



Level 5, 183 North Quay Brisbane Qld 4000
PO Box 13035, George Street Brisbane Qld 4003
T: 07 3025 3888 | F: 07 3025 3800
Freecall: 1800 012 255
ABN: 116 314 562

Mr Edward Santow,
Human Rights Commissioner,
Australian Human Rights Commission,
Level 3, 175 Pitt Street,
SYDNEY NSW 2000

By email:

tech@humanrights.gov.au

30th March 2020

RE: HUMAN RIGHTS AND TECHNOLOGY DISCUSSION PAPER 2020

We welcome and appreciate the opportunity to make a submission in relation to the issues surrounding current practices and future use of artificial intelligence systems in decision making which has a legal or similarly significant effect for an individual where the artificial intelligence system AI materially assisted in the decision-making process (“AI-informed decision making.”) The Discussion Paper raises a broad spectrum of important issues, many of which are in urgent need of addressing and many of which have largely remained invisible to date.

Preliminary Consideration: Our background to comment

The Aboriginal and Torres Strait Islander Legal Service (Qld) Limited (ATSILS), is a community-based public benevolent organisation, established to provide professional and culturally competent legal services for Aboriginal and Torres Strait Islander people across Queensland. The founding organisation was established in 1973. We now have 26 offices strategically located across the State. Our Vision is to be the leader of innovative and professional legal services. Our Mission is to deliver quality legal assistance services, community legal education, and early intervention and prevention initiatives which uphold and advance the legal and human rights of Aboriginal and Torres Strait Islander people.

ATSILS provides legal services to Aboriginal and Torres Strait Islander peoples throughout the entirety of Queensland. Whilst our primary role is to provide criminal, civil and family law representation, we are also funded by the Commonwealth to perform a State-wide role in the key areas of Community Legal Education, and Early Intervention and Prevention initiatives (which include related law reform activities and monitoring Indigenous Australian deaths in custody). Our submission is informed by four

and a half decades of legal practise at the coalface of the justice arena and we therefore believe we are well placed to provide meaningful comment. Not from a theoretical or purely academic perspective, but rather from a platform based upon actual experiences.

OVERVIEW

As noted in the discussion paper, the use of artificial intelligence assisted decision making raises a wide variety of issues surrounding their use. We are broadly supportive of the initial views of the Commission, this submission focusses on a few key topics, including transparency, accountability and the availability of effective remedies for individuals affected by the decision making and effective responses which address or stop systemic issues.

We are especially alert to the problem that some technology may both positively affect some groups and yet at the same time negatively impact other groups. Groups that are already experiencing levels discrimination or socio-economic disadvantage before a new technology is implemented are more vulnerable to experiencing negative human rights impacts from that technology.

THE PROPOSALS

Necessary Further Enquiries

We support the proposals to properly investigate:

- a) The accountability of AI-informed decision making (**Proposal 3**), and
- b) The use of AI in decision making by the Australian Government (**Proposal 17**).

At the core of all the issues is the need to consider reform or other change needed to protect the principle of legality and the rule of law. We support measures to ensure that AI-informed decision making does not evade or bypass the existing protections of the law and human rights standards which form part of domestic and international law. This is not a problem isolated to Australia, we note the work that is being done in other parts of the world in this regard, especially the European Union with respect to the General Data Protection Regulation, and we note the sense of remaining broadly consistent with other protective measures adopted in other jurisdictions.

Ethical and legal frameworks which already form part of government decision making also need to be embedded in the AI systems used to support government decision making. There needs to be, in effect, a building code for systems development with the explicit goal of protecting the principle of legality and protecting human rights. The opportunity of embedding such frameworks is that it could improve consistency and accountability in government decision making.

Because AI systems evolve over time, monitoring and evaluating the use of AI in government decision making will also require ongoing checks on the AI systems.

Remedies and Measures

We support the introduction of remedies and measures so that there is an opportunity for redress when mistakes are made due to the dataset or algorithmic design.

In order for measures and remedies to be at all meaningful, AI-informed decision making must be explainable so that a person affected by AI-informed decision making is able to assert a right to a reasonable explanation. (**Proposal 7**). Namely:

If an individual would have been entitled to an explanation of the decision were it not made using AI, then the AI-informed decision making should also be explainable.

the individual should be able to demand:

- a) a non-technical explanation of the AI-informed decision, which would be comprehensible by a lay person, and
- b) a technical explanation of the AI-informed decision that can be assessed and validated by a person with relevant technical expertise.

In each case, the explanation should contain the reasons for the decision, such that it would enable an individual, or a person with relevant technical expertise, to understand the basis of the decision and any grounds on which it should be challenged.

In order to give full effect to proposal 7, legislation should be introduced by the Australian Government to require the explainability of AI-informed decision making and we consider it fundamental to other reforms or other changes needed to protect the principle of legality and the rule of law.

We support the three-pronged approach to algorithmic accountability proposed by ASIC, requiring:

- a) a responsible person for an algorithmic system
- b) that relevant algorithmic outputs be capable of an explanation so that decisions relying on them can be meaningfully explained to customers
- c) and opportunity for redress when mistakes are made due to the dataset or algorithmic design.

Specific Issues – Facial Recognition

With respect to specific problematic issues on the use of facial recognition in decision making that has a legal or similarly significant effect for individuals, we would support a legal moratorium on the use of the technology in those circumstances until an appropriate legal framework has been put in place.

Specific Questions Surrounding Proposal 7 and Questions B and D

In response to Questions B and D:

While Question B raises an important safeguard of a rebuttable presumption that a decision was not lawfully made in circumstances where a person is responsible for an AI-informed decision and the person does not provide a reasonable explanation for that decision, it does not go far enough.

The further safeguard should be that in the absence of a satisfactory explanation for a decision, including the ability to understand the basis of the decision and any grounds on which it should be challenged, a person affected by the decision should at law be able to require a substitute decision be made by a human decision maker.

Additionally, in the circumstances where the disputed decision is being litigated in courts or tribunals, a judicial officer or tribunal member should have the power *proprio motu* to refer the matter to a human decision maker.

In our view this will be extremely important where neural networks are the form of Artificial Intelligence used to “inform” decision making. Presently there is widespread use of neural networks in a number of applications. While other forms of Artificial Intelligence exist and are capable of description of their processes, a neural network has no epistemological knowledge. A neural network will extract patterns from vast amounts of data, it emulates patterns but it does not articulate them. In other words, it does not know what it knows. A neural network becomes very good at seeing patterns, but whether those patterns are really there or valid patterns upon which to make decisions is another question.

Compounding on that problem is the problem of “drift” in neural networks. Neural networks are not static, they continue to learn and develop from their environment. A neural network may initially be operating satisfactorily but may then evolve to operate incorrectly. Or to put it in non-technical terms, a neural network may go “off-piste”. That is an unsatisfactory foundation for decision making and raises the question whether neural networks, as opposed to other Artificial Intelligence architectures, should be used at all in decision making that affects people’s rights.

Where human decision making excels over machine pattern matching and decision making is the ability to understand context in which information arrives and to recognise when there is a mismatch between the data and the context. Thus humans will not, unlike an AI systems, make the mistake of identifying faces in clouds. Humans are also much more skilled at extrapolating from old data to deal with novel situations. For those reasons, not just in airplane cockpits and nuclear powers stations, the opportunity should remain for a human to override decision making by an AI system.

Accountability Measures and Structural Reform

The regulation of AI-informed decision making requires a greater dialogue between legal and technical experts. Effective AI systems that respect the rule of law means that system developers require input from legal practitioners on ethical frameworks and parameters, and legal practitioners seeking effective legal remedies and responses will need to draw upon technical expertise and tools to effectively challenge invalid decisions.

We support **Proposal 19** for the establishment of an AI Safety Commissioner whose role would focus on preventing individual and community harm and protecting and promoting human rights with respect to the development and use of AI in Australia.

We agree that the proposed AI Safety Commissioner should:

- a) build the capacity of existing regulators and others regarding the development and use of AI;
- b) monitor the use of AI, and be a source of policy expertise in this area;
- c) be independent in its structure, operations and legislative mandate;
- d) be adequately resourced, wholly or primarily by the Australian Government;
- e) draw on diverse expertise and perspectives;
- f) determine issues of immediate concern that should form priorities and shape its own work.

Fundamental to both ideas is the capacity for decision-making systems to be tested and analysed for compliance with the principle of legality and human rights standards. We therefore support **Proposal 15** for the establishment of a regulatory sandbox to test AI-informed decision-making systems for compliance with human rights.

We thank you for the opportunity to provide feedback on the discussion paper.

Yours faithfully,

Shane Duffy - CEO