow, *Human Rights and Technology Final Report* (Sydney: Australian Human Rights Commission, 2021), 103, https://humanrights.gov.au/our-work/technology-and-human-rights/publications/final-report-human-rights-and-technology.

59 Alex Engler, "The AI regulatory toolbox: How governments can discover algorithmic harms", *Brookings TechTank,* 9 October 2023, https://www.brookings.edu/articles/the-ai-regulatory-toolbox-how-governments-can-discover-algorithmic-harms/

60 Mariana Mazzucato, *Mission-oriented public procurement: lessons from international examples*. (London: UCL Institute for Innovation and Public Purpose, 2020), https://www.ucl.ac.uk/bartlett/public-purpose/pr2020-20.

61 Mariana Mazzucato, Marietje Schaake, Seb Krier and Josh Entsminger, *Governing Artificial Intelligence in the public interest*, UCL Institute for Innovation and Public Purpose and Stanford Cyber Policy Center, Working Paper WP/2022/12 (2022), https://www.ucl.ac.uk/bartlett/public-purpose/sites/bartlett_public_purpose/files/governing_ai_in_public_interest.pdf.

62 NSW Government and Investment NSW. *Shaping the Future of NSW in science and technology-Final Report* (Sydney: NSW Government and Investment NSW, 2022), https://www.chiefscientist.nsw.gov.au/__data/assets/pdf_file/0003/508935/FINAL-Report-20-Year-NSW-R-and-D-Roadmap_WEB_DPS.pdf.

63 Ibid.

64 Mariana Mazzucato, *Mission-oriented public procurement: lessons from international examples* (London: UCL Institute for Innovation and Public Purpose, 2020), https://www.ucl.ac.uk/bartlett/public-purpose/pr2020-20.

65 NSW Government, *Discussion Paper #001* (Sydney: Office of the Anti-Slavery Commissioner NSW, 2022), https://dcj.nsw.gov.au/documents/legal-and-justice/anti-slavery-commissioner/nsw-asc-discussion-paper-nsw-public-procurement-and-modern-slavery.pdf.

66 Mariana Mazzucato, *Mission Economy: A Moonshot Guide to Changing Capitalism* (Sydney: Penguin Books Australia, 2022).

67 Elvira Uyarra, Jon Mikel Zabala-Iturriagagoitia, Kieron Flanagan and Edurne Magro. "Public procurement, innovation and industrial policy: Rationales, roles, capabilities and implementation", *Research Policy*, Vol. 49 (2020), 1-11.

68 Ibid.

69 Lance Worrall, John Spoehr and Hamish Gamble. *Innovation Procurement: Lessons for Australia.* (Adelaide: Australian Industrial Transformation Institute, Flinders University, 2022), https://www.flinders.edu.au/content/dam/documents/research/aiti/aiti_innovation_procurement_report.pdf.

70 Stéphan Vincent-Lancrin and Reyer van der Vlies, *Trustworthy artificial intelligence (AI) in education: Promises and Challenges,* OECD Education Working Papers, No.218 (2020), https://www.oecd-ilibrary.org/docserver/a6c90fa9-en.pdf?expires=1693284990&id=id&accname=guest&checksum=387FE17883D75B146E57FA9C39BC9AFD.

71 Kalervo N.Gulson, "Artificial Intelligence and a new global policy problem in education" in Fazal Rizvi, Bob Lingard and Risto Rinne (Eds), *Reimagining Globalization and Education* (New York: Routledge: 2022), 87.

72 Leslie Loble AM and Aurora Hawcroft, *Shaping AI and edtech to tackle Australia's learning divide*, (Sydney: University Technology Sydney and Paul Ramsay Foundation, 2022), https://www.uts.edu.au/sites/default/files/2022-12/PRF-UTS-Shaping-AI-EdTech-LOBLE-report.pdf.

73 José-Miguel Bello Y Villarino & Kalervo N.Gulson, "Artificial Intelligence, regulation and the purposes of education". Later version forthcoming in Regine Paul, Emma Carmel, and Jennifer Cobbe (eds.), *Handbook on Public Policy and Artificial Intelligence* (Cheltenham: Edward Elgar, 2024).

74 NSW Government, *Procurement Policy Framework* (Sydney: NSW Government, 2022), https://info.buy.nsw.gov.au/__data/assets/pdf_file/0020/1065503/Procurement-Policy-Framework-1.9-April-2022-Full-V1.pdf, (accessed 30 August 2023).

75 NSW Government, *NSW Artificial Intelligence Assurance Framework* (Sydney: NSW Government, 2022) https://www.digital.nsw.gov.au/policy/artificial-intelligence/nsw-artificial-intelligence-assurance-framework, (accessed 30 August 2023).

76 NSW Government, *Mandatory Ethical Principles for the use of AI* (Sydney: NSW Government, 2022), https://www.digital.nsw.gov.au/policy/artificial-intelligence/artificial-intelligence-ethics-policy/mandatory-ethical-principles, (accessed 30 August 2023).

77 Estonian Government, , *2023 Vision – Bürokratt: the next level of digital state development in e-Estonia* (Tallinn: Estonian Ministry of Economic Affairs and Communications, 2023), https://www.kratid.ee/en/burokratt-visioon.

78 E-Estonia.com, "Estonia partners with Microsoft in developing the virtual assistant Bürokratt" (Tallin: E-Estonia.com, 2022), 28 November 2022, https://e-estonia.com/estonia-partners-with-microsoft-in-developing-the-virtual-assistant-burokratt/ (accessed 30 August 2023).

79 Ibid.

80 NSW Government, *Procurement: Artificial Intelligence Strategy* (Sydney: NSW Government, 2022), https://www.digital.nsw.gov.au/policy/artificial-intelligence/artificial-intelligence-strategy/procurement,.

81 Ibid.

82 Ibid.

83 Ibid.

84 Ibid.

85 NSW Government, *Supplier Code of Conduct*, (Sydney: NSW Government, 2022), https://info.buy.nsw.gov.au/policy-library/policies/supplier-code-of-conduct.

86 *Modern Slavery Act 2018* (NSW) s 27(1).

87 *Modern Slavery Act 2018* (NSW) s 27(2).

88 NSW Government, "NSW Anti-slavery Commissioner and Clean Energy Council join forces to tackle modern slavery risks in renewables", 17 August, 2023, https://dcj.nsw.gov.au/justice/anti-slavery-commissioner/news-and-media/nsw-anti-slavery-commissioner-clean-energy-council-renewables.html, (accessed 30 August 2023).

89 OECD, *Regulatory Sandboxes in Artificial Intelligence* (Paris: OECD Digital Economy Papers, No.356, 2023), www.oecd-ilibrary.org/deliver/8f80a0e6-en.pdf?itemId=/content/paper/8f80a0e6-en&mimeType=pdf.

90 Ibid, 14.

91 Ibid.

92 Ibid.

93 CSIRO, *AI regulation* (Canberra: CSIRO, 2023), https://research.csiro.au/ss/science/projects/responsible-ai-pattern-catalogue/ai-regulation/ (accessed 28 August 2023).

94 OECD, *Regulatory Sandboxes in Artificial Intelligence*, OECD Digital Economy Papers No.356 (2023), 15 , www.oecd-ilibrary.org/deliver/8f80a0e6-en.pdf?itemId=/content/paper/8f80a0e6-en&mimeType=pdf.

95 Ibid,13.

96 Ibid, 15.

97 Ibid, 15.

98 United Kingdom Government, *The Guide to the Sandbox* (London: Information Commissioner's Office, 2023), https://ico.org.uk/for-organisations/advice-and-services/regulatory-sandbox/the-guide-to-the-sandbox/.

99 Shlomit Yanisky-Ravid and Sean K. Hallisey, ""Equality and Privacy by Design: A New Model of Artificial Intelligence Data Transparency via Auditing, Certification and Safe Harbor Regimes", Fordham Urban Law Journal, 46:2 (2019), 428-486.

100 Ibid.

101 Ibid.

102 Lauren Solomon and Nicholas Davis, *The State of AI Governance in Australia* (Sydney: Human Technology Institute, UTS, 2023), 30, https://www.uts.edu.au/sites/default/files/2023-05/HTI%20The%20State%20of%20AI%20Governance%20in%20Australia%20-%2031%20May%202023.pdf.

103 Alistair Reid, Simon O'Callaghan, and Yaya Lu, *Implementing Australia's AI Ethics Principles: A selection of Responsible AI practices and resources* (Sydney: Gradient Institute and CSIRO, 2023), https://www.csiro.au/-/media/D61/NAIC/Gradient-Report/23-00122_DATA61_REPORT_NAIC-ResponsibleAITools_WEB_230620.pdf.

104 Angie Abdilla, Megan Kelleher, Rick Shaw and Tyson Yunkaporta, *Out of the Black Box: Indigenous protocols for AI* (Paris: UNESCO, 2021), 2, https://static1.squarespace.com/static/5778a8e3e58c62bbf1a639ae/t/61808f1d034eda41942223a9/1635815199890/*Final+Unesco+Paper_Designed.pdf.

105 Ibid.

106 Ibid.

107 Ibid, 2.

108 NSW Government, "Digital Uplift", https://www.digital.nsw.gov.au/policy/artificial-intelligence/artificial-intelligence-strategy/digital-uplift (accessed 30 August 2023).

109 Ibid.

110 Mariana Mazzucato, Marietje Schaake, Seb Krier and Josh Entsminger, *Governing artificial intelligence in the public interest* (London: UCL Institute for Innovation and Public Purpose and Stanford Cyber Policy Center, Working Paper WP/2022/12), https://www.ucl.ac.uk/bartlett/public-purpose/sites/bartlett_public_purpose/files/governing_ai_in_public_interest.pdf.

111 Seth Lazar and Alondra Nelson, "AI Safety on whose terms?", *Science,* 381:6654 (2023), 138.

112 See for example, Fernando Santiago Rodriguez and Nadejda Komendatova, *Approaches to Participatory Policymaking Processes: Technical Report* (Vienna: United Nations Industrial Development Organization, March 2022); Helene Landemore, *Open Democracy: Reinventing Popular Rule for the Twenty-First Century* (Princeton NJ: Princeton University Press, 2020); Elizabeth Anderson, "The Epistemology of Democracy", *Episteme: A Journal of Social Epistemology,* 3:1-2 (2006), 8-22; Joshua Cohen and Charles Sabel, "Directly Deliberative Polyarchy", *European Law Journal,* 3:4 (1997), 313-342.

113 Joshua Cohen and Charles Sabel, "Directly Deliberative Polyarchy", European *Law Journal,* 3:4 (1997), 313-342, 318-319; Helene Landemore, *Open Democracy: Reinventing Popular Rule for the Twenty-First Century* (Princeton NJ: Princeton University Press, 2020), 138-140.

114 See for example, Elizabeth Anderson, "The Epistemology of Democracy", *Episteme: A Journal of Social Epistemology,* 3:1-2 (2006), pp. 8-22; Joshua Cohen and Charles Sabel, "Directly Deliberative Polyarchy", *European Law Journal,* 3:4 (1997), 313-342; Helene Landemore, *Open Democracy: Reinventing Popular Rule for the Twenty-First Century* (Princeton NJ: Princeton University Press, 2020), 170; Helene Landemore, "Deliberation, Cognitive Diversity, and Democratic Inclusiveness: An Epistemic Argument for the Random Selection of Representatives", *Synthese,* 190 (2023), 1209-1231; Mariana Mazzucato, *Mission Economy: A Moonshot Guide to Changing Capitalism* (London: Penguin Books, 2021), 199.

115 Joshua Cohen and Charles Sabel, "Directly Deliberative Polyarchy", European *Law Journal,* 3:4 (1997), 313-342, 318-319.

116 Nicole Gillespie, Steve Lockey, Caitlin Curtis, Javad Pool and Ali Akbari, *Trust in Artificial Intelligence: A Global Study* (Brisbane: University of Queensland and KPMG Australia, 2023), https://ai.uq.edu.au/files/6161/Trust%20in%20AI%20Global%20Report_WEB.pdf.

117 See, for example, Helene Landemore, *Open Democracy: Reinventing Popular Rule for the Twenty-First Century* (Princeton: Princeton University Press, 2020), pp. xiv-xv; 218; Helene Landemore, Andrew Sorota and Audrey Tang, "Why picking citizens at random could be the best way to govern the AI revolution", *Fortune,* 20 June 2023, https://fortune.com/2023/06/20/why-picking-citizens-at-random-best-way-to-govern-ai-revolution-tech-politics/.

118 NSW Government, "Have Your Say" (Sydney: NSW Government, 2023), https://www.nsw.gov.au/have-your-say.

119 University of Sydney, *2022 No Data about Us Without Us: Community Responses to the idea of a National Disability Data Asset Report* (Sydney: University of Sydney Policy Lab, 2022), 49, https://www.ndda.gov.au/wp-content/uploads/2022/11/no-data-about-us-without-us-sydney-policy-lab.pdf.

120 Australian Government, "Message from the Council July", 17 July 2020, https://www.ndda.gov.au/message-from-the-council-july/.

121 Canadian Citizens' Assembly on Democratic Expression, "Canadian Citizens' Assembly on Democratic Expression: Recommendations to strengthen Canada's response to new digital technologies and reduce the harm caused by their misuse" (Ottawa: Public Policy Forum, 2021), https://ppforum.ca/articles/recommendations-to-strengthen-canadas-response-to-new-digital-technology-and-reduce-the-harm-caused-by-their-misuse/.

122 Ibid.

123 Institute for Advanced Study, *Comment of the AI Policy and Governance Working Group on the NTIA AI Accountability Policy Request for Comment Docket NTIA-230407-0093* (Washington DC: Institute for Advanced Study, AI Policy and Governance Working Group, 2023), https://www.ias.edu/sites/default/files/AI%20Policy%20and%20Governance%20Working%20Group%20NTIA%20Comment.pdf

124 OECD, *Regulatory Sandboxes in Artificial Intelligence,* OECD Digital Economy Papers, No.356 (2023), 20 www.oecd-ilibrary.org/deliver/8f80a0e6-en.pdf?itemId=/content/paper/8f80a0e6-en&mimeType=pdf.

125 US Government, *2023 Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (Gaithersburg: Department of Commerce, 2023),https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf; Nicholas Davis, Sophie Farthing and Edward Santow, *Discussion Paper, Safe and responsible AI in Australia Submission to Department of Industry, Science and Resources,* (Sydney: Human Technology Institute, The University of Technology Sydney, 2023), 7-8, https://www.uts.edu.au/sites/default/files/2023-08/23.08.09%20HTI%20Submission%20to%20DISR%20Safe%20and%20Responsible%20AI%20in%20Australia%20DP.pdf. See also the socio-technical model of classifying AI systems developed by the OECD in the *OECD Framework for the Classification of AI Systems* (OECD Digital Economy Papers, no.322, 2022), https://www.oecd.org/publications/oecd-framework-for-the-classification-of-ai-systems-cb6d9eca-en.htm.

126 UNESCO, *2022 Working Group on AI Capacity Building Report-Artificial Intelligence and Digital Transformation: Competencies for Civil Servants* (Paris: UNESCO, 2022), 16-17, https://unesdoc.unesco.org/ark:/48223/pf0000383325.

127 Ibid, 36.

128 Ibid, 36.

129 Ibid.

130 See Neil Rees, Simon Rice and Dominique Allen, *Australian Anti-Discrimination and Equal Opportunity Law, 3rd ed.* (Sydney, The Federation Press, 2018); *Anti-Discrimination Act 1977* (NSW) ss 7, 24, 38B, 39, 49B, 49S, 49ZG, 49ZYA.

131 Neil Rees, Simon Rice and Dominique Allen, *Australian Anti-Discrimination and Equal Opportunity Law, 3rd ed.* (Sydney, The Federation Press, 2018), 53.

132 *Purvis v New South Wales* (2003) 217 CLR 92, 134.

133 See *Australian Human Rights Commission Act 1986* (Cth); *Racial Discrimination Act 1975* (Cth); *Sex Discrimination Act 1984* (Cth); *Disability Discrimination Act 1992* (Cth); *Age Discrimination Act 2004* (Cth).

134 Australian Human Rights Commission, "A Quick Guide to Australian anti-discrimination laws" (Sydney: Australian Human Rights Commission, 2014), https://humanrights.gov.au/our-work/employers/quick-guide-australian-discrimination-laws.

135 See *Australian Human Rights Commission Act 1986* (Cth), *Age Discrimination Act* 2004 (Cth), *Disability Discrimination Act 1992* (Cth), *Racial Discrimination Act 1975* (Cth), *Sex Discrimination Act 1984* (Cth) and *Fair Work Act 2009* (Cth).

136 Australian Human Rights Commission, "A Quick Guide to Australian Discrimination Laws" (Sydney: Australian Human Rights Commission, 2014), https://humanrights.gov.au/our-work/employers/quick-guide-australian-discrimination-laws; see also *Australian Human Rights Commission Act 1986* (Cth).

137 Sophie Farthing, John Howell, Katerina Lecchi, Zoe Paleologos, Phoebe Saintilan and Edward Santow, *Human Rights and Technology Final Report* (Sydney: Australian Human Rights Commission, 2021), 105, https://humanrights.gov.au/our-work/technology-and-human-rights/publications/final-report-human-rights-and-technology.

138 See, for example, The European Commission Advisory Committee on Equal Opportunities for Women and Men, *Opinion on Artificial Intelligence – Opportunities and Challenges for Gender Equality* (Brussels: European Commission, 2020); Marc Cheong, Reeva Lederman, Aidan McLoughney, Sheilla Njoto, Leah Ruppanner and Anthony Wirth, *Ethical Implications of AI Bias as a Result of Workforce Gender Imbalance* (Melbourne: School of Computing and Information Systems and The Policy Lab, University of Melbourne, 2020), 11.

139 Sophie Farthing, John Howell, Katerina Lecchi, Zoe Paleologos, Phoebe Saintilan and Edward Santow, *Human Rights and Technology Final Report* (Sydney: Australian Human Rights Commission, 2021), 105-110, https://humanrights.gov.au/our-work/technology-and-human-rights/publications/final-report-human-rights-and-technology.

140 Ibid, 106.

141 Sophie Farthing, John Howell, Katerina Lecchi, Zoe Paleologos, Phoebe Saintilan and Edward Santow, *Human Rights and Technology Final Report* (Sydney: Australian Human Rights Commission, 2021), 106, https://humanrights.gov.au/our-work/technology-and-human-rights/publications/final-report-human-rights-and-technology; United Kingdom Government, *Review into Bias in Algorithmic Decision-making* (London: Centre for Data Ethics and Innovation, 2020), 3, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/957259/Review_into_bias_in_algorithmic_decision-making.pdf.

142 Valerie Schneider, "Locked out by big data: How big data, algorithms and machine learning may undermine housing justice", *Columbia Human Rights Law Review*, 52:1 (2020), 251-305; National Consumer Law Centre, "SafeRent Solutions Accused of Illegally Discriminating Against Black and Hispanic Rental Applicants" (Boston: National Consumer Law Centre, 2022), https://www.nclc.org/saferent-solution-accused-of-illegally-discriminating-against-black-and-hispanic-rental-applicants.

143 Jeffrey Dastin, "Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women", *Reuters*, 11 October 2018, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G.

144 Australian Human Rights Commission, *Guidance Resource: Artificial Intelligence and discrimination in insurance pricing and underwriting* (Sydney: Australian Human Rights Commission, 2022), 27-37, https://humanrights.gov.au/sites/default/files/document/publication/ai_guidance_resource_december_2022.pdf

145 Tilly Cook, Aiden Greenhill and Emer Sheehy, "Discriminatory Pricing: Exploring the 'ethnicity penalty' in the insurance market (London: Citizens Advice , 2022), https://www.citizensadvice.org.uk/Global/CitizensAdvice/Consumer%20publications/Report%20cover/Citizens%20Advice%20-%20Discriminatory%20Pricing%20report%20(4).pdf.

146 Ibid, 6.

147 Adewole S Adamson and Avery Smith, "Machine learning and health care disparities in dermatology", *JAMA Dermatol*, 154:11 (2018), 1247-1248.

148 Carmen Niethammer, "AI bias could put women's lives at risk: A challenge for regulators", *Forbes*, 2 March 2020, https://www.forbes.com/sites/carmenniethammer/2020/03/02/ai-bias-could-put-womens-lives-at-riska-challenge-for-regulators/?sh=56af47fe534f.

149 Ziad Obermeier, Brian Powers, Christine Vogeli and Sendhil Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations", *Science,* 366:6464 (2019), 447-453.

150 OECD, "2019 OECD AI Principles- Transparency and Explainability-Principle 1.3" (Paris: OECD, 2019), https://oecd.ai/en/dashboards/ai-principles/P7.

151 Neil Rees, Simon Rice and Dominique Allen, *Australian Anti-Discrimination and Equal Opportunity Law, 3rd ed.* (Sydney, The Federation Press, 2018), 108.

152 *Anti-Discrimination Act 1977* (NSW) s. 4A.

153 Natalie Sheard "Employment Discrimination by Algorithm: Can anyone be held responsible?", *UNSW Law Journal,* 45:2 (2022), 617-648.

154 Australian Human Rights Commission, *Guidance Resource: Artificial Intelligence and discrimination in insurance pricing and underwriting* (Sydney: Australian Human Rights Commission, 2022), https://humanrights.gov.au/our-work/technology-and-human-rights/publications/guidance-resource-ai-and-discrimination-insurance.

155 Finn Lattimore, Simon O'Callaghan, Zoe Paleologos, Alistair Reid, Edward Santow, Holli Sargeant and Andrew Thomsen, *Using artificial intelligence to make decisions: Addressing the problem of algorithmic bias – Technical Paper* (Sydney: Australian Human Rights Commission, 2020), https://humanrights.gov.au/sites/default/files/document/publication/ahrc_technical_paper_algorithmic_bias_2020.pdf

156 CSIRO and Data61, "Responsible AI Pattern Catalogue", https://research.csiro.au/ss/science/projects/responsible-ai-pattern-catalogue/, CSIRO and Data61, 2023 (accessed 30 August 2023).

157 NSW Bar Association, "Issues Arising from the Use of AI Language Models (including ChatGPT) in Legal Practice" (Sydney: NSW Bar Association, 2023), https://inbrief.nswbar.asn.au/posts/9e292ee2fc90581f795ff1df0105692d/attachment/NSW%20Bar%20Association%20GPT%20AI%20Language%20Models%20Guidelines.pdf.

158 See Info-Communications Media Development Authority (IMDA Singapore) and Personal Data Protection Commission Singapore (PDPC), *Compendium of Use Cases: Practical Illustrations of the Model AI Governance Framework* (2020), https://go.gov.sg/ai-gov-use-cases and *Compendium of Use Cases: Practical Illustrations of the Model AI Governance Framework - Volume 2* (2020), https://go.gov.sg/ai-gov-use-cases-2.

159 United States Government, *The Americans with Disabilities Act and the Use of Software, Algorithms and Artificial Intelligence to Assess Job Applicants and Employees* (Washington DC: U.S. Equal Employment Opportunity Commission, 2022), https://www.eeoc.gov/laws/guidance/americans-disabilities-act-and-use-software-algorithms-and-artificial-intelligence.

160 United Kingdom Government, *Guidance on AI and Data Protection: What about fairness, bias and discrimination?* (London: Information Commissioner's Office, 2023), https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/how-do-we-ensure-fairness-in-ai/what-about-fairness-bias-and-discrimination/ (accessed 15 November, 2023).

Transforming public policy

Level 1, 60 Martin Place
Sydney, NSW 2000 Australia

info@jmi.org.au

jmi.org.au