
Advance Copy — Citation

Natalie Sheard, 'Employment Discrimination by Algorithm: Can Anyone Be Held Accountable?' (2022) 45(2) *University of New South Wales Law Journal* (Advance).

**EMPLOYMENT DISCRIMINATION BY ALGORITHM:
CAN ANYONE BE HELD ACCOUNTABLE?**

NATALIE SHEARD*

The use by employers of Algorithmic Hiring Systems ('AHSs') to automate or assist with recruitment decisions is occurring in Australia without legal oversight. Regulators are yet to undertake an analysis of the legal issues posed by their use. Academic literature on this topic is limited and judicial guidance is yet to be provided.

This article examines to what extent, if at all, Australian anti-discrimination laws are able to regulate the use by employers of discriminatory AHSs. First, it examines the re-emergence of blatant discrimination by digital job advertising systems. Second, it considers who, if anyone, is liable for automated discrimination. Third, it examines the law's ability to regulate 'proxy' discrimination. Finally, it explores whether indirect discrimination provisions can provide redress for the disparate impact of an AHS.

Australia's anti-discrimination laws are long overdue for reform. This article concludes that new legislative provisions, as well as non-binding guidelines, specifically tailored to the use by employers of algorithmic decision systems are needed.

* Natalie Sheard is a lawyer working at the intersection of law and technology and a PhD candidate at Law School, La Trobe University. I would like to acknowledge the helpful comments from Professor Louis de Koker and Associate Professor Karen O'Connell, as well as the anonymous reviewers of this article when under submission. I also thank Tom Read for his feedback and proofreading of earlier drafts.

I INTRODUCTION

‘Algorithmic decision-making is the civil rights issue of the 21st century.’

– Ifeoma Ajunwa¹

We increasingly live in a ‘scored society’² where algorithms³ rank and rate us, mediate our access to essential goods and services and impact important rights and freedoms.⁴ Machine learning algorithms,⁵ fuelled by large datasets, are in operation in many public and private domains. They are used to predict the likelihood of recidivism, to assess loan applications and to evaluate and score teachers.⁶ In the field of human resources, algorithmic decision systems⁷ make predictions as to which job applicant will be the ‘best performing’ or ‘best fit’ and are transforming and automating recruitment and hiring decisions.

The use of Algorithmic Hiring Systems (‘AHSs’) is widespread in other jurisdictions, particularly in sectors where employment is more precarious, such as retail, and on the rise in Australia. They promise time and cost savings for employers, improved quality of hires and, significantly, the ability to remove subjectivity and human bias from the recruitment process.⁸ But AHSs have also been described by artificial intelligence (‘AI’) experts as giving employers a ‘license to discriminate’.⁹ They may entrench and perpetuate historical discrimination, with the potential to cause harm at considerable scale.

-
- 1 Ifeoma Ajunwa, quoted in Susan Kelley, ‘Social Scientists Take on Data Driven Discrimination’, *Cornell Chronicle*, (online, 13 February 2019) <<https://news.cornell.edu/stories/2019/02/social-scientists-take-data-driven-discrimination>> (emphasis added).
 - 2 Danielle Keats Citron and Frank Pasquale, ‘The Scored Society: Due Process for Automated Predictions’ (2014) 89(1) *Washington Law Review* 1.
 - 3 An algorithm is ‘the mathematical logic behind any type of system that performs tasks or makes decisions’: ‘Algorithmic Accountability Policy Toolkit’ (Toolkit No 1, AI Now Institute, October 2018) 2.
 - 4 These include equality rights and the right to work: *International Covenant on Civil and Political Rights*, opened for signature 16 December 1966, 999 UNTS 171 (entered into force 3 January 1976) art 3; *International Covenant on Economic, Social and Cultural Rights*, opened for signature 16 December 1966, 993 UNTS 3 (entered into force 3 January 1976) art 6.
 - 5 Machine learning, a subset of artificial intelligence (‘AI’), is ‘a set of techniques and algorithms that can be used to “train” a computer program to automatically recognize patterns in a set of data’: ‘Algorithmic Accountability Policy Toolkit’ (n 3) 2.
 - 6 See, eg, Julia Angwin et al, ‘Machine Bias’, *ProPublica* (online, 23 May 2016) <<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>>; Cathy O’Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Penguin Books, 2016).
 - 7 Algorithmic decision systems are systems that ‘use automated reasoning to aid or replace a decision-making process that would otherwise be performed by humans’: ‘Algorithmic Accountability Policy Toolkit’ (n 3) 2.
 - 8 See, eg, Pymetrics, Submission to Australian Human Rights Commission, *Human Rights and Technology* (2 October 2018) <<https://perma.cc/TMY2-GKL6>>.
 - 9 Meredith Whittaker quoted in Drew Harwell, ‘A Face-Scanning Algorithm Increasingly Decided Whether You Deserve the Job’, *The Washington Post* (online, 6 November 2019) <<https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/>>.

Internationally, algorithmic decision-making and the related issues of bias, privacy and data protection have received extensive attention by regulators,¹⁰ in academic literature¹¹ and by the FAT (Fairness, Accountability and Transparency) research community.¹² In addition, there is a developed body of North American and European literature considering the use of AHSs in the employment context,¹³ including the interaction with national non-discrimination and data protection laws.¹⁴ Specific analysis regarding the bias mitigation measures adopted in these systems has also been undertaken.¹⁵

In Australia, this is a new and emerging field of law and research. Regulators are yet to undertake a comprehensive and detailed analysis of the legal issues

-
- 10 See, eg, Science and Technology Committee, *Algorithms in Decision-Making* (House of Commons Report No 4 of Session 2017–19, 23 May 2018); World Economic Forum, ‘How to Prevent Discriminatory Outcomes in Machine Learning’ (White Paper, March 2018); Claude Castellucia and Daniel Le Métayer, ‘Understanding Algorithmic Decision-Making: Opportunities and Challenges’ (Study, European Parliament, Panel for the Future of Science and Technology, March 2019); Centre for Data Ethics and Innovation, *Review into Bias in Algorithmic Decision-Making* (Report, November 2020).
- 11 See, eg, in the United States (‘US’): Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press, 2015); O’Neil (n 6); Kate Crawford and Jason Shultz, ‘Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms’ (2014) 55(1) *Boston College Law Review* 93; Citron and Pasquale (n 2). See, eg, in the European Union (‘EU’): AlgorithmWatch, *Automating Society: Taking Stock of Automated Decision-Making in the EU* (Report, January 2019); David Danks and Alex John London, ‘Algorithmic Bias in Autonomous Systems’ (Conference Paper, International Joint Conference on Artificial Intelligence, 19 August 2017); Lilian Edwards and Michael Veale, ‘Slave to the Algorithm? Why a “Right to an Explanation” is Probably Not the Remedy You Are Looking For’ (2017) 16(1) *Duke Law and Technology Review* 18 <<https://doi.org/10.31228/osf.io/97upg>>; 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* 105567:1–31. <<https://doi.org/10.1016/j.clsr.2021.105567>> (‘Why Fairness Cannot Be Automated’); Bruce Goodman, ‘A Step Towards Accountable Algorithms? Algorithmic Discrimination and the European Union General Data Protection’ (Conference Paper, 29th Conference on Neural Information Processing Systems, 2016).
- 12 See, eg, ‘Scholarship’, *Fairness, Accountability and Transparency in Machine Learning* (Web Page) <<https://www.fatml.org/resources/relevant-scholarship>>; ‘Research’, *Algorithmic Justice League* (Web Page, 2022) <<https://www.ajl.org/library/research>>; ‘Publications’, *Algorithm Watch* (Web Page, 2022) <<https://algorithmwatch.org/en/publications/>>.
- 13 See, eg, Ifeoma Ajunwa and Daniel Greene, ‘Platforms at Work: Automated Hiring Platforms and Other New Intermediaries in the Organization of Work’ in Steven P Vallas and Anne Kovalainen (eds), *Work and Labor in the Digital Age* (Emerald Publishing Limited, 2019) 61; Miranda Bogen and Aaron Rieke, *Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias* (Report, 9 December 2018). See also Colin Gavaghan, Alistair Knott and James MacLaurin, *The Impact of Artificial Intelligence on Jobs and Work in New Zealand* (Final Report, 2021).
- 14 See, eg, Solon Barocas and Andrew D Selbst, ‘Big Data’s Disparate Impact’ (2016) 104(3) *California Law Review* 671 <<https://doi.org/10.2139/ssrn.2477899>>; Jon Kleinberg et al, ‘Discrimination in the Age of Algorithms’ [2018] (10) *Journal of Legal Analysis* 113 <<https://doi.org/10.1093/jla/laz001>>; Ifeoma Ajunwa, ‘The Paradox of Automation as Anti-Bias Intervention’ (2020) 41(5) *Cardozo Law Review* 1671; Pauline Kim, ‘Manipulating Opportunity’ (2020) 106(4) *Virginia Law Review* 867; Pauline Kim, ‘Data-Driven Discrimination at Work’ (2016) 58(3) *William and Mary Law Review* 857.
- 15 See, eg, Javier Sánchez-Monedero, Lina Dencik and Lilian Edwards, ‘What Does It Mean to “Solve” the Problem of Discrimination in Hiring? Social, Technical and Legal Perspectives from the UK on Automated Hiring Systems’ (Conference Paper, Conference on Fairness, Accountability and Transparency, 27–30 January 2020); Manish Raghavan et al, ‘Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices’ (Conference Paper, Conference on Fairness, Accountability and Transparency, 27–30 January 2020).

and challenges posed by the use of AHSs by employers.¹⁶ Academic literature examining the ability of Australia's anti-discrimination framework to protect against discrimination in employment by the use of AHSs is limited.¹⁷ Judicial guidance is yet to be provided as cases involving discriminatory algorithms have not come before the courts. There is a dearth of empirical research about the existence, operation and impact of AHSs.¹⁸ In Australia, the focus of regulatory and scholarly attention has been on 'ethical AI'¹⁹ and the use of algorithmic tools in government decision-making,²⁰ particularly in the administrative law and law enforcement contexts,²¹ and, more broadly, through assessments as to compliance with rule of law principles.²²

This article begins filling this doctrinal gap. It aims to provide a detailed consideration of whether Australia's current anti-discrimination framework is able to regulate discrimination by employers using AHSs.²³ It will examine three AHSs in use by employers: digital job advertisements, Curriculum Vitae ('CV') parsing and video interviewing systems. Part II examines how these three AHSs are deployed in recruitment and hiring processes. Part III reviews some of the mechanisms by which discrimination by employers using these AHSs may occur, including through unrepresentative training data and system design bias. Part IV critically analyses whether discrimination by an employer using an AHS is unlawful by applying existing anti-discrimination law concepts of 'direct' and

-
- 16 The Australian Human Rights Commission ('AHRC'), in its *Human Rights and Technology* final report recommended that the Australian Government resource it to 'produce guidelines for government and non-government bodies on complying with federal anti-discrimination laws in the use of AI-informed decision-making': see recommendation 18 in Australian Human Rights Commission, *Human Rights and Technology* (Final Report, March 2021) ('*AHRC HR and Technology Final Report*') 108, 195.
- 17 See Mark Burdon and Paul Harpur, 'Re-Conceptualising Privacy and Discrimination in the Age of Talent Analytics' (2014) 37(2) *University of New South Wales Law Journal* 679. There is also a limited discussion of the interaction of AI-informed decision-making and Australia's anti-discrimination framework in *AHRC HR and Technology Final Report* (n 16) 105–8. See also Finn Lattimore et al, 'Using Artificial Intelligence to Make Decisions: Addressing the Problem of Algorithmic Bias' (Technical Paper, Australian Human Rights Commission, November 2020) ('*AHRC Algorithmic Bias Technical Paper*') which investigated how algorithmic bias can arise, the nature of any bias and how these problems might be addressed by businesses.
- 18 This is not unique to Australia: see, eg, Sánchez-Monedero, Dencik and Edwards (n 15).
- 19 See, eg, Department of Industry, Science, Energy and Resources (Cth), 'Australia's AI Ethics Principles', *Australia's Artificial Intelligence Ethics Framework* (Web Page) <<https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework/ai-ethics-principles>>; Toby Walsh et al, *The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve our Wellbeing* (Report, July 2019).
- 20 See, eg, Lyria Bennett Moses and Louis de Koker, 'Open Secrets: Balancing Operational Secrecy and Transparency in the Collection and Use of Data by National Security and Law Enforcement Agencies' (2017) 41(2) *Melbourne University Law Review* 530.
- 21 See, eg, Lyria Bennett Moses and Janet Chan, 'Algorithmic Prediction in Policing: Assumptions, Evaluation and Accountability' (2018) 28(3) *Policing and Society* 806 <<https://doi.org/10.1080/10439463.2016.1253695>>; Lyria Bennett Moses and Janet Chan, 'Using Big Data for Legal and Law Enforcement Decisions: Testing the New Tools' (2014) 37(2) *University of New South Wales Law Journal* 643.
- 22 Monika Zalnieriute, Lyria Bennett Moses and George Williams, 'The Rule of Law and Automation of Government Decision-Making' (2019) 82(3) *Modern Law Review* 425 <<https://doi.org/10.1111/1468-2230.12412>>.
- 23 Burdon and Harpur's analysis laid some of the groundwork: see above n 17. Consideration of the discrimination provisions in section 351 of the *Fair Work Act 2009* (Cth) is outside the scope of this article.

‘indirect’ discrimination. If discrimination by an AHS is not unlawful, there is no liability on the part of employers for resulting harm.

Part IV begins with an analysis of the re-emergence of blatant direct discrimination by digital job advertising tools and concludes that these are covered by direct discrimination provisions. Second, it considers who, if anyone, is liable for automated discrimination, that is, where the discriminatory decision is made by an algorithmic model in an AHS and not a natural person. As an algorithm cannot be considered a ‘person’ for the purpose of anti-discrimination law, the direct discrimination provisions arguably do not apply and attributing the decision of an algorithm to an employer is problematic. As a result, to come within these provisions, it becomes necessary to frame the discriminatory conduct as that of a ‘person’. This is a difficult ‘fit’ and the application of the law is made more complex by its uncertain state. Third, it examines the law’s ability to regulate discrimination by an AHS on the basis of a personal feature, such as a person’s postcode, which is not itself protected by discrimination legislation but may be highly correlated with protected attributes (known as ‘proxy discrimination’).²⁴ It concludes that protection is dependent on the availability of the ‘characteristics extension’ and the ‘mapping’ of particular attributes onto protected ones. Finally, it explores whether indirect discrimination provisions can provide redress for the disparate impact on protected groups of the use by an employer of an AHS. While this is theoretically possible, the analysis demonstrates that this will require judicial understanding of complex socio-technical algorithmic systems and engagement with difficult questions of public policy.

Although in urgent need of regulatory and academic attention, this article does not explore the significant enforcement obstacles faced by individual complainants who allege discrimination by an employer using an AHS. Chief among these is the absence of a legislative requirement that notice be provided by an employer to a job applicant that an AHS will be or has been used in the recruitment process. Nor is there a requirement that access, in an appropriate format, be provided by an employer to essential information about the AHS.²⁵ In addition, unlike some foreign jurisdictions, it is not necessary that humans oversee automated decision-making by AHSs when significant rights are affected by these systems²⁶ nor that

24 ‘Proxy discrimination occurs when a facially-neutral trait is utilised as a stand-in – or proxy – for a prohibited trait’: Anya ER Prince and Daniel Schwarcz, ‘Proxy Discrimination in the Age of Artificial Intelligence and Big Data’ (2020) 105(3) *Iowa Law Review* 1257, 1267. For example, in the US, the practice of ‘redlining’, in which financial institutions demarcated postcodes that were effectively off limits for issuing loans even for creditworthy borrowers, is a well-recognised form of proxy discrimination. As African-Americans and Latin-Americans predominantly resided in those postcodes, they were disproportionately denied access to those home loans. See, eg, Khristopher J Brooks, ‘Redlining’s Legacy: Maps Are Gone but the Problem Hasn’t Disappeared’, *CBS News* (online, 12 June 2020) <<https://www.cbsnews.com/news/redlining-what-is-history-mike-bloomberg-comments/>>.

25 This would include the ‘source code’, the training datasets, feature selection and information as to the algorithmic model’s purpose or key priorities.

26 In the EU, article 22 of the General Data Protection Regulation provides that, without explicit consent, data subjects have the ‘right not to be subject to a decision based solely on automated processing ... which produces legal effects concerning him or her or similarly significantly affects him or her’: *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the*

the employer retain data and keep records evidencing the decision and decision-making process.

II THE USE OF AN AHS IN RECRUITMENT

In recruitment, AHSs can be used during each of the four stages of the ‘hiring funnel’: sourcing, screening, interviewing and selection of candidates.²⁷ The business case presented for using these tools is strong. It is asserted that AHSs increase efficiency in the recruitment process by reducing the time it takes and the amount it costs to hire.²⁸ For example, the Hilton International hotel empire credits automated video interview assessments with reducing the average time to hire from six weeks to five days.²⁹ In addition, AHSs are touted to improve organisational performance and reduce turnover by increasing the quality of hires (by ensuring that ‘best fit talent’ is selected).³⁰ Employers seeking to increase diversity in their workplaces may also be attracted to these systems. A key promise is they can remove actual and implicit bias from the hiring process, thereby increasing the hiring prospects of groups who face barriers to economic participation, such as those from lower socio-economic groups or with a disability.³¹

Although there are a multiplicity of AHS products on the global market, there is no reliable evidence as to the market share of these systems.³² The majority of these tools have been designed and developed by private corporations in the United States (‘US’)³³ and incorporate features, such as bias mitigation, that are attuned to US legal frameworks.³⁴ The algorithms at the heart of these products are not transparent nor are they publicly available for external scrutiny or audit. The ‘black box’³⁵ of the AHS will usually be protected as commercially confidential or the company’s intellectual property.

Algorithmic hiring tools are currently used by a number of large global companies, not only to screen applicants for retail and low wage positions but also for white collar and professional positions.³⁶ It has been suggested that uptake by

Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC [2016] OJ L 119/1.

27 Raghavan et al (n 15).

28 Bogen and Rieke (n 13).

29 Harwell (n 9).

30 Pymetrics (n 8) 2.

31 Ibid.

32 Sánchez-Monedero, Dencik and Edwards (n 15) 1.

33 Raghavan et al (n 15).

34 Sánchez-Monedero, Dencik and Edwards (n 15) 2.

35 Pasquale (n 11).

36 A recent report estimated that 99% of Fortune 500 companies use AHSs to assist with human resources, recruitment and/or hiring needs: Linda Qu, ‘99% of Fortune 500 Companies Use Applicant Tracking Systems’, *Jobscan* (Blog Post, 7 November 2019) <<https://www.jobscan.co/blog/99-percent-fortune-500-ats/>>. See also Ajunwa and Greene (n 13).

Australian businesses has been lower,³⁷ but little empirical data exist to support this view. Recruitment companies, Seek and LinkedIn, currently use CV parsing systems for positions at all levels.³⁸ A leading vendor of pre-employment assessments, HireVue, asserts that it has over 700 global corporate customers, including many with a presence in Australia such as Vodafone and PricewaterhouseCoopers ('PwC').³⁹ The author is currently conducting empirical research aimed at documenting the range of AHSs in use in Australia and understanding how they are used and deployed, for example, to be made compliant with local law.

III HOW AN AHS DISCRIMINATES

Anti-discrimination statutes in Australia prohibit discrimination on the basis of particular attributes ('protected attributes'), such as race, disability or sex,⁴⁰ when engaging in specific activities in 'public life', such as work, education, the provision of goods or services or access to places and facilities. The concept of 'work' is generally considered to be broader than that of 'employment'⁴¹ and all Australian statutes prohibit discrimination by an employer against a job applicant in recruitment or hiring processes.⁴² Whether these provisions prohibit discrimination by an employer using an AHS is the key question to be answered in this article.

A Digital Job Advertisements

At the 'sourcing' stage of the hiring funnel, employers seek to attract potential candidates to apply for open positions through job advertisements, job postings

37 Jennifer Hewitt, 'Artificial Intelligence Will Decide If You Get an Interview for Your Next Job', *The Australian Financial Review* (online, 31 January 2019) <<https://www.afr.com/opinion/artificial-intelligence-will-decide-if-you-get-an-interview-for-your-next-job-20190131-h1apfh>>

38 See, eg, 'How to Get Your Resume Past the Robots', *Seek* (Web Page) <<https://www.seek.com.au/career-advice/how-to-make-sure-a-human-reads-your-resume>>.

39 'HireVue Case Studies' *HireVue* (Web Page) <<https://www.hirevue.com/case-studies>>. Pymetrics asserts that its systems are 'live and compliant in 68 countries ... [and] in 16 languages': Pymetrics (n 8) 2.

40 See, eg, *Equal Opportunity Act 2010* (Vic) ('EOA') section 6 where the protected attributes are age, breastfeeding, employment activity, gender identity, disability, industrial activity, lawful sexual activity, marital status, parental status or status as a carer, physical features, political belief or activity, pregnancy, race, religious belief or activity, sex, sexual orientation, an expunged homosexual conviction and personal association (as a relative or otherwise) with a person who is identified by any of these other attributes. See also *Anti-Discrimination Act 1977* (NSW) ss 4(1), 7, 24, 24(1B), 24(1C), 38B, 39, 49B, 49T, 49ZG, 85A ('A-D Act'); *Sex Discrimination Act 1984* (Cth) ss 5, 5A, 5B, 5C, 6, 7, 7AA, 7A ('SDA'); *Racial Discrimination Act 1975* (Cth) ss 9 and 11–15 ('RDA'); *Disability Discrimination Act 1992* (Cth) ss 5–9, 14, 15 ('DDA'); *Age Discrimination Act 2004* (Cth) ss 5, 14, 15 ('ADA'). A table of protected attributes across all Australian jurisdictions is found in Beth Gaze and Belinda Smith, *Equality and Discrimination Law in Australia: An Introduction* (Cambridge University Press, 2017) 296, tbl 1.

41 Neil Rees, Simon Rice and Dominique Allen, *Australia Anti-Discrimination and Equal Opportunity Law* (Federation Press, 3rd ed, 2018) 566.

42 See, eg, EOA (n 40) section 16 which prohibits discrimination against a person:

- (a) in determining who should be offered employment; or
- (b) in the terms on which employment is offered to the person; or
- (c) by refusing or deliberately omitting to offer employment to the person ...

and individual outreach.⁴³ Today, job advertisements are placed by employers digitally on online platforms, such as Facebook, LinkedIn and Seek. Unlike newspaper advertisements of the past, these online platforms allow employers to access microtargeting, behavioural targeting and performance-driven advertising tools developed for the broader e-commerce sector.⁴⁴ These tools draw on the detailed profiles that have been created by algorithms from data provided by users or inferred from their online activity. A detailed picture of individuals, including behavioural and personality traits and protected attributes, can be generated with even discreet online activity such as ‘liking’ things on Facebook.⁴⁵

The use of targeting tools to place a digital job advertisement may facilitate unlawful discrimination as they enable employers to select the audience who views a particular advertisement. For example, in 2017, Verizon placed an advertisement on Facebook targeted at 25 to 36 year olds who lived in the US capital, or had recently visited there, and had a demonstrated interest in finance.⁴⁶ In doing so, employers such as Verizon may exclude groups of job seekers, including those with protected attributes, from ever viewing particular job advertisements.

B CV Parsing Systems

During the ‘screening’ stage of the recruitment funnel, job candidates are assessed on the basis of their experience, skills and other characteristics to create a shortlist of candidates for interview. It has become increasingly common at this stage of the process for recruitment agencies and employers to utilise an AHS to sift through and ‘parse’ the CVs of job applicants. The AHS scans the CVs of job applicants for keywords and other information that is believed to be correlated with successful hires, such as experience, job titles, former employers, universities and qualifications.⁴⁷ A structured candidate profile is then created by the system and all candidates are scored and ranked.⁴⁸ As Seek explains, ‘[r]esumes ... are only viewed by a human if the system matches the resume to the job advertisement. The others get dumped into an electronic black hole.’⁴⁹ A CV parsing tool may automatically reject more than half of the resumes they scan.⁵⁰

A good example of how bias and discrimination may be introduced to the hiring process by this type of AHS is provided by Amazon’s AI recruiting tool. In 2018, Reuters reported that Amazon had developed a tool designed to parse CVs

43 Bogen and Rieke (n 13) 13.

44 Ibid 17.

45 Crawford and Shultz (n 11) 100.

46 Julia Angwin, Noam Scheiber and Ariana Tobin, ‘Dozens of Companies Are Using Facebook to Exclude Older Workers from Job Ads’, *ProPublica* (online, 20 December 2017) <<https://www.propublica.org/article/facebook-ads-age-discrimination-targeting>>.

47 ‘How to Get Your Resume Past the Robots’ (n 38).

48 ‘How to Optimise Your CV for the Algorithms’, *Hays* (Web Page) <<https://www.hays.com.au/career-advice/resumes-cover-letters/how-to-optimise-your-cv-for-the-algorithms>>.

49 ‘How to Get Your Resume Past the Robots’ (n 38).

50 Ibid. See also, Alexis Carey, ‘Computer Algorithms “Reject up to 75 Per Cent of CV” Before a Human Ever Gets to Them’, *news.com.au* (online, 4 July 2018) <<https://www.news.com.au/finance/work/careers/computer-algorithms-reject-up-to-75-per-cent-of-cvs-before-a-human-ever-gets-to-them/news-story/71da2bd912dfb32335169127161d5d25>>.

and give job candidates a score ranging from one to five stars.⁵¹ However, when this tool was used for software developer jobs and other technical positions it was found to be discriminatory as it favoured men.⁵² Any further development of this tool was cancelled by Amazon.

Using the taxonomies that have been developed in the international literature as to the points at which bias can enter machine learning systems,⁵³ two causes of the bias against women in Amazon's CV parsing tool can be readily identified: bias as a result of the training data and bias as a result of the system design.

I Training Data

Training datasets, often comprised of metadata or formed by aggregating previously discrete datasets, are the 'grounded truth'⁵⁴ of machine learning algorithms, including AHSs. These datasets 'train' the algorithm how to process large amounts of data and dictate 'how information in input variables is positioned to predict the value of an output variable'.⁵⁵ Biased training data leads to biased algorithmic models.⁵⁶ This is often referred to as the 'bias in, bias out'⁵⁷ problem with machine learning algorithms.

A common way for training datasets to embed bias is when they contain data that is not representative of the population under consideration (this is known as 'sampling bias').⁵⁸ Any prediction that is drawn from that dataset may systematically disadvantage those who are under- or over-represented in it.⁵⁹ Sampling bias is a particular problem for AHSs for two reasons. First, many employers, particularly smaller ones, may have only limited data about employee performance and must therefore rely on small and incomplete datasets.⁶⁰ Second, for a dataset to be free of sampling bias, it must contain data regarding 'false negative cases'.⁶¹

51 Jeffrey Dastin, 'Amazon Scraps Secret AI Recruiting Tool that Showed Bias against Women', *Reuters* (online, 11 October 2018) <<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>>.

52 Ibid.

53 See, eg, Barocas and Selbst (n 14); Brent Daniel Mittelstadt et al, 'The Ethics of Algorithms: Mapping the Debate' (2016) 3(2) *Big Data and Society* 1 <<https://doi.org/10.1177/2053951716679679>>; Kleinberg et al (n 14); Solon Barocas, 'Data Mining and the Discourse on Discrimination' (Research Paper, Princeton University, 2014) <<https://pdfs.semanticscholar.org/abbb/235fcf3b163afd74e1967f7d3784252b44fa.pdf>>; Joyce Chou, Oscar Murillo and Roger Ibars, 'How to Recognise Exclusion in AI', *Medium* (online, 27 September 2017) <<https://medium.com/microsoft-design/how-to-recognize-exclusion-in-ai-ec2d6d89f850>>; Danks and London (n 11).

54 Barocas and Selbst (n 14) 683.

55 Nizan Geslevich Packin and Yafit Lev-Aretz, 'Learning Algorithms and Discrimination' in Woodrow Barfield and Ugo Pagallo (eds), *Research Handbook on the Law of Artificial Intelligence* (Edward Elgar Publishing Limited, 2018) 88.

56 An algorithmic model is the 'accumulated set of discovered relationships ... and ... can be employed to automate the process of classifying entities or activities of interest, estimating the value of unobserved variables, or predicting future outcomes': Barocas and Selbst (n 14) 677.

57 Sandra Mayson, 'Bias In, Bias Out' (2019) 128(8) *Yale Law Journal* 2218.

58 Barocas and Selbst (n 14) 684-7.

59 Ibid 681.

60 See, eg, Raghavan et al (n 15) 474.

61 Sánchez-Monedero, Dencik and Edwards (n 15) 6.

In an employment context, ‘false negative cases’ are those of suitably qualified applicants who have been wrongly rejected. These data are rarely available.

In the Amazon tool, the training data comprised CVs submitted to the company over a 10 year period. As the tech industry was male dominated in that period, the majority of the CVs came from men. The training data therefore contained a sampling bias as it did not adequately represent the relevant population (in this case, qualified applicants) due to the lack of CVs from suitably qualified women. As the algorithm had been trained or trained itself (it is not clear which) to select job applicants for interview who possessed the same experience, education and skills as the successful male employees represented in the CVs, the algorithm systematically discriminated against women. For example, CVs containing keywords such as the verbs ‘executed’ and ‘captured’ – more commonly found on the CVs of men – were assigned more weight by the models and job applicants with those CVs therefore received a higher score.⁶² CVs that used the word ‘women’s’ as in ‘women’s chess club champion’ were downgraded and those job applicants received a lower score, as did the applicants who attended two all-women’s colleges.⁶³

There is a strong argument that unbiased datasets do not exist. Discrimination in employment is well documented and persistent.⁶⁴ In Australia, it remains the area with the largest number of complaints for two out of four of the protected attributes in federal anti-discrimination legislation.⁶⁵ Therefore, it is arguable that the use of AHSs, trained as they are on historical data, will always deliver results that are infected by structural and systemic biases. As Mayson states: ‘The deep problem is the nature of prediction itself. All prediction looks to the past to make guesses about future events. In a racially stratified world, any method of prediction will project the inequalities of the past into the future.’⁶⁶

2 Design Bias

Discrimination by an algorithm in a computational system, such as an AHS, can also occur because of the ‘design, structure, and rules of operation’ of the system.⁶⁷ Mittelstadt et al, relying on the work of Friedman and Nissenbaum,⁶⁸ articulate two ways that design bias may occur: as a result of ‘technical bias’ and ‘social bias’.⁶⁹ Technical bias occurs because of technological constraints, errors,

62 Dastin (n 51).

63 Ibid.

64 See, eg, Australian Human Rights Commission, *Willing to Work: National Inquiry into Employment Discrimination Against Older Australians and Australians with Disability* (Report, 2016).

65 In 2018–19, discrimination in employment made up 73% of complaints under the *SDA* (n 40), 61% of complaints under the *ADA* (n 40), 36% of complaints under the *DDA* (n 40) and 35% of complaints under the *RDA* (n 40): Australian Human Rights Commission, *2018–2019 Complaint Statistics* (Report, 2018–19) 2.

66 Mayson (n 57) 2218.

67 John Zerilli et al, ‘Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard?’ (2018) 32(4) *Philosophy & Technology* 661, 672 <<https://doi.org/10.1007/s13347-018-0330-6>>.

68 Mittelstadt et al (n 53) 7 ff, citing Batya Friedman and Helen Nissenbaum, ‘Bias in Computer Systems’ (1996) 14(3) *ACM Transactions on Information Systems* 330 <<https://doi.org/10.1145/230538.230561>>.

69 Friedman and Nissenbaum call this ‘preexisting bias’: Friedman and Nissenbaum (n 68) 332.

inaccurate models and bad design decisions.⁷⁰ Social bias is rooted in ‘social institutions, practices and attitudes’.⁷¹ It occurs when the conscious or unconscious biases of individuals with significant input into the design and development of a system or broader societal, structural, organisational or institutional values and norms, are transmitted to the system.⁷² Social bias is well documented, as coders tend to be drawn from narrow social groups and to be white and male.⁷³

Social bias is an acute problem for AHSs. It may occur as a result of the subjective process of interpretation and translation of project and business objectives and requirements which developers engage in when they design and develop the algorithmic models for these systems.⁷⁴ For example, it may be introduced to the AHS when developers label the training data points or decide on the ‘feature selection’, that is, what candidate attributes will provide relevant data points to be built into a model.⁷⁵ It may also be encoded in the system when an employer decides on the ‘outcome of interest’ or ‘target variable’,⁷⁶ such as ‘best performing’, ‘best cultural fit’ or ‘good employee’.⁷⁷ There is no universal definition of these terms. It is well recognised that they have always been interwoven with prejudice and discriminatory assumptions and are often based on the ‘behaviour, patterns and attributes of the historically dominant group in public life (Anglo-Australian, able-bodied, heterosexual males)’.⁷⁸

C Video Interviewing Systems

Video interview assessment systems may be used by employers to pre-interview, screen or automate the first interview with job applicants. ‘Video-based assessments’, as a leading vendor, HireVue, calls them, are ‘designed to look at a unique set of personal competencies shown to be related to success in a particular job’.⁷⁹ These competencies are typically ‘soft’ skills such as ‘communication skills, conscientiousness, problem-solving skills, team orientation and initiative’.⁸⁰

Candidates video-record their responses to a series of interview questions on their desktop computer or mobile device.⁸¹ The questions are usually written by employers or chosen from a library of questions provided by the vendor. Nathan Mondragon, HireVue’s Chief Industrial-Organisational Psychologist, estimates that in the standard 30 minute interview, containing as little as 6 interview questions,

70 Ibid 335–6.

71 Ibid 333.

72 Mittelstadt et al (n 53) 7.

73 Kate Crawford, ‘Artificial Intelligence’s White Guy Problem’ *The New York Times* (online, 25 June 2016) <<https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html>>.

74 Barocas and Selbst (n 14) 677.

75 Ibid 688. See also Kleinberg et al (n 14) 134–7.

76 The target variable is the outcome that the machine learning algorithm is designed to predict: ‘Algorithmic Accountability Policy Toolkit’ (n 3) 28.

77 Barocas and Selbst (n 14) 678–80.

78 Rosemary Hunter, *Indirect Discrimination in the Workplace* (Federation Press, 1992) 5–6.

79 ‘How to Prepare for Your HireVue Assessment’, *HireVue* (Web Page, 16 April 2019) <<https://www.hirevue.com/blog/candidates/how-to-prepare-for-your-hirevue-assessment>>.

80 Ibid.

81 Ibid.

500,000 data points may be collected including those regarding the tone and facial expressions, eye contact, word selection and emotions of applicants.⁸² After the interview, a machine learning algorithm analyses these data points to determine a candidate's 'likelihood of success' in the role and each applicant's performance is scored and ranked.⁸³

In the HireVue system, a person's score is comprised of the language⁸⁴ and words they say and the 'audio features' of their voice such as tone.⁸⁵ For example, for a call centre role, 'supportive' words might be encouraged and applicants might be penalised for using 'aggressive' words.⁸⁶ HireVue decided, in January 2021, to remove visual analysis from its new assessment models.⁸⁷ However, until facial analysis is removed from existing models,⁸⁸ the way that a person's face moves, for example, to show excitement or indicate how they would respond to an angry customer, known as 'facial action units', can also make up 29% of a person's score.⁸⁹ Based on their score, candidates are grouped into high, medium and low tiers.⁹⁰ Candidates in the lowest tier are usually filtered out and automatically rejected without human intervention or reasons being provided.

The video interviews of current employees are a common source of training data for this form of AHS.⁹¹ These data are gathered by having the entire spectrum of current employees, from 'high to low achievers', take the video assessment.⁹² An algorithmic model is then developed based on the performance and behavioural attributes found in successful employees, which may include sales quotas or time taken to resolve customer calls.⁹³ HireVue also builds models for employers based on the data it has collected of 'top performers'.⁹⁴

82 Harwell (n 9).

83 Ibid.

84 HireVue asserts that recent advances in natural language processing have significantly increased the predictive power of language and that '[b]y far, the most valuable data we can pull from a video interview is the language a candidate used': Lindsay Zuloaga, 'Nonverbal Communication in Interview Assessments', *HireVue* (Web Page, 31 March 2020) <<https://web.archive.org/web/20210205153741/https://www.hirevue.com/blog/hiring/nonverbal-communication-in-interview-assessments>>.

85 Ibid.

86 Harwell (n 9).

87 Hirevue is the only video assessment system vendor to publicly announce the removal of facial analysis from its assessment systems. This step was taken in response to an official complaint filed with the Federal Trade Commission by the Electronic Privacy Information Center alleging 'unfair and deceptive practices' in violation of section 5 of the *Federal Trade Commission Act*, 15 USC §§ 41–58 (2018 & Supp 2021); see Electronic Privacy Information Center, 'Complaint and Request for Investigation, Injunction, and Other Relief' (Complaint in *HireVue Inc*, 6 November 2019) <epic.org/privacy/ftc/hirevue/EPIC_FTC_HireVue_Complaint.pdf>.

88 Hirevue will remove facial analysis from its existing models on 'a rolling basis as models come up for annual review': O'Neil Risk Consulting and Algorithmic Auditing ('ORCAA'), *Description of Algorithmic Audit: Pre-Built Assessments* (Report, 15 December 2020) 5.

89 Harwell (n 9).

90 Ibid. See also ORCAA (n 88) 1.

91 Harwell (n 9); See also Raghavan et al (n 15) 472.

92 Harwell (n 9).

93 Ibid.

94 Ibid.

Video interview assessment systems utilise vision and speech recognition technologies. A vision system enables computers to read and identify visual images,⁹⁵ while speech recognition systems generally include a language model trained on text data and an acoustic model trained on audio data.⁹⁶ There are at least two ways these systems may introduce bias. First, as with a CV parsing system, there is a significant likelihood that the data the algorithm has been trained on contains a sampling bias. Training data of the kind required for these tools are expensive and a technical challenge to build in-house. Vendors therefore usually purchase facial and speech analysis as a service from third parties, which may result in data that are of variable quality or drawn from a narrow set of sources.⁹⁷ For example, facial recognition software has been found to be less accurate at recognising women or the faces of people of colour because it has been trained on datasets comprising predominantly white male faces.⁹⁸ An assessment of five commercial speech recognition tools – developed by Amazon, Apple, Google, IBM and Microsoft – found racial disparities in performance for African Americans as a result of insufficient audio data from this group when training the models.⁹⁹ Google’s speech recognition software is 70% more likely to accurately recognise male speech because that is what it has been trained on.¹⁰⁰ It performs poorly and is likely to mischaracterise people with regional and non-native accents.¹⁰¹

The design of video interviewing assessment systems is also likely to introduce social bias. A supervised decision system¹⁰² learns by processing training data that have labels assigned to data examples, in this case, images of facial expressions and emotions as well as audio features and language content. As training data are manually assigned class labels (by humans), this process is unavoidably subjective. Classifications are not neutral and may be open to debate.¹⁰³ Evidence suggests that it is not possible to reliably and accurately identify and label the variety of cross-cultural expressions of emotion and affect.¹⁰⁴ As Barrett et al states: ‘how people communicate anger, disgust, fear, happiness, sadness, and surprise varies

95 It is arguably an example of facial recognition software.

96 Allison Koenecke et al, ‘Racial Disparities in Automated Speech Recognition’ (2020) 117(14) *Proceedings of the National Academy of Sciences of the United States of America* 7684, 7687 <<https://doi.org/10.1073/pnas.1915768117>>.

97 Raghavan et al (n 15) 475. Emotion, affect and facial expression classification systems are usually trained on the facial expressions of paid actors: Lisa Feldman Barrett et al, ‘Emotional Expressions Reconsidered: Challenges to Inferring Emotion from Human Facial Movements’ (2019) 20(1) *Psychological Science in the Public Interest* 1, 5 <<https://doi.org/10.1177/1529100619832930>>.

98 Joy Buolamwini and Timnit Gebru, ‘Gender Shades: Intersectional Accuracy in Commercial Gender Classification’ (Conference Paper, Conference on Fairness, Accountability and Transparency, 2018).

99 Koenecke et al (n 96) 7684.

100 Rachael Tatman, ‘Google’s Speech Recognition Has a Gender Bias’, *Making Noise & Hearing Things* (Blog Post, 12 July 2016) <<https://makingnoiseandhearingthings.com/2016/07/12/googles-speech-recognition-has-a-gender-bias/>>.

101 See, eg, Bogen and Rieke (n 13) 36.

102 In supervised machine learning, an algorithm uses data to learn a pattern that it can then use to predict a particular outcome, the target variable, when it sees different data: ‘Algorithmic Accountability Policy Toolkit’ (n 3) 28.

103 Barocas and Selbst (n 14) 681.

104 Barrett et al (n 97) 46.

substantially across cultures, situations, and even across people within a single situation’.¹⁰⁵ As a result, any algorithmic model in this type of AHS may have a limited ability to analyse facial action units or the audio features of a person’s voice in a non-discriminatory way. For example, black faces have been found to be read as angrier than white faces, even after controlling for the degree of smiling.¹⁰⁶ They are also read as expressing more negative emotions compared to their white counterparts if there is any ambiguity about their facial expression.¹⁰⁷ Job applicants on the autism spectrum may be penalised at disproportionate rates by any inability to make neurotypical eye contact¹⁰⁸ where this feature is unrelated to the skills and competencies of a role. As may job applicants from non-English-speaking backgrounds whose English vocabulary is more limited than their English speaking peers.¹⁰⁹

IV CHALLENGES TO LAW

The design, development and deployment of AHSs, like other technological advancements, is occurring at a pace the law is unable to match.¹¹⁰ Barocas and Selbst assert that algorithmic decision systems have facilitated the emergence of new forms of discrimination, while at the same time enabling those actions to be hidden and masked.¹¹¹ In Australia, a comprehensive review of existing legal frameworks and their ability to regulate discrimination by automated decision systems, including AHSs, is yet to be undertaken. The Australian Human Rights Commission (‘AHRC’) recently 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 anti-discrimination laws in the use of AI-informed decision-making’.¹¹² In addition, the AHRC has indicated that as part of its ‘Free and Equal: An Australian Conversation on Human Rights’ project, it will consider broad priorities for reform of federal discrimination law, which is generally regarded to be in need of ‘renewal’.¹¹³ This will include ‘how federal discrimination law should deal with the

105 Ibid 1.

106 Lauren Rhue, ‘Racial Influence on Automated Perceptions of Emotions’ (Working Paper, 9 November 2018) <<https://doi.org/10.2139/ssrn.3281765>>.

107 Ibid.

108 For a discussion of some of the barriers faced by job applicants on the autism spectrum, and ways to mitigate them, see Jill Feder, ‘Improving the Hiring Process for Autistic Candidates’, *Accessibility.com* (Blog Post, 16 April 2021) <<https://www.accessibility.com/blog/improving-the-hiring-process-for-autistic-candidates>>.

109 In its own research, HireVue found that ‘minority candidates’ give short answers to interview questions (eg ‘I don’t know’) at disproportionate rates: ORCAA (n 88) 5.

110 It is interesting to note that academic research in the field of industrial-organisational psychology has also failed to keep pace with technological change: Raghavan et al (n 15) 470.

111 Barocas and Selbst (n 14) 692.

112 See recommendation 18 in *AHRC HR and Technology Final Report* (n 16) 108, 195.

113 Rees, Rice and Allen (n 41) 31. See also Australian Human Rights Commission, ‘Priorities for Federal Discrimination Law Reform’ (Discussion Paper, October 2019) 7–9 (‘AHRC Priorities for

potential problems of algorithmic bias’ as it presents a challenge to ‘our current law’s capacity to protect against unlawful discrimination’.¹¹⁴

Similarly, there is a lack of jurisprudence about how a court or tribunal should approach adjudication in cases alleging discrimination by algorithm. At the time of writing, a case alleging discrimination by an AHS, or algorithm of any kind, is yet to come before Australian courts or tribunals. It is not known if a complaint of algorithmic discrimination, by an AHS or otherwise, has been formally lodged with federal or state anti-discrimination bodies in Australia. The data which are reported by these bodies in annual reports and conciliation registers, where they exist, are not sufficiently detailed to draw conclusions.

This Part begins filling this doctrinal gap. It is the first detailed examination of whether employers’ use of an AHS to automate recruitment and hiring decisions, is unlawful by applying the concepts of ‘direct’ and ‘indirect’ discrimination in existing Australian discrimination law.¹¹⁵ While Burdon and Harpur consider these issues in the context of workplace analytics in Australia,¹¹⁶ their analysis is at a higher level of generality and does not consider specific legislative provisions. Instead, they focus on evidentiary and enforcement obstacles for complainants and their primary argument is that complainants should have the benefit of the process protections of information privacy law.¹¹⁷

A Re-emergence of Direct Discrimination

Today, we are unlikely to see a job advertisement indicating ‘women need not apply’. Therefore, it is often assumed that direct discrimination provisions have served their purpose. As Smith asserts, the ‘battle line has at least moved forward – it is no longer drawn over blatant and intentional exclusion, but has moved to a more indirect and structural form of discrimination’.¹¹⁸ However, the examination of the operation of digital job advertisements set out below suggests that this optimism may no longer be warranted.

Reform’); Senate Standing Committee on Legal and Constitutional Affairs, Parliament of Australia, *The Effectiveness of the Sex Discrimination Act 1984 in Eliminating Discrimination and Promoting Gender Equality* (Report, December 2008); Senate Legal and Constitutional Affairs Legislation Committee, Parliament of Australia, ‘Exposure Draft of the Human Rights and Anti-Discrimination Bill 2012’ (Exposure Draft Legislation, November 2012). The Human Rights and Ant-Discrimination Bill 2012 (Cth) (‘Human Rights Anti-Discrimination Bill’) did not progress to Parliament. See also Discrimination Law Experts Group, Submission to the Attorney-General (Cth), *Consolidation of Commonwealth Anti-Discrimination Laws*, (13 December 2011) 8–11.

114 AHRC Priorities for Reform (n 113) 9.

115 Although the *Fair Work Act 2009* (Cth) also provides a cause of action to employees who suffer discrimination in employment, it is beyond the scope of this article.

116 Burdon and Harpur (n 17).

117 Burdon and Harpur propose that complainants have the benefit of ‘info-structural due process’ which could ‘ameliorate issues of structural discrimination through the greater integration of information privacy law and anti-discrimination law’: *ibid* 680. They do not set out in any detail how this would operate but suggest it would incorporate notification strategies, limits on the use of information, de-identification structures and also compliance mechanisms: *ibid* 710–11.

118 Belinda Smith in Australian Law Reform Commission, *Equality Before the Law: Justice for Women* (Report No 69(1), July 1994) and *Equality Before the Law: Women’s Equality* (Report No 69(2), December 1994) quoted in Senate Standing Committee on Legal and Constitutional Affairs (n 113) 5.12.

Definitions of ‘direct’ discrimination, in the US referred to as ‘disparate treatment’, differ between jurisdictions in Australia. This analysis adopts the more ‘modern’ formulation found in the ACT and Victorian legislation, which does not require proof of differential treatment and the complexity of the use of a comparator.¹¹⁹ The *Equal Opportunity Act 2010* (Vic) (‘EOA’) provides that ‘[d]irect discrimination occurs if a person treats, or proposes to treat, a person with an attribute unfavourably because of that attribute.’¹²⁰ It is widely recommended that this definition replace those currently found in Federal Acts.¹²¹ Under this formulation, direct discrimination occurs where ‘the consequences of the dealing with the complainant are ... adverse to the complainant’s interests and ... the dealing has occurred because of a relevant attribute of the complainant’.¹²²

As discussed in Part II, the micro-targeting tools available to an employer¹²³ when placing a digital job advertisement on platforms such as Facebook enable the employer to select ‘targeting criteria’. Those criteria permit the employer to exclude persons with protected attributes, such as those over a particular age or with a particular ethnic background, from viewing the job advertisement. This blatant and intentional exclusion of protected groups is a straightforward case of direct discrimination by an employer.¹²⁴ For example, targeting an advertisement at individuals below the age of 50 years is unfavourable to the interests of individuals over that age, as they will not see the ad and will therefore be denied the opportunity to apply for the position. This unfavourable treatment has occurred because of the protected attribute of age and the employer’s conduct is therefore unlawful. The challenge for a complainant in these circumstances is becoming aware of this ‘opaque’¹²⁵ form of discrimination in the first place and, if they do become aware of it, garnering the evidence to prove that it has occurred.

119 See *Discrimination Act 1991* (ACT) s 8(2) (‘DA’); *EOA* (n 40) s 8(1). All other federal and state jurisdictions, with the exception of the *RDA* (n 40) and the *Equal Opportunity Act 1984* (SA), adopt the ‘standard’ definition of direct discrimination which requires proof of less favourable or differential treatment. For example, section 24(1) of the *A-D Act* (n 40) provides that

[a] person (the perpetrator) discriminates against another person (the aggrieved person) on the ground of sex if the perpetrator ... treats the aggrieved person less favourably than in the same circumstances, or in circumstances which are not materially different, the perpetrator treats or would treat a person of the opposite sex ...

120 *EOA* (n 40) s 8(1).

121 See, eg, recommendation 5 in Senate Standing Committee on Legal and Constitutional Affairs (n 113) 148; Discrimination Law Experts Group (n 113) 8–9; Human Rights and Anti-Discrimination Bill 2012 (Cth) s 19(1)–(2). The Bill did not progress to Parliament.

122 *Re Prezzi and Discrimination Commissioner* [1996] ACTAAT 132, [24]. This approach was approved by the Supreme Court of Victoria in *Kuyken v Chief Commissioner of Police* (2015) 249 IR 327, 356 [94].

123 When the employer acts through an employee the employer’s actions are captured by the attributed liability provisions: see, eg, *EOA* (n 40) s 109. The employer may also be directly liable on the basis of general agency principles: *Christian Youth Camps Ltd v Cobaw Community Health Services Ltd* (2014) 50 VR 256, 286–7 (Maxwell P) (‘*Christian Youth Camps*’).

124 As such discrimination is covert, it would not be captured by the advertising offence provisions in federal and state acts as the advertisement is not published or displayed in any way which indicates an intention to engage in discrimination. See, eg, section 86 of the *SDA* (n 40) which provides that ‘[a] person shall not publish or display an advertisement ... that indicates, or could reasonably be understood as indicating, an intention to do an act that is unlawful’.

125 Crawford and Shultz (n 11) 124.

In the US, digital job advertisements have been the subject of discrimination litigation. Between 2016 and 2018, five discrimination class action lawsuits were filed against Facebook and employers. These actions were brought by civil rights groups, a national labour organisation, workers and consumers.¹²⁶ The complaints against Facebook alleged its ad-building micro-targeting tool allowed advertisers, and Facebook itself, to prevent older applicants, women and members of minority racial groups from receiving ads in relation to employment. In addition, lawsuits were filed alleging that those same tools were used to discriminate on the basis of race, national origin, sex, familial status and disability in housing ads and credit opportunities.¹²⁷

These lawsuits did not receive judicial consideration as, in March 2019, Facebook agreed to pay \$5 million compensation to settle them.¹²⁸ An important part of the settlement was an agreement by Facebook to take action on its platforms to prevent advertisers targeting housing, employment and credit ads on the basis of age, gender, or other protected categories covered by anti-discrimination law.¹²⁹ Facebook would not, however, commit to preventing discrimination outside of products and services related to housing, employment and credit ads. Outside of these areas, advertisers are still free to upload curated, and possibly discriminatory, targeting criteria for ads. There is recent evidence that, despite this agreement, it is still possible for employers to post discriminatory ads,¹³⁰ although Facebook no longer permits advertisers to target people based on ‘multicultural’ or ‘ethnic’ affinity.¹³¹

B Automated Discrimination: Who, if Anyone, Is Liable?

CV parsing and video interview systems operate quite differently to digital job advertisements. Rather than an employer making a direct and conscious decision to exclude a person with protected attributes, it may be the algorithmic model itself that operates, in effect, to make the impugned decision. For example, it is the algorithmic model that assigns a score to applicants, thereby determining who progresses to the next stage in the recruitment process and who is rejected. These recruitment decisions are routinely automated by employers and occur without human oversight.

126 They constituted three civil rights cases before the US District Court in New York (*Mobley v Facebook* (ND Cal, No 5:16-cv-06440); *National Fair Housing Alliance et al v Facebook* (SD Cal, No 1:18-cv-02689); *Riddick v Facebook* (ND Cal, No 3:18-cv-04529) (*‘Riddick’*)) and two complaints before the Equal Employment Opportunity Commission (*Spees v Facebook* (EEOC, 2018); *Communication Workers of America et al v Facebook* (EEOC, 2018)).

127 See, eg, *Riddick* (n 126).

128 Jack Gillum and Ariana Tobin, ‘Facebook Won’t Let Employers, Landlords or Lenders Discriminate in Ads Anymore’, *Pro Publica* (online, 19 March 2019) <<https://www.propublica.org/article/facebook-ads-discrimination-settlement-housing-employment-credit>>; Alexia Fernández Campbell, ‘Facebook Allowed Companies to Post Job Ads Only Men Could See. Now That’s Changing’, *Vox* (online, 21 March 2019). <<https://www.vox.com/2019/3/21/18275746/facebook-settles-ad-discrimination-lawsuits>>.

129 *Ibid.*

130 Jeremy B Merrill, ‘Does Facebook Still Sell Discriminatory Ads?’, *The Markup* (online, 25 August 2020) <<https://themarkup.org/ask-the-markup/2020/08/25/does-facebook-still-sell-discriminatory-ads>>.

131 Julia Angwin, ‘Facebook Quietly Ends Racial Ad Profiling’, *The Markup* (online, 29 August 2020) <<https://www.getrevue.co/profile/themarkup/issues/facebook-quietly-ends-racial-ad-profiling-269635>>.

This Part of the article examines whether discrimination by an algorithm can constitute unlawful discrimination under the direct discrimination provisions of Australia's current anti-discrimination laws. Put another way, can an employer be liable under those provisions for the discriminatory decision of an algorithm?

All direct discrimination provisions in Australian law require that a 'person' engage in the discriminatory treatment. For example, the *EOA* provides that direct discrimination occurs 'if a *person* treats, or proposes to treat, a person with an attribute unfavourably because of that attribute'.¹³² But can this requirement be satisfied where the algorithmic model in the AHS has, in effect, made the discriminatory decision without any input from a natural person? Where an employer utilises an AHS to sort, score, rank and automatically cull low-scoring job applicants, has a 'person' made the decision about who is to proceed to the next step in the recruitment process and who is not? If not, can the decision of the algorithm, and therefore liability, be attributed to the employer?

On a strict view, the answer is probably 'no' to these questions. As Burdon and Harpur assert, a 'human decision-maker remains a key aspect in how anti-discrimination laws construct discrimination' in Australia.¹³³ The current legislative framework does not contemplate a situation where the discriminatory decision is made by a non-human actor, such as an algorithm. Unlike a corporation,¹³⁴ there is no legislative provision making an algorithm a 'person' for the purposes of anti-discrimination legislation. Nor, in the absence of a human decision-maker, could it be asserted that the general principles of agency apply to make the decision of the algorithm 'count as' the decision of the employer.¹³⁵

Further, as an AHS is not a legal entity, the current extended liability provisions found in Australian anti-discrimination statutes do not apply to this situation. The attributed liability provisions deem only acts committed by another legal entity (an 'employee' or 'agent'¹³⁶ or, in some statutes, a 'director, employee or agent of a body corporate')¹³⁷ to be the act of an employer (or principal or body corporate). Nor could the employer be liable pursuant to the accessorial liability provisions. These provisions are premised on one person causing, inducing, aiding or permitting another person to do an act that is unlawful.¹³⁸

132 *EOA* (n 40) s 8(1) (emphasis added). See also section 14(a) of the *ADA* (n 40) which provides that 'a person ... discriminates against another person ... on the ground of the age of the ... person if ...'.

133 Burdon and Harpur (n 17) 697. See also Barocas and Selbst (n 14) 699 where they posit that, because the doctrine of disparate treatment (the equivalent in the US of direct discrimination in Australia) 'focuses on human decision makers as discriminators', it is only able to address 'discrimination stemming from human bias'.

134 *Interpretation of Legislation Act 1984* (Vic) s 38.

135 Cf *Christian Youth Camps* (n 123) where Maxwell P of the Court of Appeal held that Parliament is taken to have intended that the general principles of agency should apply where discriminatory conduct by a corporation is alleged: 286–7.

136 See, eg, *EOA* (n 40) s 109.

137 See, eg, *ADA* (n 40) s 57. This provision details how the state of mind of a body corporate may be ascertained through that of the directors, servants or agents of that corporation.

138 See, eg, section 105 of the *SDA* (n 40) which provides that '[a] person who causes, instructs, induces, aids or permits another person to do an act that is unlawful under [this Act] ... shall, for the purposes of this Act, be taken also to have done the act.'

The only way that liability will attach to an employer using a biased AHS under the direct discrimination provisions is to assert that a ‘person’ has engaged in the discriminatory treatment. It is always a challenge in discrimination law to frame the discriminatory conduct and apply legal definitions of discrimination to real-life situations.¹³⁹ With this in mind, the discriminatory conduct on the part of a ‘person’ could be framed as: the employer treats a job applicant with protected attributes unfavourably because of those attributes when it deploys and uses a biased AHS to make recruitment decisions.¹⁴⁰

However, this scenario is markedly different to the usual factual scenarios in claims of direct discrimination. This is because the ‘treatment’ by the alleged discriminator is one step removed from the discriminatory act as it is the algorithm, through its ‘reasoning’ process, and not a person, through a human mental process, that makes the discriminatory decision.¹⁴¹ Further, in this scenario, the employer is unlikely to be aware that a particular job applicant has applied for a job, of their protected attributes and may not be aware that the AHS that has been deployed and utilised in recruitment decisions is biased against members of protected groups. Given the complexity of these systems, the absence of any legal requirement to conduct an equality or human rights impact assessment or audit prior to deployment and the fact that most are purchased from third party vendors, there is a real chance that employers will be unaware that such bias exists.

The authorities assert that possessing an intention or motive to discriminate is not an element of direct discrimination.¹⁴² In this respect, Australia has adopted a different approach to that in the US where discriminatory intention, considered to be synonymous with motive, is critical to proving a complaint of direct discrimination.¹⁴³ But, in the absence of intention or motive, is it still necessary to prove some element of state of mind on the part of the discriminator? If the employer did not have knowledge of the job applicant’s protected attribute or that the AHS was biased, has the causal nexus been proved, that is, did the employer treat the job applicant

139 Belinda Smith, ‘From *Wardley* to *Purvis*: How Far has Australian Anti-Discrimination Law Come in 30 Years?’ (2008) 21(3) *Australian Journal of Labour Law* 3 <<https://doi.org/10.2139/ssrn.1005528>>. Despite jurisprudence to the contrary (see, eg, *Waters v Public Transport Corporation* (1991) 173 CLR 349, 392–3 (Dawson and Toohey JJ) (*‘Waters’*)), the concepts of direct and indirect discrimination are better understood as overlapping and not mutually exclusive as fact situations can usually be characterised either way.

140 The attributed liability provisions apply when the employer acts through an employee: see above n 123.

141 A similar issue, albeit in an unusual case, arose in a recent case before the Full Federal Court. In *Pintarich v Deputy Commissioner of Taxation* (2018) 262 FCR 41, the Court considered whether an automated letter sent by a Deputy Commissioner to a taxpayer constituted a decision for the purpose of the *Taxation Administration Act 1953* (Cth). Although the generation of the letter did not involve an algorithm but rather a ‘template bulk issue letter’, the Court found that the letter did not constitute a ‘decision’ under the Act because it did not involve a human ‘mental process’: [143]–[152] (Moshinsky and Derrington JJ).

142 *Australian Iron & Steel Pty Ltd v Banovic* (1989) 168 CLR 165, 176–7 (Deane and Gaudron JJ) (*‘Banovic’*); *Waters* (n 139) 359 (Mason CJ and Gaudron J). See also section 8(2) of the *EOA* (n 40) which provides: ‘In determining whether a person directly discriminates it is irrelevant whether or not that person is aware of the discrimination or considers the treatment to be unfavourable’.

143 *Civil Rights Act of 1991* Pub L 102-166, 105 Stat 1071 § 107 (1991). See also *Watson v Fort Worth Bank and Trust*, 487 US 977 (1998).

unfavourably ‘because of’¹⁴⁴ the protected attribute?¹⁴⁵ Unfortunately, on the state of the law in Australia, these questions cannot be answered with any certainty.

In the most recent High Court authority, *Purvis v New South Wales* (‘*Purvis*’),¹⁴⁶ two different tests of causation¹⁴⁷ in direct discrimination complaints were applied: the ‘true basis’ test and the ‘why?’ test. The majority of the Court,¹⁴⁸ applied the ‘true basis’ test of causation. This approach was first adopted by Deane and Gaudron JJ in *Australian Iron & Steel Pty Ltd v Banovic*.¹⁴⁹ It involves asking what the underlying reason for the impugned decision is and acknowledges that ‘genuinely assigned reasons may in fact mask the true basis of a decision’.¹⁵⁰ In *Purvis*, the majority of the Court accepted that the ‘true basis’ for the principal’s decision to suspend and later expel the complainant pupil from his school was the ‘violent conduct of the pupil, and his concern for the safety of other pupils and staff members’, rather than the complainant’s disability.¹⁵¹

In this decision, the issue of knowledge on the part of the discriminator was considered by Kirby and McHugh JJ. Although McHugh J in *Waters v Public Transport Corporation*¹⁵² had asserted that ‘the status ... of the victim must be at least one of the factors which moved the discriminator to act as he or she did’,¹⁵³ their Honours recharacterised these comments. They found that McHugh J’s comments did not suggest a real difference of approach with the majority in that case,¹⁵⁴ who found that it was not necessary for the conduct of the alleged discriminator to be ‘actuated’ by knowledge of the person’s protected attribute.¹⁵⁵

The minority of the Court in *Purvis*,¹⁵⁶ in obiter, adopted the ‘why?’ approach to causation, where the central question is:

[W]hy was the aggrieved person treated as he or she was? If the aggrieved person was treated less favourably was it ‘because of’, ‘by reason of’, that person’s [protected attribute]? Motive, purpose and effect may all bear on that question. But it would be a mistake to treat those words as substitutes for the statutory expressions ‘because of’.¹⁵⁷

144 Different terminology is used between jurisdictions in Australia to describe this causal nexus, with some Acts requiring treatment to be ‘because of’ an attribute and others ‘on the ground of’, ‘by reason of’ or ‘on the basis of’. These terms are regarded as having the same meaning: Gaze and Smith (n 40) 115.

145 It is noted that, in some jurisdictions, awareness of wrongdoing or the discrimination is not relevant to a determination of whether discrimination has occurred. See, eg, *EOA* (n 40) s 8(2)(a). However, it is an open question whether awareness of an *individual act* of discrimination is the same as awareness that an algorithmic system has the potential to *systemically discriminate* against members of protected groups.

146 (2003) 217 CLR 92 (‘*Purvis*’).

147 Gaze and Smith (n 41) assert that causation is the wrong term and the ‘issue is best considered in terms of what types of reasons are unlawful, and how the reason for an action can be proved’: at 115.

148 That is, Gleeson CJ, McHugh, Kirby and Callinan JJ.

149 *Banovic* (n 142).

150 *Purvis* (n 146) 142, [157] (Kirby and McHugh JJ).

151 *Ibid* 102–3, [14] (Gleeson CJ).

152 (1991) 173 CLR 349.

153 *Ibid* 401.

154 *Purvis* (n 146) 142 (Kirby and McHugh JJ).

155 *Banovic* (n 142) 176–7 (Deane and Gaudron JJ).

156 Gummow, Hayne and Heydon JJ.

157 *Purvis* (n 146) 163 [236] (Gummow, Hayne and Heydon JJ).

Although only the minority of the Court adopted this test in *Purvis*, it is widely applied as the test for causation in direct discrimination cases.¹⁵⁸

It has been suggested that the decision in *Purvis* has left the law of causation in a ‘state of uncertainty’.¹⁵⁹ Further, the question of whether it is necessary to prove a precise state of mind on the part of a discriminator such as knowledge (as opposed to intention or motive) remains unresolved.¹⁶⁰ In cases involving algorithmic decision systems, this lack of clarity makes these tests difficult to apply and produces contradictory results.

In the scenario discussed above, imagine that an employer has direct knowledge that the AHS it has deployed to make automated decisions in the screening process operates with bias against a protected group. An application of the ‘true basis’ test of causation could mean that the employer has not engaged in unlawful direct discrimination if they could establish that the ‘real reason’ for the deployment of this system was not to discriminate against protected groups but instead to save time and money and increase efficiency. If the employer can establish that it was unaware, or perhaps even wilfully blind, that the AHS operated with bias against protected groups, the complainant’s case may be even more difficult to prove. For this reason, the ‘true basis’ approach has been criticised for allowing ‘by judicial invention rather than by legislative action, an excuse or defence of “good motive”’.¹⁶¹

A different, and arguably preferable, result is produced when the ‘why?’ approach is adopted. Unlawful direct discrimination would be established whether or not the employer has knowledge that the AHS is biased against members of protected groups. However, applying this test involves somewhat circular reasoning. The reason the aggrieved person was treated unfavourably was because of their protected attribute. The treatment is unfavourable because of a job applicant’s protected attributes because the AHS is biased against job applicants possessing those attributes. This situation is broadly analogous to one involving unconscious bias where a decision-maker is unaware of their motivating factors or reliance on stereotypes. A case of unconscious bias would also satisfy this ‘why?’ test, although this would be very difficult to prove.¹⁶²

This jurisprudential conflict over the applicable test of causation and necessary state of mind of the alleged discriminator needs to be resolved. However, this will not go far enough. The preceding analysis has shown that discrimination by an AHS will not be unlawful unless the discriminatory treatment is framed as that of a ‘person’. This is a difficult ‘fit’, particularly where the employer is not aware of the biased operation of the algorithmic system. Legislative amendment is therefore required to provide clarity and certainty regarding questions of employer liability

158 See, eg, *Sklavos v Australasian College of Dermatologists* [2017] FCAFC 128, [27] (‘*Sklavos*’), and *Taniela v Australian Christian College Moreton Ltd* [2020] QCAT 249. See also Rees, Rice and Allen (n 41) 114.

159 Rees, Rice and Allen (n 41) 115.

160 *Ibid* 109.

161 *Ibid* 111.

162 Gaze and Smith (n 40) 117.

for the operation of AHSs. In the absence of such legislation, questions will remain as to the lawfulness of discrimination by algorithm.

Detailed consideration of the exact legislative mechanism to assign liability to an employer is outside the scope of this article. The AHRC has proposed that federal legislation be introduced which creates ‘a rebuttable presumption 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 artificial intelligence’.¹⁶³ This proposal is broadly supported. However, it is preferable that any legislative amendment regarding liability for discrimination by algorithm be specifically attuned to anti-discrimination law. One obvious option is to extend the attributed and accessorial liability provisions of existing anti-discrimination laws. An express statutory rule of attribution could be employed like that in section 495A of the *Migration Act 1958* (Cth), where decisions made by the operation of a computer program deployed by a ‘person’ or an employer are ‘taken to be’ decisions of that ‘person’ or employer.¹⁶⁴ Any formulation should ensure that employers have the defence of ‘reasonable preventative action’ open to them.¹⁶⁵ In the context of an AHS, this could include obtaining an impact statement or auditing the system for unfavourable treatment and bias or ensuring that such an audit has been conducted by the third party vendor. Once the employer has primary liability for the algorithmic decision, the third party vendor of such a system may also be liable under the accessorial liability provisions for causing or assisting the employer to commit acts of unlawful discrimination. This will be an important mechanism to incentivise those vendors to conduct regular audits of their AHSs and to ensure they comply with local laws and utilise training datasets that are geographically appropriate.

C Discrimination by Proxy

AHSs, use an unprecedented number and variety of ‘features’ to assess job applicants. Many of these features appear to be neutral but may map onto and become ‘proxies’ for protected attributes.¹⁶⁶ Even if proscribed criteria are removed from the dataset, proxy discrimination may occur as membership of the protected class is ‘redundantly encoded’, that is, embedded in other data.¹⁶⁷

163 See recommendation 11 in *AHRC HR and Technology Final Report* (n 16) 78, 194.

164 Section 495A of the *Migration Act 1958* (Cth) provides that decisions made by ‘the operation of a computer program’ are ‘taken’ to be decisions of the Minister. This is also similar to attribution provisions in respect of corporations in federal legislation, such as section 57(2) of the *ADA* (n 40), which provide that ‘[a]ny conduct engaged in on behalf of a body corporate by a director, employee or agent ... within the scope of his or her ... authority is taken to have been engaged in also by the body corporate’. See also *DDA* (n 40) s 123.

165 All of the attributed liability provisions permit an employer to avoid liability for an act of unlawful discrimination committed by an employee if they can demonstrate that they took reasonable preventative action to avoid such acts of discrimination. See, eg, *EOA* (n 40) s 110.

166 Mittelstadt et al (n 53) 8.

167 Barocas and Selbst (n 14) 691.

The proxies utilised by AHSs may be different to those used historically by human decision-makers. They may not be easy to predict or detect,¹⁶⁸ as they are not based on stereotypical assumptions or generalisations but rather statistical correlation. These problems are pronounced if utilising an ‘unsupervised’ learning model,¹⁶⁹ as those systems may find unintuitive connections and patterns in those features.¹⁷⁰ For example, one CV parsing system found two features to be most indicative of job performance: that an applicant’s name was Jared and that they played high school lacrosse.¹⁷¹ Indeed, Burdon and Harpur argue that the randomness of these features or ‘informational attributes’ used by AHSs, which may include location data, device data and sociometric measurements, produce effects that are ‘not automatically covered by anti-discrimination laws because they do not habitually involve decisions regarding a protected attribute’.¹⁷²

As discussed in Part III, the features video interviewing assessment systems rely include on tone and facial expressions, eye contact and word use to assess job candidates. It is not difficult to imagine that a candidate, for example, with a disability which impacts speech or facial expressions or the ability to make eye contact, might be treated unfavourably by such as system. They may receive a low tier score and/or be subject to automatic exclusion. So too may candidates with non-native accents.¹⁷³ CV parsing tools, on the other hand, rely on a different assortment of features. These may include where a person lives, the university they attended, or the key words used in a CV. However, as with the female candidates who attended all-women’s colleges discussed in Part III, these too may lead to unfavourable outcomes for group members.

The use of proxy features to make decisions that are unfavourable to the interests of job applicants may constitute direct discrimination¹⁷⁴ where they fall within the ‘characteristics extension’. The characteristics extension expands the operation of anti-discrimination legislation. It makes it unlawful to discriminate against a person because of a ‘characteristic’ which is generally possessed by, or imputed to, people who have a protected attribute.¹⁷⁵ The policy rationale for this extension is to ensure that responsibility under anti-discrimination legislation is

168 Mittelstadt et al (n 53) 8.

169 Unsupervised machine learning algorithms do not try to predict a ‘target variable’ but merely learn patterns from the training data: ‘Algorithmic Accountability Policy Toolkit’ (n 3) 28.

170 Wachter, Mittelstadt and Russell ‘Why Fairness Cannot Be Automated’ (n 11) 5. See also Burdon and Harpur (n 17) 696–9.

171 Dave Gershgorn, ‘Companies Are on the Hook if Their Hiring Algorithms Are Biased’, *Quartz* (online, 24 October 2018) <<https://qz.com/1427621/companies-are-on-the-hook-if-their-hiring-algorithms-are-biased/>>.

172 Burdon and Harpur (n 17) 696.

173 Although vendors of video assessment systems may have measures in place to mitigate bias on the basis of protected attributes, this does not extend to proxy attributes. For this reason, Hirevue has identified the relationship between accents and competency scores assessed by video as an area for further research: ORCAA (n 88) 3.

174 Subject to the discussion above under Part IV(B) of this article.

175 See, eg, *ADA* (n 40) sections 14(b)(ii)–(iii) which provide that conduct will be unlawful if it is done because of ‘a characteristic that appertains generally to persons of the age of the aggrieved person’ or ‘a characteristic that is generally imputed to persons of the age of the aggrieved person’.

not evaded by ‘using such characteristics as “proxies” for discriminating on the basic grounds covered by the legislation’¹⁷⁶ and to discourage people from relying on stereotypes or immutable attributes.¹⁷⁷

Despite being available in most Federal and State and Territory Acts,¹⁷⁸ neither the *Racial Discrimination Act 1975* (Cth) (‘*RDA*’) nor the *Disability Discrimination Act 1992* (Cth) (‘*DDA*’) contain a general characteristics extension to direct discrimination provisions. Therefore, in the video interview examples above, unfavourable treatment on the ground of speaking English with a non-native accent is unlikely to constitute racial discrimination under the Federal *RDA*.¹⁷⁹ Discrimination on the basis of a complainant’s accent will only be established under the *RDA* where the accent can be ‘substituted’ or ‘directly linked’ to that of race, colour or national origin.¹⁸⁰ However, this unfavourable treatment is likely to constitute discrimination on the basis of race under State and Territory legislation.¹⁸¹ Although the *DDA* does not have a general characteristics extension, since 2009, the definition of ‘disability’ has included ‘behaviour that is a symptom or manifestation of the disability’.¹⁸² Facial expressions or a lack of eye contact are arguably manifestations of a disability. It is less clear if immutable attributes such as these could also be classified as a ‘behaviour’. Therefore, it is unclear whether discrimination on this basis could constitute unlawful direct discrimination.¹⁸³

In the case of CV parsing systems, it seems likely that many of the features used to assess job applicants will be too general or remote to fall within the characteristics extension. An argument could not be easily mounted, for example, that living in a particular postcode, attending a second-tier university, or using particular words in a CV appertains generally, or can be imputed, to a person with a protected attribute. Instead, persons who have been the victim of discrimination on the basis of proxies will need to rely on the indirect discrimination provisions of anti-discrimination legislation if a sufficient connection or correlation between the feature and a protected attribute can be established.¹⁸⁴ As Wachter has highlighted, algorithmic decision systems create new groups defined by features that do not map onto protected attributes but who ‘experience discrimination with comparable harmful effects via the same mechanisms as protected groups’.¹⁸⁵ These new patterns

176 *Purvis* (n 146) 134–5, [130] (McHugh and Kirby JJ).

177 *Ibid.*

178 See, eg, *SDA* (n 40) ss 5(1), 6(1), 7(1). A table of all characteristics extension provisions in Australian anti-discrimination legislation is found in Gaze and Smith (n 40) 296, tbl 1.

179 See Rees, Rice and Allen (n 41) 47.

180 See *Philip v New South Wales* [2011] FMCA 308, [225]. See also ‘Accents’, *Australian Human Rights Commission* (Web Page) <<https://humanrights.gov.au/quick-guide/11907>>.

181 See, eg, *Perera v Commissioner of Corrective Services* [2007] NSWADT 115, [111]–[113]; *Chew v Director General of Department of Education and Training* [2006] WASAT 248, [58].

182 *DDA* (n 40) s 4.

183 This analysis does not consider direct discrimination by not making reasonable adjustments: *ibid* s 5(2). As to whether such treatment could constitute indirect discrimination, see the discussion in Part D of this article.

184 Whether the use of proxy features will violate indirect discrimination provisions is considered in Part D of this article.

185 Sandra Wachter, ‘Affinity Profiling and Discrimination by Association in Online Behavioural Advertising’ (2020) 35(2) *Berkeley Technology Law Journal* 367, 414 <<https://doi.org/10.15779/Z38JS9H82M>>.

of disparity may ignite fresh debates about the scope of anti-discrimination law.¹⁸⁶ One such group attribute is social class. It is arguable that CV parsing systems use the university attended by a job applicant or that applicant's vocabulary as proxies for social class or socio-economic status. However, this is not a protected attribute in any anti-discrimination statute in Australia.¹⁸⁷

D Systemic Discrimination

The concept of 'indirect' discrimination, in the US referred to as 'disparate impact', would appear to be well-suited to the task of regulating algorithmic discrimination.¹⁸⁸ Indirect discrimination is concerned with 'practices that are fair in form but discriminatory in operation',¹⁸⁹ and is widely understood to be directed at tackling structural discrimination and achieving substantial equality.¹⁹⁰ As equality of outcome, not treatment, is the focus, many of the problems that arise from the lack of transparency in the operation of AHSs are avoided. Further, given algorithmic discrimination is always systemic discrimination, that is, discrimination at a group rather than an individual level, this approach appears to be attuned to this particular form of harm.

However, in practice, the operation of indirect discrimination provisions in Australia is fraught with uncertainty and their ability to provide effective remedies for complainants is questionable. First, the indirect discrimination provisions differ in each Australian jurisdiction and were drafted at different times, often without apparent regard for existing provisions or caselaw.¹⁹¹ Second, the High Court of Australia has considered only four cases of indirect discrimination¹⁹² and jurisprudence which has emerged about key provisions lacks consistency and clarity.¹⁹³ There is also a disparity of views both within the High Court itself and in the lower courts and tribunals. This has led to the widely held view that claims of indirect discrimination are 'extremely difficult

186 See, eg, Barocas and Selbst (n 14).

187 Although complaints in relation to discrimination on the basis of 'social origin' can be commenced pursuant to the *Australian Human Rights Commission Act 1986* (Cth) Part II and section 351 of the *Fair Work Act 2009* (Cth). It was also one of the protected attributes when connected with 'work or work-related areas' in the Human Rights and Anti-Discrimination Bill 2012 (n 113) section 17(1)(r). See also Angelo Capuano, 'Social Origin: The Misunderstood, Multifaceted and Symmetrical Attribute' (Working Paper, 27 October 2019). See also section 7 of the *DA* (n 119) which includes 'accommodation status' (defined to include homelessness) and 'employment status' as protected attributes.

188 International scholars posit that it is indirect discrimination provisions which are apposite to cases of algorithmic discrimination: Barocas and Selbst (n 14); Wachter, Mittelstadt and Russell (n 11) 19–20.

189 *Griggs v Duke Power Co*, 401 US 424, 431 (1971) ('*Griggs*').

190 Cf *Sklavos* (n 158); Michael F Foran, 'Discrimination as an Individual Wrong' (2019) 39 (Winter) *Oxford Journal of Legal Studies* 901 <<https://doi.org/10.1093/ojls/gqz026>>.

191 Rees, Rice and Allen (n 41) 144.

192 *Banovic* (n 142); *Waters* (n 139); *New South Wales v Amery* (2006) 230 CLR 174; *Lyons v Queensland* (2016) 259 CLR 518.

193 Rees, Rice and Allen (n 41) 143. See also Alice Taylor, 'The Conflicting Purposes of Australian Anti-Discrimination Law' (2019) 42(1) *University of New South Wales Law Journal* 188 <<https://doi.org/10.53637/TYBL5821>>.

to prove'.¹⁹⁴ In part, these difficulties stem from the fact that '[t]he concept of indirect discrimination delegates significant responsibility to courts and tribunals to make policy decisions of broad public importance in the absence of any legislative guidance about relevant considerations'.¹⁹⁵ This is exacerbated by the potential of these provisions to achieve radical outcomes: a redistribution of wealth and opportunities from privileged groups to those who have been historically disadvantaged.¹⁹⁶

With these difficulties in mind, we turn to consider whether the indirect discrimination provisions in Australian law protect against, and render unlawful, discrimination by an employer using an AHS. The 'modern' formulation of indirect discrimination found in Tasmania, Victoria, and under the *Age Discrimination Act 2004* (Cth) and *Sex Discrimination Act 1984* (Cth), is used in this analysis as it is widely accepted as the preferable one.¹⁹⁷ Under this formulation, an employer using an AHS engages in unlawful indirect discrimination where: (i) it imposes a 'requirement, condition or practice'; (ii) 'that has, or is likely to have, the effect of disadvantaging persons with [a protected] attribute' and (iii) 'that is not reasonable'.¹⁹⁸ In this approach, there is no need to prove that a person with a protected attribute does not comply or is unable to comply with the requirement or condition.

In this analysis, two scenarios are considered. In the first, an employer uses a CV parsing system that penalises job applicants with a broken employment history. In the second, an employer uses a video interview assessment system that is not able to recognise expressions of 'enthusiasm' on the faces of people of colour. In both of these examples, it is arguable that there is requirement, condition or practice imposed on all job applicants including those with protected attributes. That requirement, condition or practice could be framed with different levels of specificity¹⁹⁹ as:

1. a requirement that, in order to be considered for the position, all job applicants submit to CV parsing or a video interview assessment system;
2. a requirement that, in order to progress to the next stage of the recruitment process, a job applicant achieve a score above a certain level as determined by the CV parsing or video interview assessment system;²⁰⁰ or

194 Burdon and Harpur (n 17) 698. See also Margaret Thornton, 'Disabling Discrimination Legislation: The High Court and Judicial Activism' (2009) 15(1) *Australian Journal of Human Rights* 1 <<https://doi.org/10.1080/1323238X.2009.11910859>>.

195 Rees, Rice and Allen (n 41) 143.

196 Michael Connolly, *Townshend-Smith on Discrimination Law: Text, Cases and Materials* (Cavendish, 2nd ed, 2004) 238.

197 See, eg, Discrimination Law Experts Group (n 113) 8–9; Human Rights Anti-Discrimination Bill (n 113) s 19(3).

198 *EOA* (n 40) s 9(1)(a).

199 See above n 139 regarding the multitude of ways in which a complaint of discrimination may be framed and the overlap between concepts of direct and indirect discrimination.

200 The framing of the 'requirement' or 'condition' in this way was accepted by the Full Court of the Federal Court in *Nojin v Commonwealth of Australia* (2012) 208 FCR 1. This case concerned the use by an employer of an assessment tool to calculate the wages of two intellectually disabled workers. Although this assessment tool was not driven by an algorithm but rather in-person interviews and skills assessments,

3. a requirement that an application have an unbroken work history or express enthusiasm in a way that an AHS can understand or interpret.

In all three examples, it is arguable that the second element is also made out as the effect of each requirement is to disadvantage persons with a protected attribute, namely women and people of colour. Women submitting to the CV parsing system will be assessed as suitable for the position at a lower rate than men as they are more likely to have a broken employment history due to breaks in paid employment to care for children and/or elderly parents. In this example, a broken employment history is a clear proxy for women. People of colour will also be disadvantaged as, unlike their white counterparts, the AHS is not able to accurately interpret and classify their facial action units. As the AHS is more likely to interpret these negatively, they will be assigned a lower score.

In order to prove this element, it seems likely complainants would be required to obtain statistical or technical evidence. Wachter, Mittelstadt and Russell assert that ‘automated discrimination may only be observable at a statistical level’ and therefore argue that statistical evidence may need to become the default option in these cases.²⁰¹ However, in Australia, there has been a general judicial reluctance to mandate the obtaining of expert statistical evidence by complainants and we lack relevant rules and standards. In particular, unlike the ‘four-fifths rule’ in the US,²⁰² neither legislation nor the authorities provide any statistical threshold or guide as to when ‘disadvantage’ has occurred.

The last and most difficult element to consider is that of ‘reasonableness’. This element permits a justification for the requirement despite the disparate impact. Some Australian statutes provide an inclusive list of matters to be taken into account when deciding if a requirement is reasonable in the circumstances.²⁰³ The authorities suggest that the balance lies somewhere between the more onerous test of business necessity which applies in the US²⁰⁴ and that of convenience.²⁰⁵ The test is said to be objective and requires the weighing of ‘the nature and extent of the discriminatory effect, on the one hand, against the reasons advanced in favour

it scored workers and the ‘higher’ the score achieved, the higher a worker’s wages. The evidence was that intellectually disabled workers were disadvantaged by this tool in comparison to other disabled workers.

The Court rejected the characterisation by the employer that the only ‘requirement’ or ‘condition’ imposed on the workers was that they submit to the assessment tool, a requirement with which they were able to comply. Instead, the Full Court found that it was open to the workers to argue, and accepted, that the ‘requirement’ or ‘condition’ to which they were subjected was that wage increases could only be achieved by obtaining a higher score on the assessment tool: at 43–4 [121]–[124] (Buchanan J), 57–8 [186]–[189] (Flick J), 70–72 [237]–[242] (Katzmann J).

201 Wachter, Mittelstadt and Russell ‘Why Fairness Cannot Be Automated’ (n 11) 16.

202 This ‘rule of thumb’ provides guidance as to when a disparate impact case may be brought against an employer – if the selection rate for one protected group is less than four-fifths of that of the group with the highest selection rate, there may be discrimination on the part of the employer: *Uniform Guidelines on Employment Selection Procedures*, 29 CFR § 1607.4 (2016) (‘*US Uniform Guidelines on Employment Selection*’).

203 See, eg, *EOA* (n 40) s 9(3).

204 *Griggs* (n 189) 431–6.

205 *Secretary, Department of Foreign Affairs and Trade v Styles* (1989) 23 FCR 251, 263 (Bowen CJ and Gummow J).

of the requirement or condition on the other'.²⁰⁶ Relevant factors usually include the reasons advanced in favour of the requirement or condition, the consequences of a failure to comply with the requirement, the financial position of the person imposing the requirement, and the availability and cost of implementing alternative methods of achieving the alleged discriminator's objectives. It is argued that the potential scale of the harm is a factor that must always be considered in the context of algorithmic discrimination. As Ajunwa states: 'the impact of one biased human manager is constrained in comparison to the potential adverse reach of algorithms that could be used to exclude millions of job applicants from viewing a job advertisement or to sort thousands of resumes'.²⁰⁷

At first glance, the requirement in example 1 appears to be reasonable in the circumstances. In this example, determination of the question of reasonableness involves a consideration of the reasonableness of the use by an employer of a form of automation in the screening part of the recruitment process. It is arguably reasonable for an employer to require a job applicant to submit to an AHS, such as a CV parsing system or a video interview assessment, when the employer is, for example, inundated with large numbers of CVs from job applicants and alternatives for processing these applications are not cost effective or timely. However, there are also strong countervailing arguments including the nature and extent of the discriminatory impact – that is, the potential of an AHS operating with bias to lock large numbers of women and people of colour out of the workforce. In addition, important questions will arise for consideration by courts and tribunals such as: (i) how far do employers have to go to undertake costly auditing of AHSs for bias before they may reasonably be used?²⁰⁸ (ii) can employers rely on representations from third party vendors as to anti-bias mitigation measures?

The requirements in examples 2 and 3 above are potentially more complicated. This is because the reasonableness test is applied not to the use by the employer of the AHS generally but to the use of a specific AHS with particular attributes and priorities. In example 2, it will only be reasonable to require job applicants to achieve a score above a certain level when all aspects of the algorithmic model, including the choice of features to be captured and the target variable, such as a 'high performing employee' or 'best fit', are reasonable.²⁰⁹ In the case of video interviewing assessments, the lack of empirical evidence to support the existence of a causal link between features such as tone and facial expressions and workplace suitability or performance²¹⁰ provide a persuasive argument that the requirement is unreasonable. Example 3 surfaces particular aspects of the algorithmic model and similar arguments would apply. In the US, there has been a call for the law to

206 Ibid; *Waters* (n 139) 395–6 (Dawson and Toohey JJ), 383 (Deane J). Applied in *Catholic Education Office v Clarke* (2004) 138 FCR 121.

207 Ajunwa (n 14) 1679.

208 The AHRC recently provided some practical guidance for businesses that develop, design and deploy AI systems regarding the lawful and responsible use of these systems: see 'AHRC Algorithmic Bias Technical Paper' (n 17) including the 'Responsible Business Use of AI and Data' toolkit: at 55.

209 See, eg, Kleinberg et al (n 14) 146–8.

210 Bogen and Rieke (n 13) 38.

mandate that the criteria used in algorithmic hiring systems have some probative value for determining fitness to perform required job duties.²¹¹

In addition, in examples 1 and 3, courts and tribunals will be called to adjudicate complex and technical questions pertaining to the algorithmic system such as: (i) should the employer have chosen different features or a target variable that is less likely to result in disadvantage for protected groups? (ii) was the choice of training datasets (with all of their limitations) reasonable when compared to other available datasets? There will also be broad policy questions for determination such as: should an employer to bear the financial burden of modifying an AHS (that has, in all likelihood, been purchased off-the shelf from a third party vendor) by acquiring more representative data or modifying the model to reduce its discriminatory impact?

It is widely accepted that there is an urgent need for reform of the indirect discrimination provisions in Australian legislation.²¹² The suitability of these provisions for, and the significant role they could play in, regulating algorithmic discrimination strengthens this call for reform. As the above analysis has shown, the utility of the indirect discrimination provisions when there is discrimination by an employer using an AHS is dependent on judicial understandings of complex socio-technical systems and engagement with difficult questions of public policy. Therefore, as part of this reform, consideration should also be given to the introduction of guidelines²¹³ or binding standards, such as those issued under the *DDA*,²¹⁴ to assist with interpreting the indirect discrimination provisions and promote employer compliance.²¹⁵ These are available in in the employment context the US²¹⁶ and in the European Union.²¹⁷ Any such guidelines or standards issued with reforms should specifically address discriminatory algorithms and assist with

211 Ajunwa (n 14) 1718.

212 See, eg, Rees, Rice and Allen (n 40) who assert that '[t]he concept of indirect discrimination needs to be re-worked if it is to play a meaningful role in Australian anti-discrimination law': at 53; Discrimination Law Experts Group (n 113) 8–11. This must include the development of a 'consistent conceptual framework' regarding the purpose or underlying rationale of anti-discrimination law: Taylor (n 193) 209.

213 One of the functions of the AHRC is to produce guidelines for employers and other organisations to assist with compliance under federal anti-discrimination laws: *RDA* (n 40) s 20(d), *SDA* (n 40) s 48(1)(ga), *DDA* (n 40) s 67(1)(k), *ADA* (n 40) s 53(1)(f). Such guidelines are non-binding: see, eg, *Richardson v Oracle Corporation Australia Pty Ltd* [2013] FCA 102. The AHRC has produced guidelines regarding the prevention of discrimination in recruitment, but those guidelines do not refer to AHSs and are in need of updating: Australian Human Rights Commission, 'A Step-By-Step Guide to Preventing Discrimination in Recruitment' (Guidelines, November 2014).

214 Standards are legislative instruments, made by the Attorney-General and reviewed every five years: AHRC Priorities for Reform (n 113) 13.

215 The AHRC's Priorities for Federal Discrimination Law Reform project is giving consideration to how existing compliance measures under federal discrimination law can be improved, and whether any additional measures would assist to provide greater certainty and compliance with those laws: see AHRC Priorities for Reform (n 113) 14.

216 *US Uniform Guidelines on Employment Selection* (n 202).

217 The EU has four non-discrimination directives including two specifically related to employment: see *Council Directive 2000/78/EC of 27 November 2000 Establishing a General Framework for Equal Treatment in Employment and Occupation* [2000] OJ L 303/16; *Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006 on the Implementation of the Principle of Equal Opportunities and Equal Treatment of Men and Women in Matters of Employment and Occupation (Recast)* [2006] OJEU L 204/23. See also *Council Directive 2000/43/EC of 29 June 2000 Implementing the Principle of Equal Treatment between Persons Irrespective of Racial or Ethnic Origin* [2000] OJ L 180/22; *Council Directive*

the conversion of ‘general legal principles into measurable, outcome-focused requirements’.²¹⁸ For example, by delineating statistical thresholds for or other measures of disparate impact or disadvantage.

E Problems of Proof

Although outside the scope of this article, it is worth highlighting the significant obstacles faced by complainants in proving discrimination by an employer using an AHS. With the exception of the reasonableness requirement in a claim of indirect discrimination in some jurisdictions,²¹⁹ the complainant bears the onus of proof in discrimination actions.²²⁰ However, currently, there is no legislative requirement that a job applicant be provided with notice or an explanation of the operation of an AHS in the recruitment or hiring process. It will therefore be inherently difficult for a complainant to assess whether wrongdoing has occurred and marshal evidence to prove discriminatory conduct.

Further, in cases of direct and indirect discrimination by algorithmic decision systems, complainants will need to obtain statistical or technical evidence to prove their claim. For example, in direct discrimination actions, such expert evidence will be required to prove the causal nexus between the unfavourable treatment of a job applicant and a protected attribute. In indirect discrimination actions, statistical evidence may be needed to establish the requirement, condition or practice and to prove its disadvantageous effect and lack of reasonableness. This evidence will be costly and time consuming to obtain and arguably acts as a deterrent to complainants.²²¹ These difficulties for complainants are exacerbated by the fact that there are no accepted or standardised statistical techniques or methods for determining whether an algorithm has discriminated against protected groups.²²² There is therefore an urgent need for evidentiary standards identifying which types of statistical steps, tests or counterfactuals should be run to unearth discrimination by these systems.²²³

2004/113/EC of 13 December 2004 Implementing the Principle of Equal Treatment Between Men and Women in the Access to and Supply of Goods and Services [2004] OJEU L 373/37.

218 AHRC Priorities for Reform (n 113) 14.

219 The following Acts place the burden on the person who imposes the requirement, condition or practice to prove that it is reasonable in the circumstances: *SDA* (n 40) ss 7B, 7C; *DDA* (n 40) s 6(4); *ADA* (n 40) s 15(2); *EOA* (n 40) s 9(2), *DA* (n 119) ss 8(4), 70; and *Anti-Discrimination Act 1991* (Qld) s 205.

220 This has been the subject of extensive criticism, see, eg, Dominique Allen, ‘Reducing the Burden of Proving Discrimination in Australia’ (2009) 31(4) *Sydney Law Review* 579.

221 See, eg, comments in *Jordan v North Coast Area Health Service [No 2]* [2005] NSWADT 258.

222 See, eg, Sandra Wachter, Brett Mittelstadt and Chris Russell, ‘Bias Preservation in Machine Learning: The Legality of Fairness Metrics Under EU Non-Discrimination Law’ (2021) 123(3) *West Virginia Law Review* 735. Wachter, Mittelstadt and Russell identified two groups of these fairness metrics: ‘bias preserving’ metrics, in the sense that they use the status quo as a baseline and can therefore reproduce historical inequalities, and ‘bias transforming’ fairness metrics, as they can identify and thereby provide a starting point for addressing structural inequalities: at 761. See also ‘AHRC Algorithmic Bias Technical Paper’ (n 17) 21, which examined the wide range of ‘fairness metrics’ currently in use as indicators of algorithmic bias or discrimination.

223 See, eg, Wachter, Mittelstadt and Russell ‘Why Fairness Cannot Be Automated’ (n 11) 6; Sandra Wachter, Brent Mittelstadt and Chris Russell, ‘Counterfactual Explanations Without Opening the Black Box:

V CONCLUSION

This article has shown how new forms of technology-facilitated discrimination have exposed many old problems with Australia's anti-discrimination laws. Inconsistent judicial interpretation of key provisions governing direct discrimination, the delegation by indirect discrimination provisions of significant responsibility to courts and tribunals to make decisions regarding public and social policy and an 'unnecessary level of difference and complexity' between federal, state and territory laws²²⁴ throw into doubt the law's ability to respond adequately to discrimination by AHSs.

These problems are compounded by the identified gaps in these laws when applied to algorithmic decision-making by an AHS. Legislative amendment of anti-discrimination laws is urgently required to resolve questions of lawfulness and employer liability for decisions made by AHSs. Further, consideration should be given to the introduction of guidelines or binding standards specifically addressing algorithmic decision systems and issues such as: (i) appropriate thresholds for or measures of disparate impact or disadvantage when there is discrimination by algorithm and (ii) which types of statistical steps or counterfactuals should be run to unearth forms of direct and indirect algorithmic discrimination. Finally, and more broadly, we must examine how the law should approach new forms of inequality made possible by technology including discrimination against new groups defined by features, including social class, that do not map onto protected attributes.

As AHSs proliferate and increasingly dominate recruitment and hiring decisions, it is essential that the law provide protection against discrimination to job seekers. Those who are most exposed to discrimination are applicants for low wage and insecure positions – some of society's most disadvantaged and vulnerable members. Australia's anti-discrimination laws are long overdue for reform. What has emerged from this analysis is a need for new legislative provisions specifically tailored to employers' use of discriminatory algorithms. Without reform, the ability of the law to regulate AHSs and other emerging technologies which employ algorithms is limited.

VI POSTSCRIPT

This article is current as at the date it was written and accepted for publication in May 2021. Since that time, the International Standards Organisation has published a Technical Standard describing measurement techniques and methods for assessing algorithmic bias in AI decision systems.²²⁵ All AI system lifecycle

Automated Decisions and the GDPR' (2017) 31(2) *Harvard Journal of Law & Technology* 841 <<https://doi.org/10.2139/ssrn.3063289>>. These tests may need to be different for each algorithmic decision system.

224 AHRC Priorities for Reform (n 113) 8.

225 International Standards Organisation, 'Artificial Intelligence (AI): Bias in AI Systems and AI Aided Decision Making' (Technical Standard ISO/IEC TR 24027:2021, November 2021).

phases are in scope including data collection, training, continual learning, design, testing, evaluation and use.

In addition, the AHRC's 'Free and Equal: An Australian Conversation on Human Rights' project has released a Position Paper outlining an extensive reform agenda for federal discrimination laws.²²⁶ This Position Paper does not, however, contain any discussion or make any recommendations in relation to how those laws should deal with issues of algorithmic bias and discrimination.

Finally, the Department of the Prime Minister and Cabinet is currently conducting a consultation regarding the regulation of AI and automated decision-making. '[I]dentifying where new regulation may be required to minimise existing and emerging risks', including the potential of algorithmic decision systems for bias and discrimination, forms part of this consultation.²²⁷

226 'Free & Equal: A Reform Agenda for Federal Discrimination Laws' (Position Paper, Australian Human Rights Commission, December 2021).

227 See Department of the Prime Minister and Cabinet (Cth), 'Positioning Australia as a Leader in Digital Economy Regulation: Automated Decision Making and AI Regulation' (Issues Paper, March 2022) 2.