

Submission to the Inquiry into the Digital Transformation of Workplaces

Associate Professor Alysia Blackham
Melbourne Law School, University of Melbourne

I write this submission as an expert in equality law, who has written on the risks posed by adopting automated decision-making tools in the workplace, particularly in relation to discrimination and bias.¹ I draw on that research in this submission.

Terms of Reference

d): The effects of these techniques on equality, discrimination, and dignity at work

f): The effects on gender equality, Closing the Gap and disadvantaged and vulnerable cohorts of workers

There are currently significant risks with adopting automated decision-making tools in the workplace, including the risk of bias, discrimination and error. Discrimination and bias can lead to negative personal and social outcomes, including ill-health, stress and lower workplace productivity. This therefore poses a significant risk to Australians and Australian workplaces.

1) The risk of bias and discrimination

Discrimination and bias in automated decision-making tools can emerge through three key issues. First, it might originate from poor quality or inappropriate input data for training or testing ('garbage in, garbage out'), including through the use of biased, historical or out-of-date data, or data which under- or over-represents certain groups.² Second, it might reflect technical bias in the algorithm itself, derived from technical or human constraints.³ Third, there might be emergent or user bias in how the algorithm is applied or deployed, either due to new societal knowledge, a mismatch between how the algorithm was designed and how it is ultimately deployed,⁴ or poor quality organisational data in applying the algorithm.

The Amazon recruitment tool is a well reported example of the risks of deploying automated decision-making tools at work.⁵ The Amazon tool reviewed applicants' resumes, to determine which applicants were most likely to be successful recruits.⁶ Applicants were graded from one to five stars. The tool was ultimately scrapped, however, because it systematically discriminated against women applicants for software development and technical jobs. The tool had been trained on resumes (and, presumably, hiring outcomes) from 10 years of job applicants; men are significantly over-represented in the field, and were therefore significantly

¹ Alysia Blackham, 'Setting the Framework for Accountability for Algorithmic Discrimination at Work' (2023) 47(1) *Melbourne University Law Review* 63 <https://law.unimelb.edu.au/_data/assets/pdf_file/0010/4805047/02-Blackham-62.pdf>; Alysia Blackham, 'When AI Gets It Wrong, Workers Suffer', *Pursuit* (29 November 2023) <<https://pursuit.unimelb.edu.au/articles/when-ai-gets-it-wrong-workers-suffer>>.

² Alina Köchling and Marius Claus Wehner, 'Discriminated by an Algorithm: A Systematic Review of Discrimination and Fairness by Algorithmic Decision-Making in the Context of HR Recruitment and HR Development' (2020) 13(3) *Business Research* 795, 800 ('Discriminated by an Algorithm'). See also Pauline Kim, 'Data-Driven Discrimination at Work' (2017) 58(3) *William & Mary Law Review* 857, 861.

³ Köchling and Wehner (n 2) 801.

⁴ *Ibid.*

⁵ Jeffrey Dastin, 'Amazon Scraps Secret AI Recruiting Tool That Showed Bias against Women', *Reuters* (online, 10 October 2018) <<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>>.

⁶ *Ibid.*

over-represented in the pool of resumes and successful applicants. The tool ‘learnt’ that male applicants were to be preferred. The tool therefore reportedly penalised applications with the word “women’s,” or the name of all-women’s colleges.⁷ Further attempts to develop a tool that could crawl the web to identify candidates worth recruiting, trained on 50,000 terms that appeared in past candidates’ resumes, again resulted in a tool that favoured terms that appeared in men’s resumes, such as ‘executed’ and ‘captured’.⁸

Other studies have found that AI can pick up gender signals in a CV, even when a name and pronouns are removed. And, even if AI is trained to be gender-neutral, it might still discriminate against parents.⁹

The use of automated decision-making tools therefore raises considerable risks of bias and discrimination, and the risk that these in-built biases might be replicated at scale. Automated discrimination is likely to disproportionately affect those who are already most impacted by discrimination, including Aboriginal and Torres Strait Islander peoples and First Nations peoples.¹⁰ As the European Economic and Social Committee has argued,

the development of AI is currently taking place within a homogenous environment principally consisting of young, white men, with the result that (whether intentionally or unintentionally) cultural and gender disparities are being embedded in AI, among other things because AI systems learn from training data. This data should be accurate and of good quality, diverse, sufficiently detailed and unbiased. There is a general tendency to believe that data is by definition objective; however, this is a misconception. Data is easy to manipulate, may be biased, may reflect cultural, gender and other prejudices and preferences and may contain errors.¹¹

2) A lack of transparency

A further, exacerbating, risk relates to the lack of transparency in relation to how automated decision-making tools are used in the workplace. There is minimal transparency about how employers are using automated decision-making tools, and for what purposes. There are many software tools on the market that use automation to streamline human resource functions – from recruitment, to performance management, and even to dismissal. However, there is currently no requirement for employers to disclose how they use automated decision-making tools, or if there is a problem or error in how the technologies are used. A lack of transparency means that it is difficult – if not impossible – to evaluate the extent to which discrimination is

⁷ Ibid.

⁸ Ibid.

⁹ Lea Frermann et al, ‘When It Comes to Jobs, AI Does Not like Parents’, *Pursuit* (23 July 2023) <<https://pursuit.unimelb.edu.au/articles/when-it-comes-to-jobs-ai-does-not-like-parents>>.

¹⁰ For a technical example of how this might affect Aboriginal and Torres Strait Islander people, see Australian Human Rights Commission, *Using Artificial Intelligence to Make Decisions: Addressing the Problem of Algorithmic Bias: Technical Paper* (2020) 30–32 <https://humanrights.gov.au/sites/default/files/document/publication/ahrc_technical_paper_algorithmic_bias_2020.pdf>.

¹¹ *Opinion of the European Economic and Social Committee on ‘Artificial intelligence — The consequences of artificial intelligence on the (digital) single market, production, consumption, employment and society’ (own-initiative opinion)* (2017/C 288/01) para 3.5.

built into automated decision-making tools,¹² and even more complex for individuals to challenge discriminatory impacts.

For example, in the Australian Public Service, there was an attempt to use AI-assisted technology to manage promotions. Many of these promotions were later overturned for not being on the basis of merit.¹³ This was only revealed because the Public Service has a dedicated Merit Protection Commissioner. In the private sector, where most people work, these forms of review are not always in place.

A lack of transparency in automated decision-making tools themselves (as a ‘black box’) also makes it difficult for employers or users to understand how the system works, or to identify problems before they occur: the people and organisations deploying automated decision-making tools often fail to understand the limits and confines of the technology, or where predictions are ‘coming’ from.

Term of reference

e): Appropriate safeguards or regulatory interventions to guide responsible implementation in the workplace, including the digital skills and resources necessary for employers to appropriately utilise these technologies

1) Gaps in Australian regulation

Australian privacy law significantly lags behind countries like the UK, and the European Union. Automated decision-making tools are built on substantial quantities of data, yet Australian workers have few data protection rights. The *Privacy Act 1988* (Cth) contains a blanket exception for ‘employee records’ – while an employer needs consent to initially gather new data, there are no limits placed on that data once it is held. The *Privacy Act 1988* (Cth) also does not apply to small businesses – who employ the vast majority of workers in Australia. While the government has committed to considering reform in this area, this has not yet occurred.

Discrimination law might fill this gap: it might be possible to show that automated decision-making tools discriminate against certain people or groups. However, discrimination law mostly relies on individuals making a complaint¹⁴ – and few people do, even when they know they have been discriminated against. With automated decisions, we may not even know what technology has been used, let alone if it is discriminatory. This makes individual complaints a limited tool for addressing these issues.

2) International perspectives

In Europe, there is a strong regulatory framework for privacy and data protection law – the General Data Protection Regulation (GDPR) – which requires an ultimate human decision

¹² Alysia Blackham, “‘We Are All Entrepreneurs Now’: Options and New Approaches for Adapting Equality Law for the ‘Gig Economy’” (2018) 34(4) *International Journal of Comparative Labour Law and Industrial Relations* 413, 418 (‘We Are All Entrepreneurs Now’).

¹³ Merit Protection Commissioner, *Annual Report 2021-22* (2022) <<https://www.workplaceexpress.com.au/files/2022/APS%20Merit%20Annual%20Report%202021-22.pdf>>.

¹⁴ Alysia Blackham, *Reforming Age Discrimination Law: Beyond Individual Enforcement* (Oxford University Press, 2022) (‘*Reforming Age Discrimination Law*’).

maker when an automated decision significantly affects people’s lives. In the EU, gig workers have used this to challenge Uber and Ola when they were automatically terminated as drivers.¹⁵

Adopting rigorous privacy law – like the GDPR – would be an important first step for Australian law reform. In particular, this should entail new requirements for:

- **Transparency:** so the data automated decision-making tools are based on, their use, and their outcome, is understandable, communicable, and able to be effectively challenged.¹⁶
- **Substantive limits** on when data can be collected, and when automated decision-making tools can be used.
- **Oversight** of automated decisions or outcomes by a human decision-maker – an approach described as ‘human-in-command’.¹⁷
- **Evaluation and consultation** prior to the adoption of automated decision-making tools, to ensure the technology is being used for a legitimate purpose, is appropriate for achieving that purpose, and includes appropriate safeguards for individual rights. This takes the form, for example, of Data Protection Impact Assessments in the EU.¹⁸ It could also be embedded in collective bargaining and negotiation.¹⁹

The EU has recognised the need to go further, with the passage of the EU Artificial Intelligence Act. The Act generally categorises the use of AI systems at work as high risk – including where used for recruitment, for promotion, termination, the allocation of tasks or to evaluate performance (article 6, Annex III article 4) – unless the system is preparatory, used for checking or improving human decision making, or undertaking narrow procedural tasks (article 6(3)). High-risk AI systems require a risk management system and testing (article 9), appropriate transparency (article 13), and human oversight (article 14). Obligations are imposed on those who develop, provide, import, distribute and deploy high-risk AI systems. Employers need to inform workers’ representatives and affected workers before deploying a high-risk AI system (article 26(7)).

3) Discrimination law reform in Australia

Discrimination law also needs reform to address the discrimination risks of automated decision-making tools. Rather than relying on individuals to make a complaint, it is necessary to adopt positive, proactive duties on employers. This could be modelled, for example, on the existing positive equality duty under the *Sex Discrimination Act 1984* (Cth) s 47C, but aimed specifically to address the discrimination risks posed by automated decision-making tools, and covering all protected grounds.

¹⁵ Raphaël Gellert, Marvin van Bekkum and Frederik Zuiderveen Borgesius, ‘The Ola & Uber Judgments: For the First Time a Court Recognises a GDPR Right to an Explanation for Algorithmic Decision-Making’, *EU Law Analysis* (28 April 2021) <<http://eulawanalysis.blogspot.com/2021/04/the-ola-uber-judgments-for-first-time.html>> (‘EU Law Analysis’); Natasha Lomas, ‘Drivers in Europe Net Big Data Rights Win against Uber and Ola’, *TechCrunch* (5 April 2023) <<https://techcrunch.com/2023/04/05/uber-ola-gdpr-worker-data-access-rights-appeal/>>.

¹⁶ Janine Berg, ‘Protecting Workers in the Digital Age: Technology, Outsourcing, and the Growing Precariousness of Work Automation, Artificial Intelligence, & Labor Law’ (2019) 41(1) *Comparative Labor Law & Policy Journal* 69, 91 (‘Protecting Workers in the Digital Age’). See GDPR arts 13, 15.

¹⁷ Global Commission on the Future of Work, *Work for a Brighter Future* (2019) 78, 43 <https://www.ilo.org/wcmsp5/groups/public/---dgreports/---cabinet/documents/publication/wcms_662410.pdf>.

¹⁸ GDPR art 9(1).

¹⁹ See, eg, *Proposal for a Directive of the European Parliament and of the Council on improving working conditions in platform work* COM(2021) 762 final art 9.

In particular, I have recommended²⁰ the adoption of positive equality duties imposing obligations on employers that include requirements to:

- give proper consideration and take proportionate action to eliminate discrimination and advance equality of opportunity;²¹
- collect, analyse and publicize data on the protected characteristics of those in the workplace;²²
- report on what data is being collected, for what purpose, and how it is being processed and used, and data quality control measures adopted;
- report on what automated decision-making tools are being adopted, including:
 - how the technology operates, and for what purpose,
 - the training and testing and organisational data being deployed, data quality measures in place, and how recently data were reviewed or renewed,
 - how the system operates in the context of the specific workplace,²³
 - the role of human oversight in the process, and
 - outcomes of automated decision-making tools across the workforce, including across protected characteristics;
- adopt policies to demonstrate what is being done to address and eliminate discrimination; and
- consult and engage with workers and worker representatives in the collection of data, and adoption and use of automated decision-making tools.

These duties should be enforceable by a statutory regulator, like the Australian Human Rights Commission, with powers akin to those in relation to the positive duty in s 47C. They should also be enforceable by workers and worker representatives, to ensure multiple checks on compliance.

²⁰ Blackham, ‘Setting the Framework for Accountability for Algorithmic Discrimination at Work’ (n 1).

²¹ See Blackham, *Reforming Age Discrimination Law* (n 14) ch 7.

²² See, eg, *Gender Equality Act 2020* (Vic); Lauren Ryan et al, *Laying the Foundation for Gender Equality in the Public Sector in Victoria: Final Project Report* (report, University of Melbourne, 2022) <<https://doi.org/10.26188/19254539>> (*‘Laying the Foundation for Gender Equality in the Public Sector in Victoria’*).

²³ See, eg, Köchling and Wehner (n 2) 839.