Human Rights and Technology
Discussion Paper 2019
Submission

29th Feb 2020

Thank you for the opportunity to comment on the proposals submitted by this project.

My comments deal more broadly with the foundations of AI and how this impacts the development of human rights policy and laws.

Artificial Intelligence is fundamentally "artificial" intelligence. It is problematic that we refer to artificial intelligence as "AI". It is easy to forget the word "artificial". When we look at the development of medicine and food over time, there has been an explosion of artificial solutions. These have been tested under strict guidelines before being released as commercial viable products. Why is AI any different? Why have we allowed AI to develop as far as it has without the same level of control? Why have we accepted numerous forms of AI products becoming part of our daily life? Is the general public aware of AI adoption across government and business?

For example – Image recognition software makes a decision based on recognising a set of numbers in a grid. It **does not** recognise an image as a human with natural cognitive ability. The decision it makes is **artificial** and not natural.

When developing a regulatory framework around AI, can we draw on the experience from the Therapeutic Goods Administration or the Food Safety Authorities?

Explainable AI – at what level? Any business employing an AI model to make decisions impacting its customer base will not survive without a detailed explanation of the business model. The explanation will have to be accessible to the general public. A deep neural network for example, is often not understood by the developer. How will a solution like this be understood by the general public?

When developing a framework, the audience needs to be taken into account to define the level of explanation required. An explanation written for a data scientist will not be accessible by the general public.

Transparency – to what extent? True transparency with AI may involve disclosing the nature of the input data and the key classifications and weightings. It may not be enough to simply disclose an algorithm at the core of a decision making process. It is also not enough to include this information in the "Terms and Conditions". For example: A recruitment agency replying on AI to screen candidates. This needs to be disclosed as part of the job advertisement. It is not enough to have this disclosed via a link to the terms and conditions page of a website.

The framework needs to be clear on what constitutes transparency.

Augmented AI – what does the framework need to take into account if a model includes human intervention? Should the framework require the involvement of a human where the risk of human rights violation is high?

Human Rights risk matrix – to avoid an over regulated framework, how will risk be assessed? A few preliminary questions could be:

1. Is this a decision making algorithm?
2. What decision is being automated?
3. What is the human rights violation risk of a false positive/negative?
4. What data is required to make this decision fair?
5. Is this data captured? If not, can it be captured?
6. What privacy issues are triggered in relation to the collection of the required data?
7. Can we collect the data lawfully?
8. Are decisions being made on aggregated data?
9. Is there bias in the data?
10. Can the bias be removed from training data?
11. What are the ongoing data collection requirements to train the model?
12. What are the cyber security risks?

The role of data literacy and the development of AI. What is not often discussed in Australia is the role of data literacy in the world of big data. The discussion paper addresses the need for AI literacy and acknowledges the different levels at which this needs to be deployed. However, the education piece needs to address data literacy as a whole rather than "AI" literacy. The level of education required here is much more general. For example:

1. What is data?
2. What is "big data"?
3. How is it collected?
4. How is it analysed?
5. Data visualisation literacy eg: How Charts Lie by Alberto Cairo (written for a general audience)
6. Correlation is not causation
7. Statistics – fundamentals of mean, medium, standard deviation, p-values, probability etc

We are at the beginning of a transitional stage within what some call the "4th Industrial Revolution". Therefore, the level of teaching has to adapt to the different groups of our society based on current levels of data literacy. It is not enough to embed data literacy into the school system. We need to embed it within universities, business, government and general public through media campaigns. Protecting our human rights through policy and legislation needs to work hand in hand with our own defence mechanism. This can only develop by empowering the masses with basic data literacy skills.

Thank for taking the time to read my comments.

Regards,
Wendy Small

[redacted]