



**Submission of the
Australian Discrimination Law Experts Group**

in response to the

***Safe and Responsible AI in Australia: Discussion
Paper***

26 July 2023

TABLE OF CONTENTS

Australian Discrimination Law Experts Group	2
Glossary	3
1. Summary.....	4
2. Introduction to the submission.....	5
3. Potential gaps in approaches.....	6
Application of Australia's discrimination law framework to AI and ADM systems.....	6
Reform of federal discrimination laws.....	8
4. Target areas	10
5. Risk-based approaches	11
Support for a risk-based approach.....	11
Elements of the risk-based approach in Attachment C	13
Limitations of a risk-based approach.....	16

Australian Discrimination Law Experts Group

This submission is made on behalf of the undersigned members of the Australian Discrimination Law Experts Group ('ADLEG'), a group of legal academics with significant experience and expertise in discrimination and equality law and policy.

This submission focuses on key questions raised in the *Safe and Responsible AI in Australia: Discussion Paper* released in June 2023 by the Australian Government: Department of Industry, Science and Resources.

We are happy to answer any questions about the submission or other related issues, or to provide further information on any of the areas covered. Please let us know if we can be of further assistance in this inquiry, by contacting Natalie Sheard on N.Sheard@latrobe.edu.au.

This submission may be published.

This submission was coordinated and authored by:

Natalie Sheard, La Trobe Law School, La Trobe University

Written contributions were provided by:

Associate Professor Alysia Blackham, Melbourne Law School, University of Melbourne

Professor Beth Gaze, Melbourne Law School, University of Melbourne

Professor Simon Rice OAM, Sydney Law School, University of Sydney

This submission is endorsed by:

Dr Robin Banks, Faculty of Law, University of Tasmania

Liam Elphick, Faculty of Law, Monash University

Professor Beth Goldblatt, Faculty of Law, University of Technology Sydney

Associate Professor Anne Hewitt, Adelaide Law School, University of Adelaide

Dr Sarah Moulds, Senior Lecturer in Law, University of South Australia

Dr Peta Spirou, Lecturer in Law, Adelaide Law School, University of Adelaide

Dr Bill Swannie, Thomas More Law School, Australian Catholic University

Assistant Professor Alice Taylor, Faculty of Law, Bond University

Glossary

General

AI	Artificial Intelligence as defined in the Discussion Paper (page 5)
ADM	Automated Decision Making as defined in the Discussion Paper (page 5)
ADLEG	Australian Discrimination Law Experts Group
AHRC	Australian Human Rights Commission

Federal Laws

DDA	<i>Disability Discrimination Act 1992</i> (Cth)
SDA	<i>Sex Discrimination Act 1984</i> (Cth)

1. Summary

This submission is made in response to the Department of Industry, Science and Resources' *Safe and Responsible AI in Australia: Discussion Paper* ('the Discussion Paper') released in June 2023.

As set out in further detail below, our primary recommendations are as follows:

Recommendation 1: The Australian Government resource the AHRC and state and territory equality agencies to undertake an expeditious review of the extent to which Australia's current discrimination law framework protects against discrimination by AI and ADM systems.

Recommendation 2: The Australian Government strengthen federal discrimination laws by enacting a positive duty in all federal discrimination laws.

Recommendation 3: AI applications which use facial analysis techniques be banned completely as having an 'unacceptable risk' of harm.

Recommendation 4: All elements of a risk-based regulatory approach be mandated through legislation which applies equally to public and private organisations engaged in the development and/or deployment of AI and ADM systems.

Recommendation 5: The Australian Government strengthen legal frameworks for protecting and enforcing fundamental human rights, including the right to equality and private life.

Recommendation 6: The Australian Government establish an expert and well-resourced regulator to oversee the development and deployment of AI and ADM systems by public and private organisations.

2. Introduction to the submission

As experts in Australian discrimination law, we are concerned with how law can protect individuals against unlawful discrimination. The likelihood of this occurring in the use of AI and ADM systems has been well documented in recent years, and the need to ensure adequate protection for individuals affected by AI and ADM decision-making underpins this submission.

This submission builds on ADLEG's submission made in response to the Department of Prime Minister and Cabinet's, *Automated Decision Making and AI Regulation: Issues Paper* released in March 2022 ('ADLEG's first submission').

ADLEG commends the Discussion Paper for seriously considering the risks of AI and ADM systems and aiming to develop governance mechanisms to ensure the safe and responsible use of AI. The desire to avoid a piecemeal regulatory environment is also supported.

In this submission, we focus on gaps in existing regulatory approaches, specifically federal anti-discrimination law (Part 3), target areas (Part 4) and the proposed risk-based approach (Part 5).

3. Potential gaps in approaches

2. What potential risks from AI are not covered by Australia's existing regulatory approaches? Do you have suggestions for possible regulatory action to mitigate these risks?

AI and ADM systems are currently in use in Australia in a range of different areas including credit scoring,¹ workplace recruitment and performance management,² and policing.³ As identified in the Discussion Paper, one of the biggest risks or dangers of AI is that the adoption of these systems will perpetuate and amplify existing societal and historical discrimination.⁴ There is substantial evidence that the adverse effects of AI and ADM systems fall disproportionately on marginalised and disadvantaged groups.⁵ As Australia becomes an increasingly unequal society,⁶ AI and ADM systems have the potential to cause harm to these groups at scale.

Application of Australia's discrimination law framework to AI and ADM systems

Not all discrimination is unlawful. Individuals will only be protected from a discriminatory decision or treatment of an AI or ADM system where that decision or treatment is prohibited by Australia's discrimination laws. However, as identified by the AHRC in its *Human Rights and Technology* final report,⁷ Australia is yet to clarify how existing discrimination legislation applies to AI and ADM systems.⁸

-
- ¹ Daniel Ziffer, 'Threat of 'Postcode Discrimination' as Credit Scores Skewed By Where You Live', *abc.net.au* (online) 7 February 2022 <<https://www.abc.net.au/news/2022-02-07/threat-of-postcode-discrimination-in-credit-scores/100723574>>.
- ² Andrew Taylor, 'Talk to the Robot: How Technology is Changing the Way You Find a Job', *smh.com.au* (online) 14 November 2021 <<https://www.smh.com.au/national/talk-to-the-robot-how-technology-is-changing-the-way-you-find-a-job-20211112-p598ir.html>>.
- ³ Jake Goldenfein, 'Australian Police Are Using the Clearview AI Facial Recognition System With No Accountability', *The Conversation* (online) 4 March 2020 <<https://theconversation.com/australian-police-are-using-the-clearview-ai-facial-recognition-system-with-no-accountability-132667>>.
- ⁴ Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Penguin Books, 2016).
- ⁵ Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor* (St Martin's Press, 2018).
- ⁶ Australian Council of Social Service and University of New South Wales (2018) *Inequality in Australia 2018* <<https://www.acoss.org.au/wp-content/uploads/2018/07/Inequality-in-Australian-2018.pdf>>; Australian Council of Social Service and University of New South Wales (2020) *Inequality in Australia 2020 - Part 1; overview* <<https://apo.org.au/node/308062>>; Australian Council of Social Service and University of New South Wales (2020) *Inequality in Australia 2020 – Part 2: who is affected and why* <<https://apo.org.au/node/310195>>.
- ⁷ Australian Human Rights Commission ('AHRC'), *Human Rights and Technology* (Final Report, March 2021) ('AHRC – HR and Technology Report').
- ⁸ Ibid 108.

This task is an urgent one. Research in this area has found that current federal anti-discrimination laws may not provide protection against all cases of algorithmic discrimination.⁹ For example:

- i. the direct discrimination provisions in federal anti-discrimination law require that a ‘person’ engage in the discriminatory treatment. As AI and ADM systems are not legal entities, this requirement may not be satisfied when the discriminatory decision is made by those systems without any input from a natural person.¹⁰ It is argued that law reform is needed to provide certainty as to legal liability in this area.¹¹
- ii. many instances of discrimination by AI or ADM hiring systems will fall for consideration under the indirect discrimination provisions of federal anti-discrimination law. However, these provisions have been fraught with jurisprudential difficulty, and the consideration of complex socio-technical systems poses particular problems for the judiciary.¹² In addition, discrimination by AI or ADM hiring systems is statistical discrimination,¹³ but Australia currently lacks clear statistical thresholds or measures of disadvantage and there are no standardised tests for determining when AI or ADM systems have discriminated against protected groups.¹⁴
- iii. new groups of people experience discrimination by AI and ADM systems but they are not covered by anti-discrimination laws as they do not possess attributes which are protected in those laws;¹⁵
- iv. federal anti-discrimination laws do not permit intersectional claims of discrimination;¹⁶
- v. the decision of the Full Federal Court in *Sklavos v Australian College of Dermatologists* [2017] FCAFC 128 (16 August 2017) and its interpretation of the positive duty found in sections 5(2) and 6(2) of the DDA limits the scope of the protection to situations where the reason for refusing to make an adjustment is the person’s disability. Legislative

⁹ See, eg, Natalie Sheard, ‘Employment Discrimination by Algorithm: Can Anyone be Held Accountable?’ (2022) 45(2) *University of New South Wales Law Journal* 617; Alysia Blackham, ‘Setting the Framework for Accountability for Algorithmic Discrimination at Work’ (2023) 47(1) *Melbourne University Law Review* (in press).

¹⁰ Sheard (n 9) 17-22.

¹¹ Ibid.

¹² Sheard (n 9) 29-30.

¹³ Sandra Wachter, Brent Mittelstadt and Chris Russell, ‘Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI’ (2021) 41 (July) *Computer Law & Security Review* 105567.

¹⁴ Sheard (n 9) 29-30.

¹⁵ Sandra Wachter, ‘The Theory of Artificial Immutability: Protecting Algorithmic Groups under Anti-Discrimination Law’ (2022) 97 *Tulane Law Review* 149.

¹⁶ See, eg, Alysia Blackham and Jeromey Temple, ‘Intersectional Discrimination in Australia: An Empirical Critique of the Legal Framework’ (2020) 43(3) *University of New South Wales Law Journal* 773.

amendment is required to ensure that persons with disability can access reasonable adjustments to achieve a non-discriminatory assessment by AI and ADM systems.

- vi. information, knowledge and resource asymmetries in cases involving algorithmic discrimination present significant barriers to complainants being able to garner the evidence necessary to prove a case of algorithmic discrimination.¹⁷ It is therefore argued that the burden of proof in discrimination claims in Australia should be shifted such that, once an applicant establishes a *prima facie* case of discrimination, the burden moves to the respondent to establish that their conduct was not discriminatory.¹⁸

The AHRC recommended in its *Human Rights and Technology* final report that the Australian Government resource it to ‘produce guidelines for government and non-government bodies on complying with federal discrimination laws in the use of AI-informed decision making’.¹⁹ While guidelines will form an important part of the regulatory approach, alone they are insufficient. Individuals and groups require the protection of legally enforceable rights when they suffer AI and ADM-facilitated harm.

For this reason, ADLEG **recommends** that the first step must be for the **Australian Government to resource the AHRC and state and territory equality agencies to undertake a review of the extent to which Australia’s current discrimination law framework protects against discrimination by AI and ADM systems.** As AI and ADM systems are currently impacting individuals, groups and communities in potentially discriminatory ways, this work must be undertaken expeditiously. The AHRC’s *Human Rights and Technology* final report, and the consultations conducted for that project, provide a solid foundation for doing so.

As part of the review, the AHRC and state and territory equality agencies should provide recommendations addressing how any identified gaps or deficiencies in our current discrimination law framework should be remedied. This may occur through the amendment of current discrimination laws or through the regulation of specific technologies.

Reform of federal anti-discrimination laws

The difficulties identified above in applying existing federal anti-discrimination laws to cases of algorithmic discrimination are exacerbated by the fact that anti-discrimination laws are no longer fit for purpose and are in urgent need of reform.

¹⁷ Alysia Blackham, ‘Setting the Framework for Accountability for Algorithmic Discrimination at Work’ (2023) *Melbourne University Law Review* (advance).

¹⁸ See, eg, Alysia Blackham, *Reforming Age Discrimination Law: Beyond Individual Enforcement* (OUP, 2022) pp 233–36.

¹⁹ AHRC – HR and Technology Report (n 7) Recommendation 18.

In December 2021, the AHRC released its *Free and Equal: A Reform Agenda for Federal Discrimination Laws* report.²⁰ While this report does not examine discrimination by AI or ADM systems, a number of the recommendations are relevant to, and should be considered as part of, any proposed regulatory landscape for these systems.

In particular, ADLEG **supports** the recommended inclusion of a positive duty in all federal discrimination laws.²¹ Positive equality duties are a ‘fourth generation’ of equality law measures that seek to encourage proactive and preventive approaches to achieving equality. They provide a good example of how regulatory systems can work together. A positive duty requires all duty holders to take ‘reasonable and proportionate measures to eliminate unlawful discrimination’.²² Pursuant to such a duty, it would be reasonable and proportionate for public and private entities that develop and/or deploy AI and ADM systems to adopt any legislated AI governance measures, such as those set out in Attachment C, as there is a serious risk to human rights, including the right to equality, if such measures are not adopted.

Further, to ensure the effectiveness of positive duties, particularly in the context of AI and ADM, we **recommend** that a positive duty specifically require duty holders to consult those affected, and their representatives (such as unions) before adopting or deploying AI or ADM systems.²³ This would complement and strengthen the audit and impact assessment processes discussed below.

As acknowledged by the AHRC and **supported** by ADLEG, the implementation of a positive duty ‘should be accompanied by significant education ... as well as support for the Commission, legal assistance providers and business peak bodies to be able to provide clear and accessible guidance about the positive duty’.²⁴

²⁰ Australian Human Rights Commission (‘AHRC’), *Free & Equal: A Reform Agenda for Federal Discrimination Laws* (Report, December 2021) (‘AHRC – Free and Equal Report’).

²¹ Ibid 58–90; Recommendation 1. The SDA now contains such a positive duty: see section 47C.

²² Section 47C of the SDA provides that employers and persons conducting a business or undertaking must ‘take reasonable and proportionate measures to eliminate, as far as possible’ sex discrimination.

²³ Blackham (n 17).

²⁴ AHRC – HR and Technology Report (n 7) Recommendation 2.

4. Target areas

10. Do you have suggestions for whether any high-risk AI applications or technologies should be banned completely?

Any risk management approach for AI and ADM systems must have a risk level of ‘unacceptable risk’ as some AI applications or technologies need to be banned completely.

ADLEG **recommends** that AI applications which use facial analysis techniques fall into the category of those with an ‘unacceptable risk’ and should be banned completely. Facial analysis techniques rely on machine learning to identify and infer human characteristics, emotions and/or behaviours from an individual’s facial features and movements. These techniques, for example, are often used in AI-enabled applications in hiring and employee evaluation processes. These systems lack scientific validity, lead to discrimination, and have been described as ‘modern day phrenology’.²⁵ In particular, emotion analysis has been discredited²⁶ and the European Parliament has argued for a complete ban on the use of such techniques in the EU *Artificial Intelligence Act* (*‘AI Act’*).²⁷

In addition, ADLEG **recommends** that serious consideration be given to banning other high-risk AI applications and technologies as recommended by the European Parliament.²⁸ These include: ‘real time’ remote biometric identification systems in publicly accessible spaces, biometric categorisation systems using protected attributes under anti-discrimination laws and predictive policing systems (based on profiling, location or past criminal behaviour).²⁹

²⁵ Ifeoma Ajunwa, ‘Automated Video Interviewing as the New Phrenology’ (2021) 36 *Berkeley Technology Law Journal* 1173.

²⁶ Lisa Barrett et al, ‘Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements’ (2019) 20(1) *Psychology Science Public Interest* 1.

²⁷ European Parliament, ‘MEPs Ready to Negotiate First-Ever Rules For Safe and Transparent AI’, *europarl.europa.eu* (online) 21 July 2023 < <https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai>>.

²⁸ *Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))*.

²⁹ *Ibid.*

5. Risk-based approaches

14. Do you support a risk-based approach for addressing potential AI risks? If not, is there a better approach?
15. What do you see as the main benefits or limitations of a risk-based approach? How can any limitations be overcome?
17. What elements should be in a risk-based approach for addressing potential AI risks? Do you support the elements presented in Attachment C?
20. Should a risk-based approach for responsible AI be a voluntary or self-regulation tool or be mandated through regulation? And should it apply to: a. public or private organisations or both; b. developers or deployers or both?

Support for a risk-based approach

Yes, the risk-based approach as outlined in this submission is **supported** by ADLEG. Such an approach is preventative, operates at the systemic level and recognises that some uses of AI should be banned completely. It is also the approach gaining international acceptance.³⁰

ADLEG **recommends** that all elements of a risk-based regulatory approach be mandated through legislation. Public and private organisations should not be permitted to develop and/or deploy and use AI and ADM systems until they meet the requirements of that legislation.

It is essential that a binding and enforceable legal obligation to adopt AI governance measures be placed on developers and deployers of AI and ADM systems, particularly those with ‘medium’ and ‘high’ risks of harm. Such AI governance measures ensure the proactive anticipation of harms, and require the developers and deployers of such systems to take preventative action to minimise or avoid those harms.

We argue that what is needed is a coherent and comprehensive approach to legal regulation of the technologies, rather than a mere expectation that technologies can be regulated by making minimal amendments to existing legal rules. The issues raised by these technologies are too profound for piecemeal regulatory responses. Nor can they be left to self-regulation approaches alone, as they can seriously affect the rights of individuals in both public and private sectors. Coherent legal regulation is necessary to protect individuals who can be adversely affected by these technologies and, as outlined in Part 3 of this submission, may not have any or adequate remedies under existing law.

³⁰ This is the approach adopted in the EU *AI Act*, the USA’s NIST AI Risk Management Framework and the Canadian *Directive on Automated Decision Making* (pages 16-20 of the Discussion Paper).

There is evidence to suggest that a voluntary or self-regulation approach will be insufficient to address the serious risks of harm from AI and ADM systems. In her empirical research project, ADLEG member, Natalie Sheard, conducted semi-structured and in-depth interviews with recruiters and HR specialists in a range of industries to understand how they operationalise and use AI-based hiring systems ('AHSs'). She found that organisations that are already deploying and using AHSs:

- i. had little knowledge of Australia's voluntary AI Ethics Framework;
- ii. had little understanding of AI and AHSs including the risks and shortcomings of these systems; and
- iii. adopted almost no AI governance measures prior to or during their deployment and use of AHSs. Impact assessments and monitoring and evaluation frameworks were rare. Many organisations had automated the 'screening out' of job applicants from the recruitment process without appropriate human oversight. While some organisations provided notice and an explanation of the use of an AHS, these explanations lacked meaningful detail and were often mediated through AHS vendors, for example, by referring job seekers to their websites.³¹

These findings are consistent with those in the University of Technology Sydney's Human Technology Institute's ('HTI') *State of AI Governance in Australia* report.³² The HTI found that 'few Australian organisations have implemented systematic and structured forms of governance around AI systems'.³³

Legally mandated AI regulation, therefore, has a significant role to play in the education of deployers of ADM systems and providing clarity regarding their obligations when such systems are used. As one of Ms Sheard's research participants (from a large financial services organisation which deploys AHSs at scale) stated:

[Regulation] can't come soon enough. In the absence of that you're just looking around for best practice, and you're trying to solve on the fly. Yeah, absolutely without a doubt we need that regulation and governance. I don't doubt that most people would be trying to use these products with the best of intention, but ultimately the more governance we have the standards continue to lift for our providers and ultimately influence the way we use these in practice, and I think that's nothing but a positive.³⁴

³¹ Natalie Sheard, *Algorithmic Discrimination in Recruitment in Australia* (PhD Thesis, La Trobe University, 2023) (forthcoming).

³² L Solomon and N Davis, *The State of AI Governance in Australia* (Report, 31 May 2023).

³³ Ibid 28.

³⁴ Participant 5.

ADLEG **recommends** that a risk-based approach should apply to public and private organisations engaged in the development and/or deployment of such systems. While we recognise the difference in the nature of decision-making in the public and private (or business) spheres, individuals can be adversely affected in serious ways in both contexts. Using the current definitions in the Discussion Paper,³⁵ for a ‘medium risk’ application of AI, the risk is described in Box 4 (page 32) as ‘[h]igh impacts that are ongoing and difficult to reverse’. For ‘high risk’ applications of AI, the risk is described as ‘[v]ery high impacts that are systemic, irreversible or perpetual’. These risks can be introduced to AI and ADM systems not just by the designers and developers of these systems but also by those in the public and private sectors who deploy and use them. There is, therefore, no justification for differentiating between the public and private sectors or between developers and deployers of these systems. Effective protection of individuals, in both public and private spheres, will be essential to the ‘social licence’ of these technologies.

Elements of the risk-based approach in Attachment C

All of the elements outlined in Attachment C form a necessary part of a risk-based approach to AI regulation. ADLEG **recommends** they be strengthened as follows:

i. Impact Assessment

Greater clarity and guidance must be provided to organisations regarding all aspects of the requirements of, and the methodology for conducting, an impact assessment.

In ADLEG’s first submission, we advocated for a ‘technology impact assessment’. Such an assessment includes the consideration of potential adverse impacts including infringements of all relevant laws or ethical guidelines. It is focused on the specific *context of the proposed use of an AI or ADM system* and must include *consultation with the individuals, group(s) and communities likely to be affected by it*.

The advantage of a broader technology impact assessment from the legal perspective, instead of separate assessments based on single legal fields, is that it should bring together a consideration of potential harms in all the areas of law that are relevant, such as discrimination, privacy, human rights, administrative law and consumer protection. This will provide a coherent analysis of the potential impact of a technological use and a basis for identifying necessary safeguards to prevent harms arising.

For ‘medium’ and ‘high’ risk AI applications, such an assessment should be publicly available and accessible.

³⁵ Any risk-based approach must develop clear definitions for the different levels of risk.

ii. Notice and Explanation

A risk-based approach to AI regulation must be underpinned by strong protections for fundamental human rights, including individual transparency and data protection rights. These rights are an essential part of AI accountability and must be supported by a legislated right for individuals to be notified when they are subject to automated decision making.³⁶ For this reason, is it essential that the review of the Privacy Act be finalised and amendments to enshrine such a right be enacted as soon as possible.

Individuals must also have a legislated right to obtain meaningful information about the AI system when they are subjected to automated decision-making. For AI applications assessed as ‘medium risk’, the right to an explanation proposed in Box 4 (page 33) is inadequate. Developers and/or deployers of such systems should also be required to provide a system explanation, including details of the risks and limitations of such systems. For example, for ‘AI-enabled applications in hiring and employee evaluation processes’ such a requirement would ensure that job seekers are provided with an explanation of the criteria on which they will be assessed, how those criteria are weighted, and the manner of assessment. Such a right is essential to give agency to job seekers to determine if they wish to interact with the AI or ADM system, and provides a basis for determining if they require any reasonable adjustments under anti-discrimination laws. It is also an important pre-requisite to contestability of AI-enabled decisions.³⁷

ADLEG **recommends** that the following risk management elements be added to those in Attachment C:

i. AI Audits

An independent ‘AI audit’ prior to deployment, and at regular intervals thereafter, should be mandated for all developers and/or deployers of ‘medium’ and ‘high’ risk AI applications including those which assess creditworthiness, the performance of teachers and students and/or are involved in hiring and employee evaluation processes.

AI audits are a process through which an auditor evaluates an AI system or product according to a specific set of criteria and provides findings and recommendations.³⁸ AI audits have an important role to play in identifying whether AI systems meet stated performance targets, discriminate against protected groups, infringe privacy and data protection rights, provide transparency and accountability, result in environmental

³⁶ See, eg, Natalie Sheard “No Notice and No Explanation: The Incontestability of Employment Discrimination by Algorithm (2022) 35(2) *Australian Journal of Labour Law* 119.

³⁷ Ibid.

³⁸ Inioluwa Raji, Sasha Costanza-Chock and Joy Buolamwini, ‘Change From the Outside: Towards Credible Third Party Audits of AI Systems’ (Report, 2023) 8.

impacts and adhere to existing regulatory requirements.³⁹ The results of such audits should be publicly available and accessible.

An example of such an audit requirement is provided in the New York City Council's *Automated Employment Decisions Tools* legislation.⁴⁰ This law prohibits the use by an employer of an automated employment decision tool to screen a job candidate for an employment decision unless the tool has been the subject of such a 'bias audit' by an independent auditor in the year prior, and a summary of the results is made publicly available on the employer's website.⁴¹

There is, however, one major obstacle to adopting AI audits. There is a lack of internationally agreed and clearly defined standards that contain high-level expectations for AI systems and their use.⁴² In particular, standards to assess the risk of potential discrimination by an AI or ADM system have been slow to emerge and we still lack international consensus as to what test or metrics of discrimination should be used in any given use case. In Australia, it is also not clear how those metrics align with legal tests of discrimination under Australian laws. Again, this is an area in which the AHRC, state and territory equality agencies, and/or another appropriate regulatory body, should be supported to develop concrete guidance. Without this, public confidence in AI audits is likely to be eroded as they will be of variable quality and may not identify real risks of discrimination.

ii. Subject Matter Expertise

Deployers of 'medium' and 'high' risk AI applications should have and give effect to in-house subject matter expertise on the system.⁴³ Such expertise is necessary to understand the risks and shortcomings of the system, including the risk of algorithmic discrimination. This may require employing someone with a data science background.⁴⁴ Without such expertise, organisations are unlikely to understand the recommendations or predictions of their AI system or to develop effective human-in-the-loop processes to oversee these potentially harmful systems. Having such expertise also enables those

³⁹ Ibid.

⁴⁰ Automated Employment Decision Tools Act, NY Admin Law §20-871(2)(b)(1) (2021). See also the Discussion Paper, page 19.

⁴¹ Ibid s 5-301.

⁴² Raji, Constanza-Chock and Buolamwini (n 38) 15.

⁴³ This is a key recommendation of the Australian Merit Protection Commissioner in her guidance material for public sector agencies: see 'Guidance Material for Using AI-Assisted Recruitment Tools', *Merit Protection Commissioner* <<https://www.mpc.gov.au/resources/guidance/myth-busting-ai-assisted-and-automated-recruitment-tools>>.

⁴⁴ Ibid.

organisations to answer any questions raised by the individuals, groups or communities that are subject to these technologies.

As the developers and vendors of the AI system in use are not independent parties, but rather are conflicted by profit motives, they are not appropriate organisations to provide this expertise. Further, developers and vendors of AI systems rarely acknowledge or undertake responsibility for how a product is ultimately deployed, creating an accountability gap.⁴⁵ It is necessary for all those involved in the development and deployment of such technologies to have relevant in-house expertise and accountability.

Limitations of a risk-based approach

There are well documented limitations to a risk-based approach. Kaminski⁴⁶ identifies three limitations of AI risk regulation that are applicable to the proposals in the Discussion Paper. First, a risk regulation approach is not aimed at providing either compensation for injured people or civil recourse in the form of individual process.⁴⁷ This is a key criticism of the EU *AI Act* which, it is argued, has not been drafted with end-users in mind and lacks individual remedies for AI harms.⁴⁸ In particular, it has been argued that the EU AI Act fails to include avenues of redress for victims of discrimination and does not align its provisions with existing obligations under the EU *Directive on Equal Treatment in Employment and Occupation*.⁴⁹

Second, Kaminski argues that risks that cannot be readily quantified, such as risks to privacy and other human rights, may be devalued or ignored.⁵⁰ Third, she asserts that the developing law of AI has been dominated by a particular version of regulation that she calls the ‘enterprise risk management model of risk regulation’.⁵¹ This model relies heavily on self-governance and internal organisational infrastructure to mitigate risks.⁵² She argues that it differs from what stakeholders appear to want which is ‘risk regulation modelled on environmental law, which entails public transparency and public participation’.⁵³

It is therefore essential that, if a risk-based approach to AI regulation is adopted in Australia, it is accompanied by effective and enforceable avenues of redress when individuals, groups and communities suffer AI harms. Contestability ‘is a core mechanism for establishing and

⁴⁵ Blackham (n 17).

⁴⁶ Margot Kaminski, ‘The Developing Law of AI; A Turn to Risk Regulation’ (Report, 2023) <<https://www.documentcloud.org/documents/23783773-the-developing-law-of-ai-kaminski>>.

⁴⁷ Ibid 14-5.

⁴⁸ Ibid 15.

⁴⁹ Center for Democracy & Technology, ‘Discrimination in Hiring: The Case for Alignment of the EU AI Act with EU Equality Laws’ (online 20 March 2023) <<https://cdt.org/insights/discrimination-in-hiring-the-case-for-alignment-of-the-eu-ai-act-with-eu-equality-laws/>>.

⁵⁰ Kaminski (n 46) 15.

⁵¹ Kaminski (n 46) 16-7.

⁵² Ibid 16.

⁵³ Ibid.

preserving justice in the Western adversarial tradition’.⁵⁴ ‘Contestability’ is also a concept which forms an important part of ethical guidelines developed globally in response to algorithmic decision systems by governments, the private sector, research institutions and civil society.⁵⁵ Australia’s AI Ethics Principles elevate contestability to a core principle.⁵⁶ ‘Contestability’ is defined in the principles as the availability of ‘a timely process to allow people to challenge the use or output of the AI system’ when it ‘significantly impacts a person, community, group or environment’.⁵⁷

One means of achieving contestability is by strengthening legal frameworks for protecting and enforcing fundamental human rights, including the right to equality and private life. This should include:

- i. resourcing the AHRC and state and territory equality agencies to undertake an expeditious review of the extent to which Australia’s current discrimination law framework protects against discrimination by AI and ADM systems (as recommended above);
- ii. enacting the reform agenda set out in the *Free and Equal: A Reform Agenda for Federal Discrimination Laws*⁵⁸ particularly with regard to positive duties (as recommended above);
- iii. enacting national human rights protection, as is currently being investigated by the Attorney-General. Human Rights Acts have been passed in the ACT,⁵⁹ Victoria,⁶⁰ and Queensland;⁶¹ and
- iv. better resourcing the legal assistance sector, including legal aid organisations and community legal centres, to provide legal advice and representation to people who have suffered AI harms. The legal assistance sector plays a critical role in ensuring access to justice is provided for marginalised and disadvantaged individuals, groups and communities targeted by AI and ADM systems. The best recent example is provided by

⁵⁴ Margot Kaminski and Jennifer Urban, ‘The Right to Contest AI’ (2021) 121 *Columbia Law Review* 1957, 2042–4.

⁵⁵ As at April 2019, there were at least 84 documents containing ethical principles or guidelines for AI: A Jobin, M Ienca and E Vayena, ‘The Global Landscape of AI Ethics Guidelines’ (2019) 1 *National Machine Intelligence* 389.

⁵⁶ Department of Industry, Science, Energy and Resources (Cth), *Australia’s AI Ethics Principles*, Australian Government, Canberra, <<https://webarchive.nla.gov.au/awa/20220603081905/https://www.industry.gov.au/data-and-publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles>> Principle 7.

⁵⁷ Ibid.

⁵⁸ AHRC – Free and Equal Report (n 20).

⁵⁹ *Human Rights Act 2004* (ACT).

⁶⁰ *Charter of Human Rights and Responsibilities Act 2006* (Vic).

⁶¹ *Human Rights Act 2019* (Qld).

the Robodebt litigation. It was only when this litigation was funded and commenced by Victoria Legal Aid, after sustained advocacy by community legal centres, that this harmful and unlawful scheme ended.

In addition, a risk regulation approach to AI must be overseen by an active, expert and well-resourced regulator. The AHRC in its ‘Human Rights and Technology: Final Report’ recommended that the Australian Government establish an independent statutory office of an ‘AI Safety Commissioner’ focused on promoting safety and protecting human rights in the development and use of AI.⁶² The HTI’s ‘Facial Recognition Technology: Towards a Model Law’ report also calls for a ‘suitable regulator to ... oversee the development and use of [Facial Recognition Technology] ... in Australia’.⁶³

ADLEG **supports** these calls to establish a new regulator, or to expand the role and responsibilities of an existing regulator, to oversee the development and deployment of AI and ADM systems by public and private organisations. Such a regulator should be empowered to:

- i. develop and/or oversee and/or coordinate the development of authoritative AI standards or guidelines including those relating to AI audits;
- ii. oversee the development of impact assessments by public and private organisations; and
- iii. conduct ‘own motion’ investigations and audits of AI and ADM systems to identify adverse impacts.

Further, as a key way to ensure contestability of AI harms, this regulator must also be empowered to investigate and determine complaints of AI harm lodged by affected individuals or another party acting as a representative for such an individual or group of individuals.

⁶² AHRC, HR and Technology Final Report (n 7) Recommendation 22.

⁶³ Nicholas Davis, Lauren Perry and Ed Santow, *Facial Recognition Technology: Towards a Model Law* (Report, 2022).