Horizon Scanning Series

The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve Our Wellbeing

Artificial Intelligence and Indigenous Data Sovereignty

This input paper was prepared by Maggie Walter and Tahu Kukutai

Suggested Citation

Walter, M and Kukutai, T (2018). Artificial Intelligence and Indigenous Data Sovereignty. Input paper for the Horizon Scanning Project "The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve Our Wellbeing" on behalf of the Australian Council of Learned Academies, www.acola.org.

Artificial Intelligence and Indigenous Data Sovereignty

Professor Maggie Walter, University of Tasmania
Professor Tahu Kukutai, University of Waikato

Introduction

Artificial intelligence (AI) is a rapidly growing and increasingly pervasive feature of societal functioning. For better or for worse, AI is now part of the everyday lives of Māori and Aboriginal and Torres Strait Islander peoples, albeit often operating in ways that lack transparency. Recent sector reports have highlighted the benefits, and in particular the economic value, that AI innovation and adoption can bring (AI Forum of New Zealand, 2018;). While increasing efficiencies and the development of beneficial technologies can produce positive outcomes, the marginalised social, cultural and political location of our respective peoples suggest that we will not share equally in these. We are unlikely to see, for example, the immediate benefits of precision diagnostics and AI-assisted surgery in the stretched public systems where most of our Indigenous populations receive health care. The considerable risks embedded in the ubiquity of AI are also unevenly distributed, and there are significant challenges for Māori and Aboriginal peoples relating to bias, stigma and accountabilities. In this paper we discuss the potential unforeseen (and likely unseen) negative consequences of AI for both peoples. We also consider how Indigenous data sovereignty, as an emerging site of science and activism, can mediate the risk of harmful outcomes while providing pathways to collective benefits.

Algorithms and AI functions: Whose Reality do they Reflect?

The conceptualisation of the term 'AI' is very broad but is defined here as comprising the 'advanced digital technologies that enable machines to reproduce or surpass abilities that would require intelligence if humans were to perform them' (AI Forum of New Zealand, 2018). Here we focus on Machine Learning (ML), Machine Reasoning (MR) and the use of predictive analytics within social and cultural projects. Algorithms are replacing or augmenting human decision making across a rapidly expanding range of contexts, from credit card lending and consumer services to child welfare and judicial sentencing (O'Neil, 2016). In Australia and Aotearoa NZ, governments are using algorithms and tools such as Predictive Risk Modelling (PRM) in a wide variety of frontline services, motivated in part by a desire to reduce costs through targeting those most ‘at risk’ of bad outcomes, while meeting targets for service delivery (Keddell, 2015). Despite a growing call for transparency and accountability in machine-driven decision-making (Lepri et al. 2017), the logic underlying algorithms is rarely accessible to the communities that they affect (Eubanks, 2018).

Whether motivated by profit or the drive to address pressing societal problems, the appeal of predictive analytics can be partly attributed to the aura of objectivity around AI and the abstraction of algorithms from the people and issues that their models describe. But, as several high-profile authors have persuasively argued (Eubanks, 2018; O’Neil, 2016), algorithms operate on given inputs and the learner algorithm and its functions are designed by human actors. The data used in ML and MR is also a socio-cultural artefact that is the product of very human subjectivities (Walter & Andersen, 2013). Whether algorithm designers or data generators, these technical experts are not neutral entities and neither are the learnings they produce. The construction of algorithmic rules always involve choices about which assumptions are incorporated and which are not. How those choices fall is fundamentally linked to the epistemic and ontological realities of algorithm designers and data generators. In short, AI rules resemble their creators in terms of their prioritisation of knowledge holders and sources, and their perspective of how the social and cultural world operates. In the vast majority of cases those creators are not Aboriginal and Torres Strait Islander or Māori (Kukutai & Walter, 2015).

Unforeseen and Unseen Negative Consequences of AI for Indigenous Peoples

There is growing evidence that racial biases in algorithms can have very real negative impacts and unintentionally entrench, rather than adjudicate, existing inequalities. Cossins (2018), for example, cites five examples from the United States where the specific logics of artificial intelligence had resulted in prejudicial outcomes. These include: likelihood in reoffending in sentencing guidelines that predict higher rates of recidivism in African American offenders; policing algorithms predicting crimes locations being more likely to target areas where minorities are located; facial recognition software with high accuracy for White men, but not for women or darker skinned people of both genders; and advertising target software shown to be more likely to target male candidates for executive positions. These unequal outcomes are unlikely to have been intended by algorithm writers. As Bornstein (2017) argues, such results come from a combination of unrecognised assumptions underlying the algorithmic construction combined with place-based inequality. Racial minorities are much more likely to live in poorer, heavily disadvantaged areas with relatively few services, making such biases almost guaranteed in predictive results. In Aotearoa New Zealand and Australia our respective populations are the poorest group, carrying the heaviest burden of disease, over-incarceration and broad spectrum inequality. This shared positioning is not co-incidental. Rather, it is directly related to our history as the colonised and dispossessed Indigenous peoples of our lands and the related ongoing integrational impacts of social, cultural and political marginalisation. Yet as Bornstein (2017) states, algorithms do not understand social, historical contexts. The youthful Indigenous demographic profile also contributes to the likelihood of being targeted with young people more ‘at risk’ of interactions with the justice system, unemployment (Jackson, 2002) and child welfare.

In many respects the profiling of Indigenous populations and the targeting of services is not new; surveillance by the state, its institutions and agents has long been an enduring characteristics of colonialism (Berda, 2013). What is new in the social policy arena are the opaque, complex and increasingly automated processes that shape targeting and profiling (Henman 2018). As ‘data subjects’ (Van Alsenoy et al. 2009), Indigenous peoples are included in a diverse range of data aggregations, from self-identified political and social groupings (e.g., tribes, ethnic/racial group), to clusters of interest defined by data analysts and controllers. In the latter case, those identified are often completely unaware of their assigned status and the associated implications. This raises the important point that Indigenous identifiers need not be explicitly included in algorithms for Indigenous peoples to experience the disproportionate impacts of algorithm-informed decision-making. For example, a study using PRM to predict child maltreatment in Aotearoa/NZ excluded ethnicity as it added little explanatory power to the models once socioeconomic risk factors were accounted for (Vaithianathan, Maloney et al. 2013). However, Māori children were much more likely to be exposed to the risk factors associated with maltreatment, reflecting the aforementioned inequities in access to the determinants of wellbeing. More generally, Māori and Aboriginal and Torres Strait Islander children and their families are disproportionately affected by the use of potentially biased algorithms in child protection. In Aotearoa New Zealand more than half of children in state care are Māori even though they only comprise one quarter of the child population (Office of the Children’s Commissioner, 2015). In Australia, Aboriginal and Torres Strait

Islander are nearly seven times as likely to be in state care as non-Indigenous children (AIFS 2017). The complex relationships between structural inequalities, ethnicity, patterns of system contact and system bias are still not well understood (Keddell & Davie, 2017). Marked spatial differences in child protection substantiations relative to notifications suggests system bias is one of several explanatory factors at work. There are other signs of bias. For example, the overrepresentation of Māori children increases at each decision point within the child protection system, with 40 per cent of children notified being Māori, increasing to 60 per cent by the time decisions to remove children into foster care are made (Keddell & Davie, 2017).

The prejudicial outcomes of discriminatory policies have been unwound in both countries to some extent by Indigenous activism and social justice movements over many years. With social problem identification decision-making now increasingly deferred to algorithms the likelihood of Indigenous linked injustice reworking its way back into the system rises, exponentially. Reworking the old adage around data: if the algorithm data 'rules' target problems where Indigenous peoples are over-represented; then the problematic Indigene will be the target. And unlike in earlier times, we cannot lobby, write letters to, or protest outside of an algorithm. The 'hands off' myth of AI acts as justification for social intervention injustice.

Mediating the Well-Being Risks: Indigenous Data Sovereignty

AI and machine learning are data driven which rely on ongoing access to data. But data are not just a free-floating phenomena. While increasingly electronic in form they have an underlying reality. They also have a tangible value as a cultural, strategic, and economic resource that is only growing with the advent of AI and other data technologies. So who owns the data? How should it be used? Who should have access to the data and under what circumstances? And who makes the decisions about the ownership, use, control and access to data and its value?

These questions have been of increasing concern and interest for Indigenous peoples around the globe. One response has been the emergence of the Indigenous Data Sovereignty (ID-SOV) movement. Indigenous Data Sovereignty is concerned with the rights of Indigenous peoples to own, control, access and possess data that derive from them, and which pertain to their members, knowledge systems, customs or territories (Kukutai & Taylor 2016; Snipp 2016). ID-SOV is supported by Indigenous peoples' inherent rights of self-determination and governance over their peoples, country (including lands, waters and sky) and resources as described in the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP). Implicit in ID-SOV is the desire for data to be used in ways that support and enhance the collective wellbeing of Indigenous peoples. In practice that means Indigenous peoples need to the decision-makers around how data about them used or deployed, including within social program algorithms.

ID-SOV movements are active in Aotearoa NZ and Australia and are grappling with the complexities of Indigenous data usage in AI. In Australia, the Maiam nayri Wingara Indigenous Data Sovereignty Collective, in partnership with the Australian Institute of Indigenous Governance, issued a communique from a 2018 national meeting of Aboriginal and Torres Strait Islander leaders. This communique stated the demand for Indigenous decision and control of the data ecosystem including creation, development, stewardship, analysis, dissemination and infrastructure. In Aotearoa New Zealand the Te Mana Raraunga Māori Data Sovereignty Network Charter[1] asserts Māori rights and interests in relation to data and requires the quality and integrity of Māori data and its collection. Māori have often been at the sharp end of intrusive data surveillance and misuse but have well-tested ‘tikanga’ (ethics, processes, principles) around the protection and sharing of knowledge for collective (versus individual) benefit. Groups like Te Mana Raraunga are exploring ways that tikanga can be used to rethink scientific approaches to data governance, use and validation. In a country that aspires to be a ‘world leader in the trusted, inclusive and protected use of shared data’ (New Zealand Data Futures Forum, n.d.), issues relating to ethics, trust and confidence are both timely and critical. For advocates of Māori data sovereignty, the goal is not only to protect Māori individuals and communities from future harm and stigma, but also to safeguard Māori knowledge and intellectual property rights, and to ensure that public data investments create benefits and value in a fair and equitable manner that Māori can fully share in.

Conclusion

The potential of AI to provide benefit to Māori and Aboriginal and Torres Strait Islander peoples is not able to be fully developed in this paper. However, such

[1] The TMR Charter can be accessed here: https://www.temanararaunga.maori.nz/tutohinga/

potential exists. In Aotearoa New Zealand, for example, AI is being used for language revitalisation with tribal radio stations Te Hiku Media creating language tools that will enable speech recognition and natural language processing of Te Reo Māori (Collins 2018). In Australia, Aboriginal technology entrepreneur Mikaela Jade is using augmented and mixed reality technologies to tell stories on country in Indigenous communities (Powell 2018). Both these examples also highlight the essential message of Indigenous Data Sovereignty; that harnessing the potential of AI for Indigenous peoples in Australia and Aotearoa New Zealand is closely aligned with Indigenous leadership and Indigenous governance on the processes of how, when and in what circumstances these technologies are applied.

References:

AIFS (2017) ‘Child protection and Aboriginal and Torres Strait Islander children’, *CFCA Resource Sheet— August 2017.* Australian Institute of Family Studies https://aifs.gov.au/cfca/publications/child-protection-and-aboriginal-and-torres-strait-islander-children

Artificial Intelligence Forum of New Zealand. (2018) *Artificial Intelligence: shaping a future New Zealand*. Retrieved from https://aiforum.org.nz/wp-content/uploads/2018/07/AI-Report-2018_web-version.pdf

Berda, Y. (2013) Managing dangerous populations: colonial legacies of security and surveillance. *Sociological Forum, 28* (3), 627.

Bornstein A. (2017) Are Algorithms building the new infrastructure of racism?
How we use big data can reinforce our worst biases—or help fix them. *Nautilus*, Dec 21. http://nautil.us/issue/55/trust/are-algorithms-building-the-new-infrastructure-of-racism

Bottou, L. (2014). From machine learning to machine reasoning. Machine Learning, 94(2), 133-149.

Capatosto, K. (2017). *Foretelling the future: a critical perspective on the use of predictive analytics in child welfare.* Columbus, OH: Kirwan Institute for the Study of Race and Ethnicity, Ohio State University.

Collins, M (2018) Te Hiku Media project teaching machines to speak te reo Maori, New Zealand Herald 24 February 2018. https://www.nzherald.co.nz/northern-advocate/news/article.cfm?c_id=1503450&objectid=11998971

Cossins, D. (2018) Discriminating algorithms: 5 times AI showed prejudice. *New Scientist* https://www.newscientist.com/article/2166207-discriminating-algorithms-5-times-ai-showed-prejudice/

Eubanks, V. (2018). *Automating inequality. How high-tech tools profile, police and punish the poor.* New York: St Martin's Press.

Henman, P. (2018): Of algorithms, Apps and advice: digital social policy and service delivery. *Journal of Asian Public Policy*, DOI: 10.1080/17516234.2018.1495885

Jackson, N. (2002). The doubly-structural nature of Indigenous disadvantage: a case of disparate impact? *New Zealand Population Review, 28*(1), 55-68.

Keddell, E. (2015). The ethics of predictive risk modelling in the Aotearoa/New Zealand child welfare context: child abuse prevention or neo-liberal tool? *Critical Social Policy, 35*(1), 69-88.

Keddell, E. (2018). Substantiation decision-making and risk prediction in child protection systems. *Policy Quarterly, 12*(2), 46-56.

Keddell, E. & Davie, G. (2018). Inequalities and child protection system contact in Aotearoa New Zealand: developing a conceptual framework and research agenda, *Social Sciences,* DOI:10.3390/socsci7060089

Kukutai, T. & Taylor, J. (2016) Data sovereignty for Indigenous peoples: current practice and future needs. In T. Kukutai and J. Taylor (eds.) *Indigenous data sovereignty: toward an agenda* (pp. 1-24). CAEPR Research Monograph, 2016/34. ANU Press. Canberra. https://press.anu.edu.au/publications/series/centre-aboriginal-economic-policy-research-caepr/indigenous-data-sovereignty

Kukutai, T. & Walter, M. (2015) Recognition and indigenizing official statistics: Reflections from Aotearoa New Zealand and Australia. *Statistical Journal of the IAOS 31*(2), 317– 326.

Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2017). Fair, transparent, and accountable algorithmic decision-making processes: the premise, the proposed solutions, and the open challenges. *Philosophy & Technology*. https://doi.org/10.1007/s13347-017-0279-x

Maiam nayri Wingara Indigenous Data Sovereignty Collective and the Australian Institute of Indigenous Governance (2018) *Indigenous Data Summit Communique*, 20 June, Canberra. Retrieved from https://maiamnayriwingara.org

New Zealand Data Futures Forum (n.d.). *Harnessing the economic and social power of the data*. Retrieved from http://datafutures.co.nz/assets/Uploads/Data-Futures-FORUM-NZDFF-harness-the-power.pdf

Office of the Children's Commissioner. (2015). *State of care 2015*. Retrieved from http://www.occ.org.nz/assets/Publications/OCC-State-of-Care-2015.pdf

O'Neil, C. (2016) *Weapons of math destruction.* New York: New York Crown.

Powell. D. (2018) 'How Mikaela Jade built augmented reality startup Indigital from deep in Kakadu National Park'. *Smart Company*, Monday, March 12, 2018 https://www.smartcompany.com.au/startupsmart/news-analysis/mikaela-jade-augmented-reality-startup-indigital-kakadu-national-park/

Snipp, M. (2016) What does data sovereignty imply: what does it look like? In T. Kukutai and J. Taylor (eds.) *Indigenous data sovereignty: toward an agenda* (pp. 39-56). CAEPR Research Monograph, 2016/34. ANU Press. Canberra.

Stats NZ. (2018) Integrated Data Infrastructure. Retrieved from http://archive.stats.govt.nz/browse_for_stats/snapshots-of-nz/integrated-data-infrastructure.aspx

Te Mana Raraunga Māori Data Sovereignty Network (2018) https://www.temanararaunga.Māori.nz/tutohinga/

TVNZ (2018) Free te reo language software developer upset about multinational's forays

Vaithianathan, R., Maloney, T., Putnam-Hornstein, E. & Jiang, N. Children in the public benefit system at risk of maltreatment. Identification via predictive modelling. *Am J Prev Med, 45I(3), 354-359.*

Van Alsenoy, B., Ballet, J., Kuczerawy, A., Dumortier, J. (2009). Social networks and web 2.0: are users also bound by data protection regulations? *Identity in the Information Society, 2*(1), 65-79. doi: https://doi.org/10.1007/s12394-009-0017-3

Walter, M. & Andersen, C. (2013) *Indigenous statistics: a quantitative methodology*. New York: Routledge.