

Submission to the Australian Human Rights Commission

Preliminary observations on the Human Rights and Technology Issues Paper (2018)

Dr David Glance, *Senior Research Fellow, University of Western Australia, School of Physics, Mathematics & Computing*

Dr Marco Rizzi, *Senior Lecturer, University of Western Australia, Law School*

Background

The Australian Human Rights Commission's Human Rights and Technology Issues Paper (the Issues Paper) details a number of fundamental questions that need to be addressed to develop a governance framework that ensures technology supports, and does not violate, peoples' human rights.

All digital technologies have the potential to impact human rights. This often happens as an oversight on the part of the developers, but it can also be a consequence of deliberate abuses for commercial and/or political gain. A close look at current practices indicates that exploitation of private data and discrimination of individuals is the norm in the design and function of digital technologies. Whether this is due to the negligence or lack of awareness of developers, or to deliberate calculation instead is debatable and the answer would have to be nuanced. Yet the drivers behind this reality have largely been:

1. the provision of "free" services on the Internet supported by the liberal use of private data for the purpose of advertising;
2. the capillary use of surveillance systems operated by governments, adopted and justified in the wake of terrorist attacks, the use and abuse of which verges into forms of social control;
3. the unwillingness to make investments towards providing services and products that accommodate the disabled or other minority groups because this is not justified on a cost-benefit basis;
4. the unwillingness to make investments (or take political action) towards modifying the business model in [1] by adding security that properly protects private data.

Problematic aspects of Human Rights & Technology in Australia

The Australian legal landscape in this field suffers from the lack of a Bill of Rights. The protection of privacy and anti-discrimination laws, especially in relation to digital services and products are largely inadequate: weak laws accompanied by weak enforcement.

The technologies that have sparked the current discussion have been impacting human rights for quite some time now (on the basis of the 4 drivers depicted above). It is no surprise that the companies that are leaders in both Artificial Intelligence (AI) and Internet of Things (IoT) are Google and Facebook whose business models are entirely based on targeted advertising. Amazon and Apple are, in turn, leading in the use of AI and IoT for the sale of their services and products.

The Issues Paper raises important questions about topics like algorithmic bias. There are legal, technical and social aspects to this issue, which has gained momentum and international attention in recent years. This discussion is particularly pertinent in countries that have a constitutional basis for human rights and advanced laws on privacy such as the EU. In countries like Australia, however the debate suffers from the lack of foundations upon which to develop solid protective regulatory frameworks.

While there would be little disagreement that Australia needs significant strengthening of its legislation, there is little evidence that an Australian Government would act to strengthen human rights in the current political environment.

On the other hand, the option of self-regulation presents its own set of challenges. Again, there is little evidence that industry would self-regulate to effectively set limitations to the current prevailing business model based on targeted or personalised advertising. The fact that such a business model supports a large share of the digital economy and forms the principal source of income of tech giants illustrates how high the stakes are. The widespread and systematic violations of privacy, to which we have become accustomed in recent times, represent a true ‘devil’s bargain’ of the internet. The impact of such violations is further exacerbated by the consequences of the ‘attention economy’, which is producing a multiplication of fake news, deep fake videos and audio leading to political and social manipulation not only of individuals, but societies at large.

While issues such as AI, IoT, the blockchain and robots are currently dominant in public discourse, this is largely due to the fact that we are currently in the ‘hype’ segment of the cycle for these technologies, and their ultimate success is far from certain. One could argue convincingly that many of the challenges cited in the context of these new technologies are nothing new. Decision support systems, even those driven by machine learning, have been in existence since the days of the mainframe. Exclusion of groups with disabilities from the use and benefits of technology have also been a constant over the years – to give but one obvious example, the internet remains to this day, inaccessible to the deaf and blind.

Turning our attention to the goal of the Issues Paper, inquiring how to legislate and/or regulate technology in ways that protect and promote human rights in Australia, it must be kept in mind that the centrality of human rights is as much a product of legal and cultural traditions as it is of positive legislative and policy efforts. This explains why Europe has one of the most comprehensive legislative and regulatory frameworks, while Australia is currently lagging behind. Cultural attitudes can change over time and the Australian public’s attitude towards its own privacy may harden and become more demanding. If and how this will inspire legislative, governmental or judicial response remains to be seen.

Scope and aims of this submission

This submission constitutes a preliminary compilation of critical areas that will require in-depth research to inform the Commission’s action plan on the regulation of emerging technologies. Both empirical research and theoretical reflections are still needed to ground the process that will lead to the adoption of a final report in 2020.

As this is a preliminary exercise, this submission is structured around four main themes that cut across the various questions raised by the AHRC in the ‘Human Rights and Technology Issues Paper’ (2018). Detailed answers to each of the specific questions raised by the Commission will be possible further down the consultation process, once the necessary research will have been carried out. The thematic areas identified in this submission are:

1. **General need for evidence**: ensuring the debate is grounded on genuine data and not sets of assumptions
2. **Regulation of the technology itself and its use**: ensuring adequate procedures for decision-making review;
3. **Governance of the increasing pervasiveness of the technology**: focusing regulatory efforts on the right questions and the right regulatory targets;
4. **Protection of affected individuals**: understanding the place of collective rights and interests in an over-individualized framework.

I. General need for evidence

As a preliminary general observation, it must be noted that much of the rhetoric around technology and its future is based on speculation at this stage. It is ironic that, while there is much talk about how data will be used through machine learning and Big Data processing, there is a paucity of open, verifiable and trusted data on how technology is currently impacting human rights.

To inform the process undertaken by the Commission, we need data on the impact technologies are having on human rights in Australia. The measurement of impact in human rights has always been a contentious issue but it is possible. The United Nations' *Human Rights Indicators: A Guide to Measurement and Implementation* provides a framework to develop assessment methods and indicators of the impact of technologies on human rights. These indicators could help evaluating the nature and scale of the problem as well as its perception and impact on the community. This evidence could include, for example, data on fines, court cases, reported infringements of data protection and antidiscrimination laws through digital technologies.

II. Regulating the technology and its use

The Issues Paper is principally focused on the question of how to regulate emerging technologies such as AI, IoT, machine learning processes and their impact on particularly disadvantaged groups (with a focus on people with disabilities). This is a critical angle that essentially asks two separate but interrelated questions:

- How to regulate the use of emerging tech *within* sectoral regulatory frameworks?
- How to regulate the technology itself?

We submit that satisfactory answers to these fundamental questions require in-depth research to ground further theoretical and practical reflections on three overarching points.

1. **Correlation vs Causation: a potential clash of methods and goals.** AI and machine learning processes in general are characterised by a fundamental feature. That is, they produce predictive results by drawing correlations and identifying patterns in large datasets (Big Data as defined in the Issues Paper). This mechanism of 'correlative prediction' is fundamentally different from the causal inquiry, which performs an explicative function and plays a central role in many areas of the law (chiefly the law of torts and criminal law, but also, for instance, aspects of administrative law). With AI and learning algorithms being increasingly relied upon as complementary decision-making tools by law enforcement agencies, judiciaries and administrations, an essential preliminary point of inquiry should focus on whether or not and to what extent correlative and causal methods can coexist, integrate and complement each other.

2. **‘Merits review’ of AI and machine learning outputs.** The second preliminary aspect to be investigated has to do with the ‘reviewability’ of the output produced by algorithms and AI. A central tenet of the rule of law as developed in western liberal democracies over the past centuries is the fact that it must be possible to review the merits of decisions affecting individuals. A defining characteristic of the functioning of AI and learning algorithms, according to both computer and data scientists, is that it can be extremely difficult, if not impossible altogether, to review in detail how an algorithm came to a specific conclusion in an individual case. This poses a potentially insolvable conundrum in a society that seeks to preserve the rule of law.
3. **The ‘anchoring’ phenomenon.** In light of the previous two points (the potential clash of methods and goals, and the difficulties of a merits review system), a third issue requiring specific analysis is that of ‘anchoring’. Provided the final decision-maker remains human, to what extent does he or she anchor to the machine generated prediction? The anchoring phenomenon has been thoroughly explored by behavioural scientists following the pioneering studies of psychologists Daniel Kahneman and Aaron Tversky, and those methods should be adaptable to research that seeks to identify to what extent decision-makers ‘anchor’ to machine-generated predictions. Needless to say, anchoring to the output of a mechanism that can hardly be subject to merits review and obeys a fundamentally different logic from that traditionally relied upon in democratic legal systems is a risk calling for thorough assessment.

III. Focusing regulatory efforts

The third thematic area identified in this submission shifts the focus away from the technology itself to suggest that equal attention should be dedicated to the regulation of the raw data and materials that are necessary for the technology to function in the first place. A good example can be the blockchain.

The blockchain is said to be a transparent ‘ledger’ that can be constantly updated by any of its users, and simultaneously. As per the Issues Paper, the blocks of information that form the chain “can be stored across the internet, cannot be controlled by a single entity and have no sole point of failure.” The process of creating new blocks of information to update the chain is called ‘mining’. The computational power required for large scale mining, with its consequential use of natural resources and environmental impact, is significant. The notion that no single entity can control the ledger is tempered by the fact that only a relatively restricted number of users have the capacity to ‘mine’ and therefore control the blockchain (for whatever purpose the technology is used, cryptocurrency or other).

It bares reflecting whether instead of (or alongside) attempts to regulate the technology, efforts should concentrate on the regulation and accessibility of data-mining resources, as well as the regulation of the environmental impact of data-mining activities.

The same shift in focus can apply to other areas. For instance, aside from attempting to regulate machine-learning and AI technologies, it would be advisable to explore the possibility of getting the widespread phenomenon of targeted or personalised advertisement (discussed above) under control.

IV. The place of collective and individual rights

The final aspect this submission seeks to bring to the attention of the AHRC is that of individual and collective rights.

A preliminary observation is that, in the face of technological changes that happen at increased speed, attempts to regulate the technology run a significant risk of failure. A preferable approach would therefore be to focus on the goals that regulatory efforts aim to achieve, and confer sets of rights and remedies to that effect. These would be of two kinds:

- Individual rights: *based on the classic liberal understanding of the phrase* – the rights in question would involve chiefly the protection of personal data, access to resources and a right to review the process when one is affected by it;
- Collective rights: *based on the effects of the technology* – algorithmic data processing is designed to identify correlations and patterns. Typically, results of this form of processing take the form of artificially created profiles, groups or categories. Individual fitting certain profiling characteristics are assigned to a group (purposes can be varied and may include insurance, employment, criminal sentencing and bail, and many more). While traditional human rights protection is geared towards the liberal conception of individual rights, this submission posits that the complex challenges posed by the digital revolution would be better addressed if individual rights were to be complemented by collective ones, in two forms:
 - i. Rights attributed to individuals due to their allocation to a specific group. These rights should be developed to ensure that, in the face of profiling practices, individuals are guaranteed proportionate protection. We are often unaware that we exist within specific virtual groups, and we are allocated to these profiles on the basis of large datasets that we can hardly control. Yet, as our rights and material conditions are affected by these ‘memberships,’ it is paramount that adequate protections are put in place;
 - ii. Rights attributed to groups as such. Arguably, the groups (profiles, or categories) themselves deserve dedicated protection. The very act of grouping, based on correlative associations, is liable to create new forms of vulnerable minorities. An effective protection of vulnerable groups entails the creation of dedicated collective rights, as single individuals may often not be in position to enforce their individual rights effectively (e.g. the role of unions in labour law).

Conclusion

The process that will lead Australia to develop and adopt a legal and regulatory framework of emerging technologies (broadly speaking) that effectively protects human rights will be a lengthy and difficult one. It is also an inescapable challenge for contemporary societies, and as such, it needs to be addressed with rigor, time and dedication. As repeatedly stated, both empirical and theoretical research are paramount to the success of the venture.

Thank you for considering this submission, we look forward to the next steps in the process.