

29 September 2023

NSW Attorney General
By email: nsw-lrc@dcj.nsw.gov.au

Anti-Discrimination Act review

About us

The **UNSW Allens Hub for Technology, Law and Innovation** ('UNSW Allens Hub') is an independent community of scholars based at UNSW Sydney. As a partnership between Allens and UNSW Law and Justice, the Hub aims to add depth to research on the diverse interactions among technology, law, and society. The partnership enriches academic and policy debates and drives considered reform of law and practice through engagement with the legal profession, the judiciary, government, industry, civil society and the broader community. More information about the UNSW Allens Hub can be found at <http://www.allenshub.unsw.edu.au/>.

About this Submission

We are grateful for the opportunity to make a preliminary submission on the [terms of reference](#) for the Anti-Discrimination Act review. Our submission reflects our views as researchers; they are not an institutional position. This submission can be made public.

We focus on the first proposed term, "whether the Act could be modernised and simplified to better promote the equal enjoyment of rights and reflect contemporary community standards" and the related fourth proposed term "whether the existing tests for discrimination are clear, inclusive and reflect modern understandings of discrimination". In relation to both, we urge the review to focus *inter alia* on the question of the impact of socio-technical systems for decision-making in general and the impact of machine learning approaches in particular.

"Algorithmic bias" and discrimination law

Data-driven influencing, whether using machine learning or statistical techniques, is based on the idea that if we can understand empirical connections between variables, we can predict other variables. When these variables involve human behaviour and result in decisions that affect those humans, fairness and anti-discrimination principles are critical. Currently, discrimination law protects against discrimination on the basis of protected attributes in a range of contexts but does not protect against many examples of "algorithmic bias" because the laws were written at a time when the primary concern was human animus and cognitive limits rather than bad (machine learning) models.

Thus, in the context of machine learning, discrimination law does not operate as effectively as it might. Organisations may well seek to avoid direct discrimination by removing variables without eliminating disparate impact. Complex machine learning algorithms do not necessarily set a

“requirement or condition” of (say) being male, rather they factor in correlates with being male among many other variables in ways that influence the outputs, and hence the decisions. Organisations will often also be able to avoid accusations of indirect discrimination by relying on the reasonableness test – to the extent the system sets a “requirement or condition”, that it is reasonable to use it where it is generally useful in, say, filtering job applications. The primary problem is that discrimination law does not currently require any testing of black-boxed systems.

It would be desirable to reform discrimination legislation so that the need for testing for discriminatory outcomes when a decision affecting a human is made in part or entirely on the basis of data driven inference are laid out more clearly. Guidance for such testing can be found in both international standards and the work of organisations such as the Gradient Institute. Testing may also be required in the context of generative AI and search, to ensure that people do not, for example, only “see” white males in professional roles. Legal changes could reduce the incentive to avoid direct discrimination by deleting variables, which restricts the ability to test for disparate impact. If done well, such requirements would not only apply to artificial intelligence or machine learning but to any potentially discriminatory data-driven process.

Reforming discrimination law versus regulating the technology

We are making a separate submission to the *Inquiry into artificial intelligence in New South Wales*. However, in both, we emphasise that law reform in this area is better based around values (such as non-discrimination) than creating new technology-specific regulation. There are several reasons for this, which we will go into more detail on in that submission. In brief:

1. Decision-making is complicated and is often a combination of humans and outputs of AI systems. For example, an AI system might ‘score’ applicants, leaving the decision to a human. Regulating “decisions made by AI systems” separately from “decisions made by humans” is thus unhelpful. It is better to begin with the issue (like anti-discrimination) and design laws that are effective in preventing unfair discrimination in all contexts.
2. It is difficult to define a technology (say, “artificial intelligence”) to be regulated. This is both because there is no good definition of artificial intelligence that places in scope systems where there is a risk of unfair bias and systems where there is no such risk. As mentioned earlier, data-driven inferencing can involve pen, paper and statistics although there are obviously efficiencies and other improvements in using more modern techniques such as machine learning. But from a scoping perspective, the problem lies in reliance on a particular form of inferencing and particular assumptions about underlying data sets rather than a particular set of techniques.

Engaging with the review

This preliminary submission seeks to highlight the need to include discrimination issues arising out of the use of socio-technical systems, particularly the use of machine learning techniques, in the scope of the review. We are keen to continue the conversation and discuss our thoughts on specific proposals in more detail as the review progresses.

Yours sincerely,
Lyria Bennett Moses (Director, UNSW Allens Hub)