Response to Issues Paper on Human Rights and Technology

About Us

The Allens Hub for Technology, Law and Innovation ('The Hub') is a community of scholars at UNSW Sydney aiming to add breadth and depth to research on the interactions among law, legal practice and technological change in order to enrich scholarly and policy debates and enhance understanding and engagement among the legal profession, the judiciary, industry, government, civil society and the broader community. The views of those participating in this submission are our own, based on our research, and do not represent the official views of UNSW Sydney or Allens.

Introduction

This document attempts to summarise some of the research conducted at the Allens Hub, which may assist the Commission in exploring issues relating to technology and human rights. We are aware that some of the research referred in our paper may relate to very specific topics, as they were not originally written for this paper. Nonetheless, we hope our research will aid the Commission in exploring issues relating to technology and human rights, particularly in the areas of social media regulations and AI-informed decision making. Overall, we are grateful for the opportunity to present our views and hope this paper will help the Commission in completing his final report.

Primary authors: Adam Yu and Amanda Lo, student interns (in succession)
Participants/Researchers: Roger Clarke, Bassina Farbenblum, Daniel Joyce, Marc De Leeuw, Lyria Bennett Moses, Kayleen Manwaring, Justine Nolan, Monika Zalnieriute.

Question 1) What types of technology raise particular human rights concerns? Which human rights are particularly implicated?

While this question is a broad one, we focus here on a particular technological architecture, the Internet, as well as surveillance technologies and explore how they raise specific or general human rights concerns. We draw upon several strands of Allens Hub research we think would be particularly helpful. The first strand explores whether access to the internet should be a human right.[1] The second strand of argument focuses on whether human rights law can be used to protect citizens from being subject to mass surveillance conducted by their own governments.[2] The third strand focuses on the disproportionate impact of digital censorship and surveillance for marginalized groups.[3] The Commission may also find work on eObjects (enhanced objects) and their implications for consumer rights[4] to be useful, as the challenges identified also have negative implications for human rights to privacy, safety and security, non-discrimination and equal treatment.

Access to the Internet as a Human Right?

As the internet becomes an essential tool for expressing political opinions, academics and politicians alike have experimented with the idea of treating Internet Freedom as a human right.[5] Indeed, the Arab Spring and 'Me too' Movement both serve as powerful reminders on how important the internet is at influencing social change. However, there is unequal access to media, information and communications infrastructure, posing important questions on whether 'we can adapt the right of freedom of expression to extend it to deal with the imbalances that exist regarding communication flows and access to communications?'[6]

In 2010, Hillary Clinton outlined the US commitment to 'Internet Freedom,' alongside references to the role 'online organising' has played in human rights advocacy.[7] UN Special Rapporteur Frank La Rue has gone as far as suggesting that 'access to the Internet should be considered in human rights terms and that achieving universal access to the Internet should be a priority for all states.'[8] Overall, viewing the internet as a 'public good' contrasts with the current approach, where the internet is viewed as a 'private sphere' where entrepreneurs have free-reign to create world-shaping companies.

Protecting Privacy through Human Rights Law?

The Snowden revelations illustrated how public and private organisations have the capacity to use a range of surveillance technologies and act in ways that are 'invasive and detrimental to our liberty.'[9] The General

---

[1] Daniel Joyce, 'Internet Freedom and Human Rights' (2015) 26(2) *European Journal of International Law* 493.
[2] Daniel Joyce, 'Privacy in the Digital Era – Human Rights Online?' (2015) 16(1) *Melbourne Journal of International Law* 270.
[3] Zalnieriute, Monika. "The anatomy of neoliberal Internet governance: A queer critical political economy perspective." In D. Otto, *Queering International Law*. Routledge, 2017. 53-73.
[4] Kayleen Manwaring, 'Emerging information technologies: challenges for consumers' (2017) 17(2) *Oxford University Commonwealth Law Journal* 265; Kayleen Manwaring, 'Kickstarting reconnection – An approach to legal problems arising from emerging technologies (2017) 22 (1) *Deakin Law Review* 53.
[5] Joyce, above n 1.
[6] Ibid
[7] Hillary Rodham Clinton, Remarks on Internet Freedom, 21 January 2010, available at <http://www.state.gov/secretary/20092013clinton/rm/2010/01/135519.htm>.
[8] Frank La Rue, Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression (Special Rapporteur's Report), Human Rights Council, A/HRC/17/27, 16 May 2011.
[9] Joyce, above n 2.

Assembly resolutions on the right to digital privacy seek to attempts to extend human rights to online contexts.[10] However, the complexity of the digital environment, as well as the UN's institutional limitations, means that the practice of surveillance is difficult to curb.

Academics have debated over the utility of international law in protecting privacy. On the one hand, it has been argued that international human rights law offers non-discrimination in terms of treatment of citizens in different states and helps to overcome the partisan approach of domestic constitutions to such issues. However, critics have pointed to the limitations of privacy in the era of big data. Despite the development of privacy jurisprudence in domestic and international contexts, few concrete protections are in place.

Disproportionate Impact of Surveillance and Censorship for Marginalized Groups

The third strand of our research focuses on the narratives of the liberatory role of the Internet and digital technologies for marginalized groups and discusses how the neoliberal model has been used to repress and limit the rights of LGBTI people (among others) and how such repressions have been justified.[11]

Conclusion

Responding to the popularity of the Internet and surveillance technologies, attempts have been made at the international level to expand the scope of human rights. The Commission should carefully observe these latest developments when formulating policies which address the impact of technology on human rights.

[10] *The Right to Privacy in the Digital Age,* GA Res 68/167, UN GAOR, 3rd Comm, 68th sess, 70th plen mtg, Agenda Item 96(b), UN Doc A/RES/68/167 (21 January 2014, adopted 18 December 2013).
[11] Monika Zalnieriute, 'The anatomy of neoliberal Internet governance: A queer critical political economy perspective' In D. Otto, *Queering International Law*. Routledge, 2017. 53-73.

Question 3) How should Australian law protect human rights in the development, use and application of new technologies?

We have two strands of research that can potentially assist the Commission in answering this question. Firstly, we suggest that reforms in this area should not focus on new technologies per se, but instead the change brought about by these new technologies. Secondly, and far more specifically, using social media regulations in Germany and US as case studies, we argue that legislation that empowers social media companies to monitor and edit user comments may be dangerous to human rights and other approaches may be more suitable.

a) The right question

The question "What can we learn about the need for regulating new technologies" in question 3(b) is potentially the wrong question.[12] Essentially, it ignores the ways in which law that does not target a particular technology (practice or artefact) nevertheless influences the design of things and conduct relating to those things. Instead, the question the Commission should be asking is "How can Australian law adapt to ongoing technological change in ways that protect (or continue to protect) human rights?" It is the *change* in technological possibilities that creates the potential for new things, activities and relationships that *may* in some circumstances, without law/regulation, infringe human rights.[13]

Once the correct question is posed, it is easier to see what can be learnt from other countries. In Australia, the development of the law is left to a variety of institutions,[14] including law reform commissions and the Productivity Commission.[15] There are other models, including the use of Participatory Technology Assessment, common in Europe, that rely on technical rather than legal or economic expertise.[16] Overall an integrated, interdisciplinary approach could take advantage of the best features of both models.[17]

b) Social media regulations (a specific case study)

Rules and regulations are being imposed to prevent internet companies from disseminating harmful content. In regulating online content, Australian lawmakers need to carefully consider its potential impact on the freedom of speech.[18]

In 2017, Germany introduced the 'Network Enforcement Act' which make social media sites with at least 2 million users monitor its content and remove anything that is illegal within twenty-four hours (can be

---

[12] The reasons why "regulating technology" is the wrong phrase are discussed in Lyria Bennett Moses, 'Regulating in the Face of Socio-Technical Change' in *Oxford Handbook of the Law and Regulation of Technology* (2017) and Lyria Bennett Moses, "How to Think about Law, Regulation and Technology: Problems with "Technology" as a Regulatory Target" (2013) 5(1) *Law, Innovation and Technology* 1-20.

[13] See generally Lyria Bennett Moses, "Recurring Dilemmas: The Law's Race to Keep Up with Technological Change" (2007) 7 *University of Illinois Journal of Law, Technology and Policy* 239-285.

[14] Lyria Bennett Moses, 'Agents of Change: How the Law 'Copes' with Technological Change', (2011) 20(4) *Griffith Law Review* 763-794

[15] Lyria Bennett Moses, N Gollan and K Tranter, The Productivity Commission: a different engine for law reform? (2015) 24(4) *Griffith Law Review* 657-686.

[16] See European Parliamentary Technology Assessment, available at <http://www.eptanetwork.org/>.

[17] See Lyria Bennett Moses, "Bridging distances in approach: Sharing ideas about technology regulation" in Ronald Leenes & Eleni Kosta (eds), Bridging distances in technology and regulation (Wolf, 2013) 37-51.

[18] Justine Nolan, Wadhwa and Baumann-Pauly, 'A Regulatory Renaissance: The Role and Responsibilities of Internet Companies in Protecting the Right to Freedom of Expression Online (draft, not yet published).'

extended to seven days for more complex cases).[19] Indeed, self-regulation may be the only means of regulating online content considering the limited resources regulators can mobilise in addition to the sheer complexity of monitoring platforms with large amounts of users. However, such methods have come under criticism as the German government essentially 'outsourced the decision of what is lawful and what is not' to Facebook.[20] The US has adopted a different approach to regulating online content, with the 'Honest Ads Act'[21] targeting digital campaign advertisements by internet companies with at least 50,000,00 views. The focus of American legislation is transparency, with the act requiring companies to keep a public profile of all election communications purchased by a person who spent more than $500.

Overall, legislating on content hosted by internet platforms requires considerations into free speech and harm minimisation. The Australian government should create 'frameworks of accountability that require increased transparency of the definition of harmful content and actions taken to inhibit its dissemination.'[22] Practically speaking this may include additional disclosure on who is publishing the content, regular reporting of the amount of harmful content removed as well as the justification of removal.

Conclusion

The Commission needs to consider how ongoing technological changes can affect human rights. As can be seen in the case study, this may involve a balancing of rights.

[19] *Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken* (Netzwerkdurchsetzungsgesetz – NetzDG) (NetzDG), Paragraph 3 Umgang mit Beschwerden ueber rechtswidrige Inhalte <https://www.gesetze-im-internet.de/netzdg/__3.html>.
[20] Katrin Bennhold, 'Germany acts to tame Facebook, learning from its own history of hate' *New York Times* (19 May 2018) < https://www.nytimes.com/2018/05/19/technology/facebook-deletion-center-germany.html> accessed 21 May 2018.
[21] S. 1989 *Honest Ads Act* (US).
[22] Justine Nolan, Wadhwa and Baumann-Pauly 'A Regulatory Renaissance: The Role and Responsibilities of Internet Companies in Protecting the Right to Freedom of Expression Online (draft, not yet published).

Question 5) How well are human rights protected and promoted in AI-informed decision making? In particular, what are some practical examples of how AI-informed decision making can protect or threaten human rights?

Artificial intelligence as a category is not necessarily the most useful lens for understanding the types of tools that may be employed. Particularly when it comes to understanding the ethical implications and risks of deploying particular tools, a far more fine-grained analysis is necessary. The real question is not whether AI is involved but rather what the tool does and whether it is appropriate in the context in which it is being deployed.

For example, pre-programmed logic can be built into a system for decision-making. This can guide a decision-maker through the logic inherent in a piece of legislation or a government policy. The logic for such systems is programmed into the system – the intelligence really lies in the human programmer (and others that they may consult) rather than the system itself. Nevertheless, such tools can replace human decision-making in specific instances. It is only as good as the logic built into it (as RoboDebt demonstrates).

Such tools are quite different from those that rely on patterns and trends in historic data, as is the case for machine learning tools. Here, the logic is not entirely pre-programmed in but rather "learnt" by the algorithm based on patterns in how humans identify discoverable documents in a "training" data set. So, for example, a predictive policing program may "learn" that crime is most likely to occur in particular locations at particular times. Not all machine learning techniques work in the same way. They vary along several dimensions. An algorithm may be unsupervised, meaning it detects clusters and patterns in a data set that has not been subjected to classification by a human. An algorithm may be more or less able to give "reasons", comprehensible by humans, for classifications it makes or clusters it identifies. It may give more or less weight to outliers (that have an unusual classification, for example). It may prefer false positive or false negatives (to varying extents), or weight them evenly. There are even approaches that remove discriminatory impacts. Particular algorithms may have particular properties and be more useful at some tasks compared to others. The overall point here is that tools need to be assessed at the micro-level (of a particular tool used in a particular context) rather than at the macro-level of "artificial intelligence".

We have done some work that looks at "artificial intelligence" in the particular context of law enforcement and legal decision-making.[23] One area in which regulation should be considered is the question of whether there should be requirements for transparency of tools used in particular sensitive contexts.[24] An example of the problematic use of legal tools occurred in the case of *Wisconsin v Loomis*.[25] In this case, risk assessment tools, which relied on machine learning to predict the risk that an individual would re-offend based on

[23] See Lyria Bennett Moses and J Chan, Algorithmic prediction in policing: assumptions, evaluation and accountability, (2018) 28(7) *Policing and Society* 806-822, http://www.tandfonline.com/doi/full/10.1080/10439463.2016.1253695, Lyria Bennett Moses and J Chan, "Using Big Data for Legal and Law Enforcement Decisions: Testing the New Tools" (2014) 37(2) *University of New South Wales Law Journal* 643, Lyria Bennett Moses, Artificial intelligence in Legal Practice, Academia and the Courts: Understanding the Implications' (2017) 91(7) *Australian Law Journal* 561.
[24] Lyria Bennett Moses and L de Koker, 'Open Secrets: Balancing Operational Secrecy and Transparency in the Collection and Use of Data for National Security and Law Enforcement Agencies' (2017) 41(2) *Melbourne University Law Review* 530).
[25] 881 NW 2d 749 (Wis, 2016)

similarities with other offenders who had done so, were used by a trial judge in sentencing. Specifically, the circuit court had stated in the context of sentencing:

> *You're identified, through the COMPAS assessment, as an individual who is at high risk to the community. In terms of weighing the various factors, I'm ruling out probation because of the seriousness of the crime and because your history, your history on supervision, and the risk assessment tools that have been utilized, suggest that you're extremely high risk to reoffend.*

Neither the defendant in that case nor the primary judge were given the opportunity to access the algorithm or the data from which it drew, due to the COMPAS algorithm being a trade secret of Northpointe, Inc. Not only does this raise natural justice concerns in a particular case, it also makes it more difficult to detect when such algorithms have differential impact on particular communities. For example, ProPublica found that the COMPAS algorithm discriminates against African Americans, in the sense that there is a significantly higher probability that they will be a "false positive" compared to the general population. Without access to the algorithm and the data on which it relies, it is difficult to determine the reasons for such discrimination, or the extent to which it could be remedied through deployment of a different machine learning algorithm.

While non-transparent machine learning algorithms may be able to achieve higher levels of predictive accuracy in some circumstances compared to non-transparent tools, there will be some circumstances where their use is not appropriate due to the risks of unfair discrimination and/or inappropriate reliance in decision-making. In the example of COMPAS, it is not clear that a defendant *ought* to have a sentence affected by personal characteristics that correlate with "dangerousness" among the general population. For example, if people whose parents have divorced when they were young are more likely to go onto commit multiple crimes, does that mean that a defendant whose parents divorced should be sentenced more harshly based on that fact? Only by understanding the way a particular tool works can judges and lawyers retain the opportunity to challenge the inappropriate use of particular tools. As a starting point, if information would not be admissible as evidence on sentencing, it shouldn't be indirectly admitted through a machine learning inference.

More broadly, it is arguable that there is a human right to transparency in automated decision-making that affects an individual in important ways. Even more broadly, it may be appropriate to develop an ethical framework for the deployment of machine learning or data-driven decision-support methodologies by government. This should take into account rule of law values such as predictability and consistency, transparency and accountability, and equality before the law (the subject of a current project).

To date, evolution of 'artificial intelligence' tools is associated with a decline in transparency. As data analysis techniques become more sophisticated, it has become increasingly difficult for the system to 'explain' its conclusions in a human-interpretable way. The technical sophistication of modern-day computer systems can be seen through neural networks. Neural networks are seeded by a small amount of pre-thought meta-data, such as labels and relationships assigned by the human creators; thereafter, the process is entirely empirical in the sense that it is based on mass amounts of data being processed to identify correlations.[26] Whereas the

[26] See Roger Clarke, 'A Contingency Approach to the Software Generations' Database 22, 3 (Summer 1991) 23 - 34, PrePrint at http://www.rogerclarke.com/SOS/SwareGenns.html.

previous generations of computing 'still involved humans to express a model of the problem-domain, neural networks generated its own implicit model.' Thus, the inferences drawn by using neural network technologies are literally inscrutable,' making accountability and transparency in decision making very difficult to achieve.

Conclusion

As computing technologies have become more sophisticated, the reasoning behind their inferences have become more opaque. The Commission should seek ways to make AI-informed decision making more transparent and accessible. The need to make AI-informed decision making transparent is made more urgent as governments start using AI-informed decision making technologies in areas such as sentencing.

Question 7) In addition to legislation, how should Australia protect human rights in AI-informed decision making?

As data science begins to be applied in business and government, questions are being asked. Auditors, public relations executives and Board Directors want to be satisfied that negative impacts have been recognised and mitigated and that risks have been managed. Proponents of data analytics need to be able to show that their projects will not cause data subjects harm or infringe upon their right to privacy.

In recent publications, Clarke has established a framework for ensuring that data analytics projects are undertaken responsibly. It comprises 'Guidelines for Responsible Data Analytics', comprising a checklist of Do's and Don'ts, mapped onto a business process model in order to identify the appropriate point at which each of the checks is most usefully performed.[27] It is recommended that the Commission look through the checklist which encompasses guidelines on areas including data acquisition, data analysis and use of inferences. Such checklist can be applied to broad range of data-intensive technologies. This could be a useful reference point for best practice or as an industry standard.

[27] Roger Clarke 'Guidelines for the Responsible Application of Data Analytics' Computer Law & Security Review 34, 3 (May-Jun 2018) 467- 476, https://doi.org/10.1016/j.clsr.2017.11.002, PrePrint at http://www.rogerclarke.com/EC/GDA.html

Question 8) What opportunities and challenges currently exist for people with disability accessing technology?

The Hub and its affiliated researchers has also conducted a report on digital technology initiatives seeking to engage migrant workers and other low waged workers, as well as digital platforms designed to facilitate migrant workers' access to justice.[28] Adopting a worker-centred lens, the Report critically analyses the risks to users of the various digital platforms and the challenges confronting developers who seek to improve conditions for workers through the use of technology. It considers a range of practical, ethical, and legal challenges, many of which are generalizable to digital tools developed for vulnerable individuals beyond the migrant worker context and relate to issues outlined in the *Human Rights and Technology Issues Paper*. These include, for example, factors that determine the effectiveness of digital tools in terms of yielding clear outcomes for vulnerable individual users; privacy and security risks as well as defamation risks to vulnerable individual users; and challenges in design and implementation to ensure accessibility and uptake by vulnerable individuals.

Further research has been conducted on the potential of e-governance and the digitisation of migrant recruitment as promising means to facilitate safe and responsible labour migration and reduce forced labour and exploitation.[29]

[28] Bassina Farbenblum, '*Transformative Technology for Migrant Workers: Opportunities, Challenges and Risks*' available online at <https://www.mwji.org/>.
[29] See further, Bassina Farbenblum and Justine Nolan, 'The Business of Migrant Worker Recruitment: Who Has the Responsibility and Leverage to Protect Rights?' (2017) 52 *Texas International Law Journal* 1, Bassina Farbenblum, 'Governance of Migrant Worker Recruitment: A Rights-Based Framework for Countries of Origin' (2017) 7 *Asian Journal of International* Law 152, Bassina Farbenblum, Laurie Berg and Angela Kintominas, *Transformative Technology for Migrant Workers: Opportunities, Challenges and Risks* (Open Society Foundations, September 2018).