Amazon Web Services Submission Regarding Australian Human Rights Commission: Human Rights and Technology Discussion Paper

Australian Human Rights Commissioner
Email: tech@humanrights.gov.au

I would like to thank the Australian Human Rights Commission (the "**Commission**") for giving Amazon Web Services (**AWS**) the opportunity to provide feedback on the **Human Rights and Technology Discussion Paper** (henceforth referred to as the "**Discussion Paper**").

Artificial Intelligence (**'AI'**) encompasses a wide variety of technologies. Potential use cases are vast. Human rights issues, and corresponding possible regulations, will vary broadly depending on the nature of the deployment. For example, human rights issues in the context of autonomous vehicles are very different from those that pertain to an AI deployment that aids a judge in criminal sentencing. Consequently, there should be no one-size-fits-all regulatory response. AWS supports a risk-based approach to AI regulations and standards. Regulation of applications of AI that pose high risk, or are of significant consequence to individuals, should be different from those that present a low-risk to individual human rights. In this regard, the Commission's approach of focusing on high-risk deployments of AI, or "edge-cases" which could result in harm to an individual's human rights, is an appropriate method for considered regulation.

This submission sets out AWS's feedback on the Discussion Paper.

1. Proposed thresholds for the definition of "AI-Informed decision making" should include a clear link to human rights impacts (in response to Question A of the Discussion Paper).

The Discussion Paper goes into detail identifying the scope of "AI-informed decisions" that would have the greatest impact on human rights. It proposes two tests for when such decisions would be within scope for "human rights" concerns. Namely, that 'AI-informed decision making': (a) involves a decision that has a legal or similarly significant effect for an individual; and (b) AI must have materially assisted in the process of decision-making. These two tests are broad and inadequately defined in the Discussion Paper. As proposed, it could result in the over-indexing of use-cases. Given the focus on human rights harms, the tests should more directly account for both the risks to human rights as well as the risk of actual harm to the individual. In addition, the Commission could consider further clarifying the concept of "material". In Chapter 5.3 of the Discussion Paper, two scenarios are outlined that constitute materiality: (1) where key elements of the decision-making process are automated; and (2) where AI is used in a decision-making process to generate a data point that "bears in a material or significant way on the ultimate decision". The Commission may want to consider further refining the definition of 'AI-informed decision making' to reflect the threshold of a "risk of harm to human rights", for example:

> AI-informed decision-making involves a decision that has a legal or similarly significant effect for an individual *that adversely impacts human rights;* and AI must have materially assisted in the process of decision-making*, where "material" refers to a deployment of AI that has been designed to include little to no human oversight or where a human has over-relied on data generated by the AI resulting in an adverse impact on human rights*.

The term "AI-informed decisions" could confuse if taken outside the scope of this paper. An alternative definition with more wide-ranging application such as "human-rights impacted AI-informed decisions" might be more appropriate.

2. The creation of a statutory cause of action for a "serious invasion of privacy" would be inefficient and ineffective in addressing serious privacy concerns. Creation of a right should be considered as part of a broader discussion on the reform of Federal Privacy legislation (Proposal 4).

Experience in other jurisdictions with statutory torts or causes of action in relation to privacy (e.g. the US and Canada), demonstrates that when it comes to privacy interests, "harms are largely inchoate and intangible".[1] In the US, for example, data has shown that even where class members have suffered a concrete injury, they are unlikely to receive material compensatory or injunctive relief through private litigation. Private rights of action also risk undermining appropriate agency enforcement and increase the risk of inconsistent application of the law, particularly in cases where the law is not already well defined or the regulation relates to new technology or new areas of the law. A statutory cause of action would allow the courts, plaintiffs and lawyers to effectively set policy, rather than allowing expert regulators to shape and balance policy and protection outcomes. The proposal to create a separate statutory cause of action could lead to significant detrimental effects on commerce, including chilling innovation and delaying or deferring deployments of AI. Even where there is little to no risk of "serious invasions of privacy", the existence of such a cause of action increases uncertainty as it threatens companies with litigation. Particularly for nascent areas such as AI deployment, this could significantly affect investor sentiment and have an extremely detrimental effect for technology start-ups in Australia.

Any proposal for a statutory cause of action for serious invasions of privacy requires a circumspect approach, and its scope must be subject to an appropriate balancing exercise. Crucially the threshold for "serious invasion" has to incorporate a clear test for "actual harm", for the purpose of establishing liability under the statutory cause of action, beyond just circumstances where there was a "reasonable expectation of privacy" and the "act or conduct" is "highly offensive to a reasonable person".[2]

AWS recognises the need to protect individuals against invasions of privacy however this needs to be considered holistically and should be developed through a review of existing privacy legislation to address any gaps, rather than the introduction of a separate statutory cause of action. Privacy regulators are best placed to enforce a "serious invasions of privacy" and amendments can be made to the privacy legislation to allow for direct remedies to individuals who have suffered significant harm, based on a finding of breach by the Privacy regulator. Any review of privacy frameworks should be principle-based, technology neutral and not limited or tied specifically to "AI-informed decisions".

3. Avoid the creation of firm and inflexible regulations or laws that are overbroad and prevent the deployment of AI in legitimate and beneficial use-cases.

The use of AI must comply with all laws, including laws that protect human rights. There is therefore no ambiguity that existing laws apply to and may restrict the use of this technology in some circumstances. The creation of additional laws or regulations focused on AI or specific deployments, for example Facial Recognition Technology (**FRT**), may result in broad prohibitions and limits on the use of AI systems. This could result in a net loss for society when, because of a lack of certainty, organisations avoid the use of AI. This will be the case even if there are legitimate and beneficial outcomes that could arise from deployments. The two proposals that raise particular

[1] US Chamber, Ill-Suited Private Rights of Action and Privacy Claims, July 2019 https://www.instituteforlegalreform.com/uploads/sites/1/Private_Rights_of_Action_-_Ill_Suited_Paper.pdf

[2] For Your Information: Australian Privacy Law and Practice (ALRC Report 108), August 2008, Pg. 2584, Recommendations 74-1 and 74-2.

concern are (a) prohibitions on the use of an AI system if no "reasonable explanation can be provided" (Proposal 8) and (b) a moratorium on the use of FRT that has a legal or similarly significant effect on individuals (Proposal 11).

The proposal to prohibit the deployment of AI systems if "reasonable explanations for its decisions" cannot be produced (Proposal 8) needs to be further clarified and scoped**.** In scenarios where an AI-informed decision impacts an individual's human rights, the Commission has proposed that the AI deployer provide a "reasonable explanation that a layperson can understand". AWS recommends that the concepts of "reasonable explanation" and "decisions that *could* infringe the human rights of an individual", particularly where this relates to "technical explanations", should be more clearly tied to a set of information that is *necessary* to provide a layperson such explanations.

In this regard, the Commission might want to consider clarifying that "reasonable explanation" means an explanation that enables an individual layperson "affected by an AI-informed decision", to have access to an explanation that is accurate and sufficient to enable the individual to understand the decision. In an area such as AI, where there is rapid innovation and a lack of clear consensus on what constitutes a "reasonable explanation", having an outright prohibition on deployment, subject to this bar, would likely have a prohibitive effect on innovation and the deployment of such AI-enabled tools. With constant evolution and improvements in both AI systems and explainability tools, policymakers should consider whether it would be a socially beneficial outcome if complex deployments of AI, like a neural network, were restricted or discouraged. Such deployments can generate better predictions and improve lives but is not practically explainable in a manner that meets the above bar due to ongoing developments in explainability.

In this regard, the Commission might want to consider incorporating a balancing test for the risk of harm to human rights vis-à-vis the benefit of the deployment. Where the risk of harm is disproportionate to the positive outcomes, organisations should avoid deployment without the ability to reasonably explain how the decision was made, or ensure that mitigations (e.g. inclusion of human oversight, redress mechanisms) are put in place.

Related to the above, the suggestion to impose a "rebuttable presumption" that AI-informed decisions are not lawfully made if a legal person does not provide a reasonable explanation for that decision (Question B) is out-of-line with the existing legal regime in Australia. It is disproportionate and could create a culture of fear around the deployment of AI that would otherwise be socially beneficial. The Commission may therefore wish to re-consider the suggestion to impose such a rebuttable presumption.

The proposal for a moratorium on the use of FRT that has a legal or similarly significant effect for individuals until an appropriate legal framework has been put in place would cause uncertainty and impede the adoption of FRT. It will negatively affect Australian citizens and Australian innovators. Society has already seen FRT being used to prevent human trafficking, reunite missing children with their parents, improve the physical security of a facility by automating access, and moderate offensive and illegal imagery posted online for removal.

It is imperative that the formulation of a legal framework surrounding the use of FRT be delineated conceptually for the human rights issues that could arise. FRT is often used to 'narrow the field' from hundreds of thousands of potential matches, to a handful. This capability benefits society by making it easier and more efficient to complete tasks that would take humans far more time. However, FRT should not be used to make fully automated, final decisions, that might result in a violation of a person's human rights. In these situations, human review of facial recognition results should be used to ensure rights are not violated.

An over-broad and unqualified ban on FRT uses that have "legal or similar effect" could impede the beneficial use of FRT. Any proposed moratorium on the use of the technology would not add to existing protections, and could instead be counter-productive to the protection of an individual's human rights, if all the focus is placed on restricting the

use of the technology. Rather the focus of debate and regulation should be on the pertinent issue that often gets lost – distinguishing between what FRT technology does (automating image comparison) versus where it simply exposes or exacerbates existing social issues.

It should be mandatory to apply a balancing test for the risk of harm versus the benefit to society prior for any deployment of FRT. International best practices should be discussed and encouraged for certain higher-risk deployments. For example, AWS recommends that where law enforcement uses FRT for identification, or in a way that could threaten human rights, human review as well as a 99% confidence score threshold be required. Confidence scores are predictions and should be used as one element of the investigation and not as a sole determinant. AWS also supports the creation of a national legislative framework covering FRT through video and photographic monitoring on public or commercial premises, and we encourage deeper public discussion and debate about whether the existing video surveillance laws should be reviewed and updated. Our view is that the same notice framework should cover FRT and video/photo surveillance.

4. The proposal to impose a "rebuttable presumption" that the legal person who deploys AI-informed decision-making system is legally liable for the use of the system (Proposal 10) is not the appropriate standard for the deployment of *all* AI systems. If included its applicability should be narrowed with a clearer distinction in applicability between a technology deployer versus a technology provider.

As discussed above, the ways in which AI is deployed are extremely varied. Creating a rebuttable presumption that would apply across numerous contexts would be over-broad. Instead, a case-by-case approach should be taken. The bar applied by the Commission, that the main responsibility ought to rest on the deployer of the AI system, is appropriate. We encourage the Commission to consider further clarifying the distinction between a technology deployer and a technology provider. A technology provider, like a Cloud Service Provider, may provide an AI technology that allows the technology deployer to utilise machine-learning services or build and use their own custom machine learning models and capabilities. However, it is generally the technology deployer that controls and determines the way in which models and capabilities are used, data sets selected, and decisions made in connection with the results of such models.

5. Reform of Australian law to make it "easier to assess the lawfulness of an AI-informed decision making system", by providing better access to "technical information" including algorithms, must be clearly limited and scoped (Question C). Overreach by the government or the courts, including through broad powers to force source-code disclosures by service providers and/or developers as a means of transparency, threatens IP and may chill innovation/deployment of AIML technologies. AWS submits that any consideration of reform of access to "technical information" be limited, and exclude the disclosure of proprietary or confidential information.

Sincerely,

[redacted]

Head of Public Policy, Austral a & New Zealand
Amazon Web Services