Māori Perspectives on Trust and Automated Decision-Making

Report for the Digital Council

Kiri West, Daniel Wilson, Ari Thompson, Maui Hudson

13 November 2020

Contents

Executive Summary

The following report is a presentation of findings from a literature review and expert wānanga looking at Māori perspectives on trust and automated decision making. The research contributes to the Digital Council's Work Programme for 2020 which included a focus to technological issues facing New Zealand relating to trust. This project explores understanding the trust Māori have in automated decision making technologies and how this influences perceptions of the harm and benefit of technology.

The project was conducted in two parts. The first (Part A) was a literature review which included a synthesis of existing literature relevant to Māori perspectives on trust, Māori data sovereignty, and Māori perspectives on technology. The key findings were:

- There is an emerging body of literature about Māori Data Sovereignty
- Primary focus of data access and governance
- Comments on appropriate use and analysis but limited discussion of algorithms and ADM at this stage
- Generally reflects concerns expressed by minorities in other jurisdictions

The second part (Part B) was an expert wānanga where we brought together Māori experts to consider their views on trust and automated decision making. The key findings were:

- There is a need to consider the way that the Crown is framing 'trust' and ensure that it does not fall into a deficit thinking approach
- Data cannot be separated from algorithms and algorithms can't be separated from the process/system within which they operate
- Meaningful participation and partnership is critical in the development of governance mechanisms around ADM

Insights emerging from the expert wānanga informed the recommendations outlined below which are intended to give effect to Te Tiriti o Waitangi and promote greater transparency about the use of ADM as it affects Māori communities.

1. Build Māori Data and Digital Capacity within both Māori communities and across networks of Māori practitioners.
2. Develop robust equity assessment protocols for algorithms.
3. Ensure meaningful Māori participation in Institutional algorithm self-assessment processes.
4. Support collaborative partnership in project governance and the development and use of algorithms.
5. Create a Māori values framework and tikanga guidelines to support ADM design, development, use and maintenance.
6. Explore Te Ao Māori use-cases involving te reo Māori, tikanga Māori, and mātauranga Māori in ADM

Action was seen by the Māori experts as time critical because ADM systems are impacting Māori communities now.

Part A: Literature Review on Māori perspectives on Trust and Automated Decision Making

Data are the single most significant asset shaping our present and future realities. Data are driving national and global economies, and are presented as the evidential basis for the development of policies; they are framing political landscapes and radically transforming what it means to live in a democratic state. Underlying the datafication of our common realities is a persistent rhetoric that data are objective and free from bias. These claims of neutrality filter through commonplace justifications for the use of automated decision making (ADM) technologies, including algorithms. So much so that ADM's are now being elevated as a mechanism for the removal of human bias in important decision making.[1]

ADM offers the potential for greater efficiency, particularly in sectors where there is a demand for decisions to be made quickly. For example, in the face of the health care system potentially being overwhelmed by Covid-19 patients requiring ICU beds and ventilator support, the 1000minds decision-support system was configured to provide prioritisation recommendations for these patients based primarily on clinical criteria (1000minds, 2020). Aside from efficiency, algorithms are also considered useful for making decisions quickly where there is a high volume of cases; this is the case in New Zealand with ACC where an algorithm is used to make decisions in the claims process. Where the decision is straightforward, the algorithm can reduce the wait time from weeks in some cases to seconds; in instances where the decision is not straightforward, the decision is deferred to a human to process (Accident Compensation Corporation (ACC), 2017; Health Informatics New Zealand (HINZ), 2019). Another argument in favour of the use of ADM tools is that they are capable of processing large data-sets and learning from the decisions being made (Blackmore, 2020). The use of RoC*RoI (risk of recidivism * risk of imprisonment) in Aotearoa is an example of this as it holds the criminal histories of over 130,000 offenders and uses weighted criteria derived from all of these profiles in each decision it makes - there is no way that a person, or even team of people would be capable of processing that much information in such a short period of time as an algorithm can (Blackmore, 2020).

While there are certainly many potential benefits to using ADM, there is also significant potential for them to (re)produce harm for Māori. Primarily because, technologies rely on the availability of data to inform their processes. It will be brought to light in subsequent sections of this literature review that there are grounds for concern among Māori about the quality of data that presently exists *about* us within the system, and therefore there are concerns around the decision making tools that these data-sets inform. Given the current investment at the government level into investigating the potential use of ADM's, it is both timely and necessary to consider how Māori will be impacted by these developments.

[1] Consider, for example, the following quote justifying the use of the ROC*ROI algorithm: "The Department of Corrections' rationale for pursuing a statistical method of risk assessment, rather than continuing to rely solely on the professional judgements of probation officers and psychologists, was that statistical calculations of risk are more accurate than professional human judgement. [...] Professional judgement may also overestimate the level of risk because of biases based on stereotypical beliefs, including the risk of reconviction for people from different ethnic groups." (Waitangi Tribunal, 2005, p. 34)

Given that most of the ADM cases considered in the literature review involve government organisations, it is particularly important to acknowledge Te Tiriti rights and obligations and the United Nations Declaration on the Rights of Indigenous Peoples (United Nations, 2007). See for, example, the charter of Te Mana Raraunga, the Māori data sovereignty network, which affirms that "Māori data is subject to the rights articulated in the Treaty of Waitangi and the UN's Declaration on the rights of Indigenous Peoples, to which Aotearoa New Zealand is a signatory" (Te Mana Raraunga, 2016), and TMR's recent reiteration of these obligations specifically in relation to AI in a statement on the Department of Internal Affairs' recent procurement of a facial recognition system:

> Issues with the accuracy of facial recognition technologies, particularly amongst racialised minority populations, are widely documented. So too are the risks and impacts of implementing new technologies without sufficient engagement, assessment, controls, and governance. **If Aotearoa is to be a world leader in the trusted and ethical use of data, transparency and genuine engagement in decision-making is crucial. As a Crown agency, DIA has obligations to uphold the principles of te Tiriti o Waitangi and the principle of Free Prior and Informed Consent as set out in the United Nations Declaration on the Rights of Indigenous Peoples**. Māori data governance over Māori data is critical, and our rights to be involved in all decisions that affect us need to be upheld. (Te Mana Raraunga, 2020)

The above quote speaks to dominant principles recognised by the judiciary in relation to Te Tiriti obligations that are significant for ADM systems including: partnership, active protection, and redress (Waitangi Tribunal, undated). Similarly, Karaitiana Taiuru proposes the following Te Tiriti compliant guiding principles for artificial intelligence systems or systems that employ algorithms (Taiuru, 2020, p. 18-19):

1. Consultation, negotiation and free and informed consent are the foundations for systems with Māori data.
2. The responsibility for consultation and negotiation is ongoing. Consultation and negotiation are a continuous two-way process.
3. Consultation and negotiation should achieve mutual understanding about the proposed systems, data storage, machine learning and algorithms.
4. Māori knowledge systems and processes must be respected.
5. The intellectual and cultural property rights of Māori must be respected and preserved.
6. The negotiation of outcomes should include results specific to the needs of Māori.
7. Negotiation should result in a formal agreement for the conduct of a research project, based on good faith and free and informed consent.

As will be seen in what follows, these rights and obligations inform much of the expectations about what counts as safe and positive outcomes for Māori with respect to ADM systems.

ADM and trust

The widespread acceptance and use of algorithms and other ADM technologies relies heavily on whether or not people trust the systems. The digital council applied the following definition of trust:

> We are using the working definition that trust is about whether people are comfortable in a situation where they are vulnerable to the consequences of someone else's actions

Trust can be seen to be limited when the internal workings of an algorithm are not transparent (Quince, 2020). Generally speaking then, trust stems from whether or not the public understand the tools themselves (Ngomo, 2019), or at the very least, understand the basic process behind how an algorithm reaches a particular decision (Quince, 2020). This assessment from both Ngomo (2019) and Quince (2020) is mirrored in New Zealand's Algorithm Charter, which identifies transparency and accountability as critical for "...ensuring that the public can trust and support the government to use these tools in appropriate ways." (Stats NZ, 2020). In the absence of understanding, to generate trust amongst communities in Aotearoa New Zealand, the concept of social licence has been critical in communicating a mandate.

Trust, Social Licence, & Cultural License

At its core, social licence refers to the process of conferring permission to act, (Jenkins, 2018), usually in ways that would otherwise be seen to be operating outside of socially accepted norms (Gulliver, et al,2018). It is a concept which has risen in prominence in industries where there is potential for harm to occur, for example in extractive industries such as mining (Gulliver, et al, 2018) and as data is increasingly viewed in the same light (as extractive and as having significant potential for harm to occur) social licence is also being applied in this space. Most notably by Stats NZ in the development of the Integrated Data Infrastructure (IDI) (Moses, 2020). Importantly, since those operating under a social licence model are often participating in activities which sit outside of accepted social norms, the absence of sanctions is also critical (Gulliver, et al, 2018; Jenkins, 2018)

Conferring permission to act, and specifically, permission to act without sanction, requires high levels of trust among the communities which may be impacted. This presumption of trust limits the extent to which it may be deemed acceptable to continue to rely on social licence in the development and use of ADM's, from a Māori perspective. This is because, as this literature view has highlighted, Māori have heightened levels of distrust towards data and more specifically in the systems in which data operate.

In 2016, Stats NZ, working with the Data Futures Partnership, invested time and resources to measure the extent of their social licence and to develop models for evaluating social licence (Moses, 2020). The primary focus of this research was to consider what factors contribute to the comfortability of New Zealanders when it comes to data use.

> Trust in data use is an important part of social licence: when people trust that their data will be used as they have agreed and accept that enough values will be created, they are more likely to be more comfortable with its use. (Data Futures Partnership, 2017, p. 5)

Whether or not people have enough trust in our statistical infrastructure to infer social licence is currently contested, this literature review has highlighted that. Central to this contestation is that in order to have trust in a system, you need to understand how it works, therefore in the context of Aotearoa, social licence is limited when less that 25% of a surveyed population agree that they know Stats NZ reasonably or very well (Kukutai & Cormack, 2019). Further, there are additional layers of complexity for Māori because as TMR pointed out in their statement on social licence:

> ...we view **Social licence** as the ability of an organisation to use and share data in a legitimate and acceptable way, based on the trust that **individuals** have. (Te Mana Raraunga,2017)

Though it is important that individuals have a level of comfortability with data use, it is also critical that there is recognition of the fact that what can be considered safe use for individuals is not necessarily safe at the collective level. The NZ Police's national wastewater drug-testing programme is an apt example of this individual-collective data rights tension. As part of this programme ESR, tests community wastewater to detect the levels of methamphetamine and other illicit drug use (Kukutai, et al, 2020). Because the data cannot identify people at the individual level, the testing programme is compliant with privacy laws. However, the aggregated datasets do tell stories about the communities from which the data is derived, and in this case the programme can be seen as an extension of systems of surveillance into affected communities (Kukutai et al, 2020). This example highlights how a focus on social licence and trust at the individual level is not sufficient.

As an alternative to Social Licence, TMR (2017) proposed the notion of cultural licence which considers the sharing of data in a way which reflects the collective trust that iwi and Māori Treaty partners have (TMR, 2017; Data Futures Partnership, 2020). Cultural Licence is a particularly useful conceptual tool when thinking about data sets which can be aggregated and used to represent groups, particularly when considering the potential risks and benefits of data use (Data Futures Partnership, 2020).

> As the Māori Data Sovereignty Network, Te Mana Raraunga is committed to protecting and securing Māori rights and interests in data. Our view is that the proposed Guidelines for Social Licence should acknowledge the importance of Cultural Licence, the distinctive rights and interests of iwi/Māori as Treaty Partners, and Iwi/Māori aspirations to derive equitable benefits from data as a counterbalance to the significant collective risks. For Iwi/Māori to derive clear and equitable value from data use the guidelines would need to make provision for 'data for governance' (access to data for iwi/Māori decision making), and 'governance of data' (involvement in decisions about data access and use). (TMR, 2017)

Māori Data Sovereignty

Where mainstream definitions of data sovereignty focus on the jurisdictional control nation-states have over data, Indigenous data sovereignty (ID-Sov) asserts that the rights to control data remain with the nation from which the data originated (Kukutai & Taylor, 2016). This centres the rights of Indigenous peoples to retain control over data concerning their lifeways and territories irrespective of the geographic location of the data (Walter and Suina, 2019; Cormack, et al, 2020). Importantly, nationhood is conceptualised more broadly than traditional understandings of the 'nation-state' and also includes tribes and other kinship based collectives (Lovett et al, 2019). ID-Sov responds to the rapidly evolving challenges associated with global data environments which are increasingly open and integrated (West, et al, 2020). At the international level, these rights are supported by the United Nations Declaration on the Rights of Indigenous Peoples (Kukutai, 2018) and have been asserted as necessary by the UN Special Rapporteur on the Rights to Privacy (Cormack, et al, 2020). Conceptually speaking, the ID-Sov movement is relatively new, gaining traction primarily through the publication of the book *Indigenous Data Sovereignty: Toward an Agenda* in 2016. However, the central ideas forming the foundation of ID-Sov, have been the focus of critical engagement by Indigenous scholars and allies for decades. ID-Sov draws together relevant discourses relating to cultural and intellectual property rights, Indigenous research ethics and Indigenous governance (West, et al, 2020).

In the context of Aotearoa, Te Mana Raraunga: The Māori data Sovereignty Network (TMR) alongside the Data Iwi Leaders Group (Data-ILG) have been critical in asserting Māori data rights (Cormack,et al, 2020; Sporle, et al, 2020). Using the mana-mahi framework, outlined in the Te Mana Raraunga Charter (2016) TMR and the Data-ILG are able to focus their efforts respectively to support Māori data-sov rights (Sporle, et al, 2020). The Iwi Leaders' Group work in the Mana space and are actively engaging in the national policy developments (Sporle et al, 2020).One such policy development is the Mana Ōrite agreement.

Mana Ōrite Agreement:
The Mana Ōrite agreement is an agreement between StatsNZ and the Data ILG reflecting a Te Tiriti ō Waitangi derived relationship between the two groups (StatsNZ 2020e). Importantly, emphasis has been placed on how the agreement is a relationship - not a partnership - where both signatories (the Crown and the Data-ILG as Māori representatives) have equal explanatory power (Sporle, et al, 2020). The work programme for 2020 is made up of four key projects:

1. Examine and develop ways of addressing disproportionate effects for iwi of 2018 census results
2. Improve administrative data to ensure a sustainable and diversified flow of relevant iwi data for Māori
3. Develop a Māori data governance proposal

4. Develop a scope of work proposal for potential te reo Māori specific datasets (StatsNZ, 2020d)

The work programme aligns with the goals of the agreement which are to:

- Meeting the current and future data needs of Māori
- Develop stronger relationships between StatsNZ and whānau,hapū and iwi Māori
- Identify and address key data gaps for Māori
- Improve equity of outcomes
- Embed a te ao Māori lens in data decision making (Stats NZ, 2020e)

A key goal of particular relevance to the discussion around Māori and automated decision making is that 'iwi-Māori have improved access to iwi-Māori data and enhanced opportunities to co-create and co-develop future systems and data design across the public sector data eco-system' (StatsNZ, 2020e, p.2).

The work that the Iwi Leaders' Group does is significant in asserting Māori data governance rights. Contributing to the mahi space, TMR are actively engaged in "...defining the parameters of Māori Data Sovereignty by publishing definitions for Māori data, Māori Data Sovereignty, Māori data governance, and Māori Data Sovereignty principles' (Sporle et al, 2020, p.67). There are 6 principles of Māori Data Sovereignty (TMR, 2018) which lay out an ethical approach to data ecosystems for collective and distributed benefit (Kukutai, et al, 2020) and to enhance Māori wellbeing (Sporle, et al, 2020). These principles and their sub-principles are:

- Rangatiratanga / Authority
    - Control
    - Jurisdiction
    - Self-determination
- Whakapapa / Relationships
    - Context
    - Data disaggregation
    - Future use
- Whanaungatanga / Obligations
    - Balancing rights
    - Accountabilities
- Kotahitanga / Collective benefit
    - Benefit
    - Build capacity
    - Connect
- Manaakitanga / Reciprocity
    - Respect
    - Consent
- Kaitiakitanga / Guardianship
    - Guardianship
    - Ethics
    - Restrictions (Te Mana Raraunga, 2018)

The work of both TMR and the Data Iwi Leaders Group has provided clear pathways for Māori to be involved in setting the agenda for data sovereignty and data governance in Aotearoa (Sporle, et al, 2020). This is evident in the influence that the Māori Data Sovereignty principles have influenced agency responses to data sovereignty issues (Including the development of the Ngā Tikanga Paihere framework) as well as the aforementioned Mana Ōrite agreement.

Models of data governance

A central function of Indigenous data sovereignty movements has been the development of data governance models. Kamira (2003) describes governance as the processes drawn upon by delegated bodies to make decisions to suit the needs and values of relevant stakeholders. Broadly speaking then, data governance refers to the broad range of mechanisms implemented and adapted to encourage desirable behaviour in the use of data (Wende, 2007). These include, as Panian (2010) notes the"...processes, policies, standards, organization, and technologies required to manage and ensure the availability, accessibility, quality, consistency, auditability, and security of data in an organization" (p.139). Present in each layer of governance are what Bruhn (2014) refers to as agents of data governance; they are producers of data, users of data and governing actors. These categories are not mutually exclusive and it is possible that there may be multiples 'agents' present within each mechanism of governance.

At present, there is widespread recognition across academic and flaxroots communities that data sovereignty is a pressing issue. In saying this, it is likely that the framing of the issues sitting within the language of Indigenous Data Sovereignty is specific to academia and that at the community level there will be varying interpretations and experiences of the issues within depending on context. A leading scholar in the field of Indigenous data sovereignty, Maggie Walter along with colleagues identified the following agenda for IDS movements:

> *Globally, an early goal is for Indigenous leaders to agree on a set of data protocols, operating at the community, First Nations and national level that quantify and prescribe the parameters of Indigenous data governance* (Walter, et al, 2018, p.3)

Indigenous data sovereignty networks have been developing models for governance which look to protect communities from harmful or exploitative research in the first instance - but also secure protection of data in an ongoing capacity. One such model is OCAP™. OCAP™ is considered the de facto ethical standard for guiding both the ethical conduct of research using First Nations data as well as for the management of First Nations information (Bruhn, 2014). It is also regularly referenced in ID-Sov discourse as a pioneering document with regards to establishing a basis for models for governance that protect communities from harmful or exploitative research, but also create collective and individual benefit (Sporle, et al, 2020).

OCAP™ is an acronym for ownership, control, access and protection. Importantly these four principles of OCAP™ reflect the goals of information-sharing and information-governance from a First Nations perspective, within a frame of community privacy and collective well-

being. OCAP™ was developed out of the desire from First Nations communities in Canada to ensure the collective ownership and privacy (at the community level) of information derived from them (Bruhn, 2014)

Historic roots of mistrust

Before there is any direct engagement with issues relating to trust and ADM's specifically, there needs to first be recognition and acknowledgement of the reality that any distrust expressed by Māori towards ADM's is historically grounded. The whakapapa of distrust is rooted in a broader distrust of the systems in which ADM's are embedded in.

It is widely accepted in the literature that Māori are a critically over-researched and over-surveilled population, with very little by way of progress to show for their engagement with data-generating activities. Research though, has not historically been an activity designed with the interests of Indigenous peoples, including Māori, at heart. In the context of Aotearoa, Pool (2016) notes how research and data collection were part of Britain's broader 'civilising mission'. To this end missionaries utilised 'imported data methodologies' to gather information which would label Māori often in contradictory terms as murderers, brutes and noble savages (Pool, 2016). The 'research' activities of missionaries, referred to by Pool (2016) have informed a 'collective memory of imperialism', the effects of which are still visible today (Smith, 2012). As Linda Smith (2012) sums it up:

> The word itself, 'research', is probably one of the dirtiest words in the indigenous world's vocabulary. When mentioned in many indigenous contexts, it stirs up silence, it conjures up bad memories, it raises a smile that is knowing and distrustful. (p.1)

Aside from the lasting impacts of distrust among indigenous peoples towards research and data collection noted by Smith (2016), an equally devastating impact of research is the influence that it has had on the way that non-indigenous peoples see us. Over the decades, the data generated through research activities as well as in official statistics and administrative data *about* Māori has by and large continued to (re) produce deficit narratives.

It is important to understand what the present data records tell us about Indigenous populations, particularly across what are commonly referred to as the CANZUS states (Canada, Australia, New Zealand and the United States of America). The data held by these nation-states about their Indigenous peoples varies in breadth, depth and quality, however there is a common base across the board. What is consistent across the data-sets is a narrative of the statistical Indigene (Walter, 2016) who is marked by poor health and education, high levels of unemployment, homelessness and incarceration as well as increased risk of morbidity and subsequently younger mortality rates (Kukutai & Walter, 2017).

In some cases, access and visibility of such terrifying and dismal data-sets may motivate and inspire the kinds of changes required to radically shift the social positioning of Māori. For example, the knowledge of language loss brought to light in the 1970's inspired a Māori cultural renaissance which saw the creation of kohanga reo, kura kaupapa Māori and wharekura. What

is important to note here though is that the 'Māori cultural renaissance' was a movement driven from the flax-roots organisations including the Māori Women's Welfare League and Ngā Tama Toa. While the language regeneration story is a positive one, there is still the risk that the story dictated by data may in fact build apathy and identify cultures and communities as the root causes of the problems being experienced, as opposed to seeing the failures of racist systems. Unfortunately, this story is far too frequent.

For example, at the same time that language attrition data was coming to light, there were also reports coming out of the departments of Justice and Social Welfare which highlighted significant disparities existing within the system (See He Whaipaanga Hou, 1987 & Pu-ao te-ata-tu, 1988). In He Whaipaanga Hou, part 2, Jackson (1988) noted that the modes of gathering official information and statistics about Māori were flawed, and noted that:

> Unfortunately, the statistics are often used, in some research and the media, with apparent disregard for the difficulties involved. Instead, they are presented as unqualified statements about Maori offending and even, in some instances, as explanations rather than descriptions of the situation. Such a use is methodologically dishonest. It is also socially mischievous because the constant use of such statistics not only produce an erroneous perception of the extent of Maori offending, it also contributes to the creation of negative stereotypes about Maori behaviour in general. (p.19)

Despite having these reports which pointed to systemic racism as a contributing factor in the significant and increasing disparities between Māori and and non-Māori, it remains today, some 32 years later, that the social problems reported in 1987/1988 remain, and in some cases have worsened. Further, what the quote above also highlights, is that the problems embedded in official statistics have been raised by Māori historically, and yet the system is still riddled with issues. The 2018 census is a clear example of this. This failure on the part of successive governments to be responsive to the rights and interests of Māori in the face of such damning statistics, has created tension when there is a need to have heightened trust in experts (Cook, et al, 2020).

> Given the history of state data relations and practices for Māori, and the evidence surrounding algorithmic bias and injustice, Te Mana Raraunga call for a stronger approach to regulating government use of algorithms. A voluntary charter is unlikely to provide the necessary protections and safeguards for Māori in relation to government use of algorithms. Our main feedback, therefore, is that an 'Algorithm Charter' is insufficient to protect Māori rights and interests and is highly likely to fail. Such failure will not only be to the detriment of Māōri, but also puts at risk the trust in the public data system that is foundational to the willingness to participate and provide high quality data. Regulation that includes mechanisms for accountability and redress is necessary. Such regulation would need to include Māori data governance at all levels... (TMR Submission - draft algorithm charter)

Hēmi Whaanga states,

> AI will be a game changer that challenges the foundations of our knowledge systems. Thus, it is critically important that we envision and shape how AI could be part of a revolution that is productive for our knowledge systems, our languages, and our

futures. We need to be part of the dialogue on establishing global principles and standards for the use of AI to ensure that is not used to perpetuate societal biases, inequalities and global homogenization."(Whaanga, 2020)

Contexts for use of ADM

In what follows, we consider four contexts of use for ADM systems that are either currently in use in Aotearoa, or which have been trialed or proposed. Examination of these case studies raise numerous issues that might inform whether it would be reasonable for Māori to trust in ADM tools embedded in these systems given current practices.

ROC*ROI and WAI 1024

ROC*ROI was developed by the Department of Corrections to generate a score for the risk of reconviction of an individual within five years. The model was trained on data containing the histories of over 130,000 individuals who had been convicted of an offence. According to the Department of Corrections report, the resulting initial model weighted Māori ethnicity as a contributing predictor of recidivism (Bakker, O'Malley, Riley, 1999). This meant that simply being Māori would increase the value of the predictive score. Roc*Roi is used to inform decision-making in the following contexts (Waitangi Tribunal, 2007):

(1) ROC*ROI informs probation officer's pre-sentence report to the court
(2) After sentencing, the ROC*ROI score informs the kinds of interventions that will be received
(3) Informs the parole process

Generating a higher score which then informs these decision processes on the basis of one's ethnicity is unjust; this is clearly discriminatory against Māori. The injustice of discriminating by ethnicity resulted in the WAI 1024 Waitangi Tribunal claim by Tom Hemopo on behalf of Ngāti Kahungunu. Hemopo's claim quoted in the final report as follows:

> The allocation in the roc /roi of a higher score to all Maori offenders compared to nonMaori, based purely on ethnicity, is discriminatory. The roc /roi is drafted on the presumption that simply because an offender is Maori there is a higher risk that they will re-offend. (Waitangi Tribunal, 2005, p. 61)

The models provide no account of what the *cause* might be for Māori to more likely be convicted again, only that the cause(s) impacted Māori with respect to recidivism outcomes in a way that is quantifiable. This ethnicity variable *could*, for instance, be in part a measure of the effects of racism within the justice system and the impact of this on Māori. In this case, the ethnicity variable would be a proxy for racism in the system.

Upon investigation by the Department of Corrections, it was found that it was possible to reduce the effect of the ethnicity variable to zero by redistributing the weight to other variables. This was done in 2004. Here is the Waitangi Tribunal's interpretation of this action:

> From the evidence on this complex matter, we understand that the ethnicity variable would have remained in roc *roi but for the fact that the concerns raised about its negative connotations caused a re-examination of its particular contribution to the tool's predictive accuracy. It was then found that, because of the high correlation of ethnicity with other variables, the predictive accuracy of roc *roi could be maintained by recalibrating other variables and reducing the effect of the ethnicity variable to zero. This is what was done. Had it been found, however, that the effect of the ethnicity variable on roc *roi's accuracy could not be replicated by such changes, it would have remained. (Waitangi Tribunal, 2005, p 126)

The reduction of the weighting of the ethnicity variable was deemed sufficient by the Waitangi Tribunal to address this part of Hemopo's claim. In the tribunal's words: "we consider that the removal of the variable's effect answered most of the claimant's original concerns about roc *roi" (Waitangi Tribunal, 2005, p. 127).

However, re-weighting the other variables to absorb the effect of the original Māori ethnicity weighting in no way guarantees that this action addresses the underlying issues regarding predicted Māori recidivism outcomes. If the accuracy of the model remains the same--which the tribunal said was the case--then the modified weightings of the other variables simply act as proxies for (Māori) ethnicity. It is misleading to claim otherwise. Rather than addressing the underlying problem, the "fix" has simply masked the issue and made it more difficult to audit. Consequently, from a statistical point of view, it is unsurprising to see the following result after the ethnicity variable was removed:

> Mr Hemopo ran several roc *roi assessment tests and found that, when identical variables except for ethnicity were entered into the data set, Maori offenders continued to receive higher scores. Consequently, the reduction of the ethnicity variable to zero had not diminished Mr Hemopo's concerns about the roc *roi assessment. (Waitangi Tribunal, 2005, p. 62)

Māori lawyer Khylee Quince (2020) reminds us that, like in the initial implementation of roc*roi in Aotearoa, algorithms can include "factors that explicitly consider a characteristic, such as race or gender, which leads to problematic logic that connects being Māori with being criminal for example. However the removal of the explicit factor does not eliminate racism if other factors are effectively a proxy for race." (Quince, 2020)

Whare Hauora and the Kāinga Ora Smart Home sensor initiative

In 2019, Whare Hauora raised surveillance concerns about a Kāinga Ora pilot initiative to install smart home sensors in state house homes (Whare Hauora, 2019). Tenants in the pilot would agree that sensors could measure temperature, humidity, CO2, light and air pressure.

Some properties would also have electricity usage measured. The identifying information would contain the address of the home. Further, the Policy for Smart Homes stated that data may be made to MSD and Police, if requested ("under any enactment") (Housing New Zealand, 2019).

Whare Hauora's concerns centered around the lack of control of the data by tenants, but also in the potential for widespread surveillance of Kāinga Ora tenants. Surveillance concerns were downplayed by the Privacy Commissioner and Kāinga Ora (Keall, 2019), very similar concerns have been raised in the past. In one study on the use of similar sensors in an office environment, a surveillance concern was that "Data on [indoor environmental quality] retrieved by means of CO2 sensors, as well as data from the operation of the dimming lighting sensors, could provide information on the occupancy of offices and could be potentially used to control the presence of people in their working place." (Casconea et al. 2017) Indeed, there has been machine learning research specifically to estimate human occupancy based on data from CO2 sensors (Arief-Ang, Salim and Hamilton, 2017). In the current environment--without adequate protections--there is both plausible means and motive for government organisations who, for example, may be interested in fraud detection, to employ algorithms on sensor data to support decision-making. The dismissal of these issues suggests that plausible concerns regarding secondary use of data for machine learning applications are not always taken seriously.

Predictive risk modeling for child harm

In a recent article, Rhema Vaithianathan, director of the Centre for Social Data Analytics (CSDA) at AUT proposed that a predictive risk model for identifying children at high risk of child maltreatment currently in use in Allegheny County in the United States to inform referral judgements by child and protection services could be adapted for use in Aotearoa (Morton, 2020). This ADM system is the Allegheny County Family Screening Tool (AFST), which generates a score that is used to inform call screeners when deciding whether to refer a case for further investigation. This kind of proposal is not new. There was earlier research into the use of predictive risk models by the Ministry of Social Development (MSD) from 2012, however a proposed study to examine the performance of this model was stopped in 2015 by the Social Development Minister Anne Tolley (MSD, 2015; Jones, 2015; RNZ, 2015).

The initial MSD research into predictive risk models (2012-2015) included a number of reports regarding ethical issues for Māori in predictive risk modelling with specific reference to the identification of high-risk new-born pēpi.

Anton Blank, Fiona Cram, Tim Dare, Irene De Haan, Barry Smith and Rhema Vaithianathan co-authored a report for MSD that utilised the Te Ara Tika framework to generate recommendations based on the principles of tika (appropriate design), manaakitanga (cultural and social responsibility), whakapapa (relationships), and mana (justice and social equity). The framework describes three levels of relationships: "Minimum standard: Consultation"; "Good Practice: Engagement"; and "Best Practice: Kaitiaki" (Hudson, 2010, pp. 6-8). However, the authors state that "Only the minimum standard and good practice levels of each dimension are examined here" because the PRM "would be implemented by a government department rather than by Māori" and under these conditions the "best practice" level of Māori governance

and control is “unobtainable” (Blank et al., 2013, p. 6). The authors proposed the following recommendations:

- Māori need to be consulted on the benefits and risks of the tool and ways to mitigate the latter – especially the potential problems of hyper-vigilance and stigmatisation.
- Māori need to determine what the potential benefits of this tool might be and whether the (mitigated) risks outweigh the benefits given the manner in which the tool will be deployed.
- We believe that a gap analysis of the needs of vulnerable Māori families is required to ensure that a Māori-centred / whānau-centred screening/intervention pathway is developed.
- A Māori-centred/whānau-centred screening/intervention pathway should be developed
- As soon as more details on the deployment of PRM are available, we suggest that there be a systematic consultation with Iwi, in particular to provide them with opportunities to validate the model.
- The interventions need to be developed in consort with PRM to ensure that the programme does not codify structural issues as “parental deficits”.

Dr Cindy Blackstock from the First Nations Child and Family Caring Society of Canada commented on the 2013 version of the above ethical review. In contrast to the “good practice” considerations from the Blank et al. report, Blackstock asserts the importance of indigenous rights in answering any of the substantive questions of the ethical review of the predictive risk model: "Consistent with the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP) consideration of these and other questions relating to the model must be informed by the fundamental principle of Maori self-determination and free, prior and informed consent." (Blackstock, 2013, p. 4)

One of the critical points Blackstock raises is the relationship of the development of a PRM model to the development of any interventions, and the importance of Māori in the design of these co-related systems. She states:

> The identification and development of interventions may be slated as a secondary step, however, I would argue that it is so critical to PRM aims and to the Mana requirements that the predictor model and the intervention models should be developed simultaneously. This is particularly true given that existing child welfare interventions often codify structural issues as parental deficits for Indigenous peoples. In Canada, and perhaps in New Zealand as well, this false codification has resulted in the unfortunate situation where Indigenous families are actually held accountable for the manifestation of the very colonial policies that disadvantaged them. (Blackstock, 2013, p. 7)

Further, the model does not separate child neglect—which often is poverty related and has societal causes—from other forms of child harm. Blackstock claims that this unfairly places the blame and stigma on whānau for causes that are instead due to the impacts of inequities in a colonised society:

> The roots of poverty in Indigenous communities are often linked to historical and contemporary disadvantage and thus it is important that the PRM and other risk assessment tools distinguish between personal and societal locus of control to better target interventions and ensure that families are not held accountable for factors beyond their control. (Blackstock, 2013, p. 5)

In concluding, Blackstock notes the necessity of meaningful Māori engagement not just in PRM development but in the development of the services that these ADM tools support:

> Regardless of the constellation of intervention programs developed, it is essential that Maori are meaningfully engaged, that interventions build on Maori cultural strengths and target both individual and structural risk factors and that child welfare center poverty, substance misuse and housing inadequacy in child welfare legislation, policy and practice. (Blackstock, 2013, p. 8)

The above comments speak to how the design of ADM systems are inextricably influenced by the attitudes and assumptions embedded in the systems they are meant to support, and can themselves increase the scale and speed at which these attitudes and assumptions are applied. Systems that exhibit structural racism and that have a long history of negative outcomes for Māori are extremely unlikely candidates for safe ADM systems without radical system change, and certainly effective Māori involvement in governance. This suggests that there are contexts where “good practice” engagement is not sufficient, and “best practice’ relationships are most appropriate. I will finish this section with the conclusion of a recent report by the Children’s Commissioner into the practice of uplifting new-born pēpi by Oranga Tamariki, which reiterates concerns about racism in this system, and which intervention would include the proposed ADM in its causal pathway, if it were implemented:

> What we heard raises questions about the impact of the underpinning culture of the current statutory care and protection system. This leads us to question if the current system is creating an environment where unprofessional social work practice, racism and discrimination, distrust and fear of supports and services is leading to negative outcomes and harmful experiences for whānau and their pēpi. (Office of the Children’s Commissioner, 2020, p. 74)

ADM and the health sector

In the health sector there are risks that ADM tools developed to solve particular issues, for example, the allocation of ICU beds should the health system be overwhelmed with Covid-19 patients or else the expediting of elective surgery approval for “low risk” patients, risks exacerbating existing health inequities.

Take for example the nzRISK pre-operative mortality risk prediction model. This has been developed to identify the mortality risk of a patient within 30 days, one year or two years of undergoing one of a large number of non-cardiac surgeries (nzRISK, 2020). The model is trained on the data of over 360,000 participants from the New Zealand National Minimum Data Set for patients having surgery between January 2013 and December 2014 (Campbell et al,

2019). The model contains an explicit ethnicity variable, which has a weighting indicative of worse mortality outcomes for Māori than for those of European descent. A comparison of the variable weightings in the model shows that Māori ethnicity impacts the predictions of mortality from the model approximately the same as taking a patient of European of Asian descent and adding nearly 10 years to their age (Campbell et al. 2019, p. 1555). That is, a Māori patient will have approximately the same risk profile for a particular surgery as a patient of European descent with the same clinical conditions but ten years older.

This situation of worse health outcomes for Māori are not new. But let us now consider plausible, potential use cases for models like nzRISK and 1000minds that may use predictions of clinical outcomes to inform various judgements in the health system.

First, imagine the use case that the nzRISK score is used to expedite surgeries for patients that are deemed low risk. The goal is to assist in reducing the backlog of elective surgeries. However, there are already inequities in relation to Māori receiving specialist care in comparison to non-Māori. Consider the following comment from a recent report on the health system: Hospital appointments are not accessible for more Maori adults than non-Maori adults [….] Moreover, specialist appointments have unacceptably long wait times and occur less frequently for Maori" (Health Quality and Safety Commission New Zealand, 2019, p. 9). Fast-tracking those patients deemed low risk may likely exacerbate these negative outcomes when there is effectively the equivalent of a ten-year age penalty for mortality risk calculations for Māori patients.

Moving on to 1000minds, it should also be noted that clinical criteria were proposed for the configuration of the 1000minds system for ICU bed triage (1000minds, 2020). The decision of whether or not to allocate a ventilator to a Covid-19 patient who meets the ventilator eligibility criteria will determine whether or not the patient has a better chance at surviving Covid-19. If such a decision system was, for instance, optimised to maximise the number of likely survivors, then one method may be to base the prediction of the decision tool predominantly on clinical criteria, which would then advantage those groups for whom the health system is already working. Under these conditions--where Māori are experience disproportionate health inequities (Waitangi Tribunal, 2019)--Māori would be further negatively impacted.

The above are plausible cases that illustrate the risks associated with poorly conceived ADM systems being implemented in historically biased systems. It is not merely technical design decisions that are involved in these cases, but value judgements and broader systemic considerations of the contexts in which ADM are employed. This reiterates the point that Māori need to be meaningfully engaged and involved in the design of systems that employ ADM.

Existing assurance frameworks/standards/guidance in Aotearoa

In this section we examine commentary on existing frameworks and guidelines that are currently in use by government organisations related to data use. Examples include *Ngā Tikanga Paihere,* the *PHRaE framework*, and the *Algorithm Charter*.

Nga Tikanga Paihere

Ngā Tikanga Paihere is a trust model developed to guide data use. It draws on 10 traditional Māori concepts identified by Hudson et al (2017) was originally designed to build and maintain public trust and confidence in the way Stats NZ manages access to microdata in the Integrated Data Infrastructure (IDI). Ngā Tikanga Paihere aligns to the 5 Safes Framework and supports responsiveness to Māori.

Figure 1: Ngā Tikanga Paihere

The PHRaE Framework

The Ministry of Social Development created the *Privacy, Human Rights and Ethics* framework (PHRaE) to identify and assess risks when using personal data (Ministry of Social Development, 2020). Core to the framework is a questionnaire that is filled in when proposing and when implementing systems that use personal information. This framework also applies to proposed operational algorithms (Stats NZ, 2018, p. 29).

The PHRaE includes a questionnaire section on potential impact on Māori and the MSD's obligations under Te Tiriti. However, it should be noted that this questionnaire is effectively self-regulation by the Ministry, which then requires the appropriate in-house expertise and engagement processes in order to be effective. There is little relevant existing commentary on the PHRaE. However, Gavaghan et al. (2019, p. 72), when commenting on the PHRaE in relation to assessing Māori engagement and impact note that "there is an opportunity for the Ministry to improve and strengthen its in-house work by engaging with Māori, including Māori data scientists and other experts, when using the PHRaE."

The Algorithm Charter

In late 2019, a draft version of the "Algorithm charter for Aotearoa New Zealand" was circulated for the purposes of garnering feedback about a proposed opt-in set of principles to provide guidance to Government agencies, thus "ensuring New Zealanders have confidence in how government agencies use algorithms." Again, this is an opt-in framework that uses self-assessment. The final version of the charter was released in July, 2020.

Some of the submissions on the draft that have been published publicly by Te Mana Raraunga and Karaitiana Taiuru expressed views about the inadequacy of the algorithm charter to protect Māori. Te Mana Raraunga emphasise the following point in their response (Te Mana Raraunga, 2019):

> Our main feedback [...] is that an 'Algorithm Charter' is insufficient to protect Māori rights and interests. Regulation that includes mechanisms for accountability and redress is necessary. Such regulation would need to include Māori data governance at all levels.

Further specific points include: that the wording of the charter was vague: it was unspecific whose notion of "fair" was being applied in relation to algorithms, nor what "transparent and accountable" use of operational algorithms means in the circumstances of self-assessment and self-regulation by ministries (with no independent oversight). Moreover, there was little evidence of processes to "address issues relating to Māori data sovereignty, or broader ethical issues about whether what has been done is ethical, equitable or acceptable to communities and individuals." The lack of accountability mechanisms or redress for racialised inequities, which are highly likely to occur in colonised societies, was noted. Also, there is inadequate recognition of the various obligations to Māori that ought to be met.

> Any framework on the use of government algorithms needs to actively recognise government obligations under the Treaty of Waitangi and the United Nations Declaration on the Rights of Indigenous Peoples. It is not clear what "Embed a Te Ao Māori perspective in algorithm development or procurement" means, how this would be monitored, and why Treaty of Waitangi and other obligations are not directly referenced.

The lack of recognition of Te Tiriti and the United Nations Declaration on the Rights of Indigenous People were also noted in Karaitiana Taiuru's response to the draft algorithm charter (Taiuru, 2019).

Despite the pointed criticisms, Stats NZ's summary of feedback reported that the draft algorithm charter was on the right track: "Half of all submissions supported the Te Ao Māori commitment but suggested it go further and provide more clarity about how agencies will embed a Te Ao Māori perspective." (Stats NZ, 2020a, p. 6.)

Taiuru noted that Māori Data Sovereignty did not appear in the draft charter (Taiuru, 2019). If we move on from considering comments on the *draft* algorithm charter to considering the *final* version, Māori data sovereignty is mentioned only to exclude it from the charter's scope, stating that the charter "cannot fully address important considerations, such as Māori Data Sovereignty, as these are complex and require separate consideration" (Stats NZ, 2020b). The explicit lack of a commitment to Māori Data Sovereignty principles is concerning and not justified by the provided explanation, particularly given that *algorithms* are trained on *data* to create *models*, the latter of which are the real topic of discussion when talking about "operational algorithms." Moreover, the standard job title for the professionals who implement these models is "data scientist." So data sovereignty rights are often inextricably entwined with algorithm development, particularly for government services.

Taiuru submission on the draft algorithm charter states that 'The term "te Ao Māori" is used without substance nor understanding its meaning. Nothing at all in the draft charter refers to "Te Ao Māori" except once in word only' (Taiuru, 2019). Indeed, a stronger and more convincing commitment could certainly be made to ensure that tikanga is observed with respect to ADM tools and that kaupapa Māori frameworks are employed in design, implementation, use, and maintenance in ways that honour Te Tiriti.[2]

The final version of the Algorithm charter listed the following principle under the heading of "partnership": "Embedding a Te Ao Māori perspective in the development and use of algorithms consistent with the principles of the Treaty of Waitangi" (Stats NZ, 2020b). While this principle explicitly acknowledges the Treaty of Waitangi (which was missing from the draft), it is not exactly clear what the claim means. The English wording of the principle is ambiguous with respect to assuming a single Māori worldview--Taiuru notes that "there is no one Māori world view" (Taiuru, 2020, p. 13)--or else "embedding" one of a multitude of Māori world views. The reo Māori version suggests the latter ("tētahi tirohanga Ao Māori" (Stats NZ, 2020c)). Further, the notion of *embedding* a Te Ao Māori perspective in the English version allows for weaker and stronger interpretations; while the use of *whakapūmau* in te reo Māori suggests the stronger interpretations of persisting or maintaining that view throughout development and operation. But in any case, this is not a replacement for Māori participation in governance decisions (Te Mana Raraunga, 2020). Additionally, Māori Data Sovereignty--which is strongly influenced by Te Tiriti rights--is deemed a separate concern and yet it is integral to model development. Further, while the section heading is "partnership", additional long-recognised Te Tiriti principles include the addition of (at least) *participation* and *protection* (Taiuru, 2019).

[2] Refer to Karaitiana Taiuru (2019) *Treaty of Waitangi / Te Tiriti and Māori Ethics Guidelines for: AI, Algorithms, Data and IOT* for further discussion of these points.

The final version of the algorithm charter includes a risk matrix for government organisations to assess the risk, which includes an estimate of the likelihood of an adverse event combined with an estimate of the severity of the event. According to the risk matrix, the algorithm charter must be applied in situations of high likelihood and high severity of impact, but the matrix leaves much room for judgement calls and it seems that there is no external accountability if the principles of the algorithm charter is not applied (Stats NZ, 2020b).[3]

Te Mana Raraunga has utilised Māori Data Sovereignty principles to propose recommendations for algorithm/models which are reproduced as Appendix A (Te Mana Raraunga, 2019).

References

1000minds (2020). Covid-19 Patient Prioritisation. Retrieved from: https://www.1000minds.com/sectors/health/covid-19-prioritization Date accessed: 4/11/2020

Accident Compensation Corporation [ACC], (September 28, 2017). How we use predictive analytics to improve our clients' experience. Retrieved from: https://www.acc.co.nz/newsroom/stories/how-we-use-predictive-analytics-to-improve-our-clients-experience/ Date accessed: 4/11/2020

Arief-Ang, I, Salim, F., Hamilton, M. (2017). CD-HOC: Indoor Human Occupancy Counting using Carbon Dioxide Sensor Data. Retrieved from: https://arxiv.org/abs/1706.05286 Date accessed 6/11/2020

Bakker, L., O'Malley, J., Riley, D. (1999). Risk of Reconviction - Statistical Models that Predict Four Types of Reoffending. Retrieved from: https://www.corrections.govt.nz/__data/assets/pdf_file/0020/10667/roc.pdf Date accessed: 6/11/2020

Blackmore, B. (2020). Predictive Risk Models in Criminal Justice: Te Matapae Tūraru ā-hiko me te pūnaha ture taihara. In Chen, A (ed). *Shouting zeros and ones: Digital technology, ethics and policy in New Zealand.* Bridget William Books: Wellington. Pp. 102-118

Blackstock, C. (2013). An Analysis of the Ethical Issues for Māori in Predictive Risk Modelling to Identify New-Born Children who are at high risk of Future Maltreatment: November 14, 2013. Retrieved from: https://www.msd.govt.nz/documents/about-msd-and-

[3] A controversial example of self-assessment is evidenced in the Department of Internal Affairs (2020) Privacy Impact assessment: Threshold check - Facial Recognition System, pp 10-12. It is clearly indicated in the initial risk assessment that the project involves highly sensitive information, and several high and medium level risks are identified. This fits the description of a "High Risk" project in the 'Summary of privacy impact', and yet the project is rated only as medium risk. As already mentioned, this anomaly is noted as a failure to uphold Te Tiriti and UNDRIP obligations in Te Mana Raraunga's statement regarding this project (Te Mana Raraunga, 2020).

our-work/publications-resources/research/predictive-modelling/ethical-issues-for-maori-blackstock-review.pdf Date accessed: 7/11/2020

Blank, A., Cram, F., Dare, T., De Haan, I., Smith, B., Vaithianathan, R. (2013; updated 2015). Ethical Issues for Māori in Predictive Risk Modelling to Identify New-Born Children who are at High Risk of Future Maltreatment. Retrieved from: https://www.msd.govt.nz/documents/about-msd-and-our-work/publications-resources/research/predictive-modelling/00-ethical-issues-for-maori-in-predictive-risk-modelling.pdf Date accessed: 7/11/2020

Bruhn, J. (2014). Identifying useful approaches to the governance of indigenous data. The International Indigenous Policy Journal, 5(2). http://doi:10.18584/iipj.2014.5.2.5

Campbell, D., Boyle, L., Soakell-Ho, M., Hider, P., Wilson, L., Koea, J., Merry, A. F., Frampton, C., Short, T. G. (2019). National risk prediction model for perioperative mortality in non-cardiac surgery. British Journal of Surgery 106: 1549–1557. Retrieved from: https://bjssjournals.onlinelibrary.wiley.com/doi/10.1002/bjs.11232 Date accessed: 8/11/2020

Casconea, Y., Ferraraa, M., Giovanninia, L., Seral, G. (2017). Ethical issues of monitoring sensor networks for energy efficiency in smart buildings: a case study. 9th International Conference on Sustainability in Energy and Buildings, SEB-17, 5-7 July 2017, Chania, Crete, Greece. Retrieved from: https://www-sciencedirect-com.ezproxy.auckland.ac.nz/science/article/pii/S1876610217346696 Date accessed 6/11/2020

Cook, L., & Gray, A. (2020). Official statistics in the search for solutions for living with COVID-19 and its consequences. Statistical Journal of the IAOS, 36(2020): 253-278. https://doi.org/10.3233/SJI-200671

Cormack, D., Kukutai, T., & Cormack, C. (2020) Not one byte more: From data colonialism to data sovereignty - Kia kaua tētahi paita anō: Mai i ngā whakatōpū raraunga ki te mana motuhake o ngā raraunga. In Chen, A (ed). *Shouting zeros and ones: Digital technology, ethics and policy in New Zealand.* Bridget William Books: Wellington. Pp. 71-83

Couture, S. (August 5, 2020). *The Diverse Meanings of Digital Sovereignty.* Retrieved from: http://globalmedia.mit.edu/2020/08/05/the-diverse-meanings-of-digital-sovereignty/ Date accessed: 5/10/2020

Data Futures Partnership (2017). *A path to social licence: Guidelines for trusted data use*

Data Futures Partnership (2020). *Trusted Data Use Guidelines: For Aotearoa New Zealand.*

Department of Internal Affairs (2020). Privacy Impact assessment: Threshold check - Facial Recognition System. Retrieved from: https://assets.documentcloud.org/documents/7224765/DIA-Privacy-Threshold-OIA-2021-0150-Oct-2020.pdf Date accessed: 9/11/2020

Edwards, P., Fleming, A., Lacey, J., Lester, L., Pinkard,E., Ruckstuhl,K., Bezuidenhout, C., Payn, T., Bayne., & Williams, T. (2019). Trust, engagement, information and social licence - insights from New Zealand. *Environmental Research Letters, 14*(2019). https://doi.org/10/1088/1748-9326/aaf33c

Gavaghan, C., Knott, A., Maclaurin, J., Zerilli, J., Liddicoat, J. (2019). Government Use of Artificial Intelligence in New Zealand. Retrieved from: https://www.cs.otago.ac.nz/research/ai/AI-Law/NZLF%20report.pdf Date accessed: 4/11/2020

Gulliver, P., Jonas, M., Fanslow, J., McIntosh, T., & Waayer, D. (2018). Surveys, social licence and the Integrated Data Infrastructure. *Aotearoa New Zealand Social Work, 30*(3): 57-71. https://dx.doi.org/10.11157/anzswj-vol30iss3id481

Health Quality and Safety Commission New Zealand (2019). He Matapihi Ki Te Kounga O Ngā Manaakitanga Ā-Hauora O Aotearoa 2019. A Window On The Quality Of Aotearoa New Zealand's Health Care 2019. Retrieved from: https://www.hqsc.govt.nz/our-programmes/health-quality-evaluation/publications-and-resources/publication/3721/ Date accessed: 8/11/2020

Health Informatics New Zealand [HINZ], (2019). Algorithm automates ACC's claims process. Retrieved from: https://www.hinz.org.nz/news/435335/Algorithm-automates-ACCs-claims-process.htm

Housing New Zealand (2019). Official Information Act Response to Alex Harris. Retrieved from https://fyi.org.nz/request/10361-privacy-policy-and-assessment-for-smart-homes#incoming-35207 Date accessed: 6/11/2020

Hudson, M., Milne, M., Reynolds, P., Russell, K. and Smith, B. (2010). Te Ara Tika. Guidelines for Māori research ethics: A framework for researchers and ethics committee members. Retrieved from: https://www.hrc.govt.nz/resources/te-ara-tika-guidelines-maori-research-ethics-0 Date accessed: 9/11/2020

Hudson, M., Anderson, T., Dewes, T. K., Temara, P., Whaanga, H., & Roa, T. (2017) "He matapihi ki te mana raraunga": Conceptualising big data through a Māori lens. In Whaanga, H., Keegan, T. T. A. G., & Apperly, M. (eds) *He Whare Hangarau Māori - Language, Culture & Technology.* Te Pua Wānanga ki te Ao / Faculty of Māori and Indigenous Studies, The University of Waikato: Hamilton, New Zealand. Pp. 64-73

Jenkins, K. (2018). Can I see your social licence please? *Policy Quarterly, 14*(4): 27-35

Jones, N. (2015). Anne Tolley scraps 'lab rat' study on children. NZ Herald online. Retrieved from: https://www.nzherald.co.nz/nz/anne-tolley-scraps-lab-rat-study-on-children/C7GIGYW2467HG327FKXFRJDPEM/ Date accessed: 7/11/2020

Kamira, R. (2003). Kaitiakitanga: Introducing useful indigenous concepts of governance. In Coiera, E., Chu, S., & Simpson, C. (eds). HIC 2003 RACGP12CC [combined conference]:

Proceedings. Brunswick East, Vic.: Health Informatics Society of Australia ; Royal Australian College of General Practitioners. p.499-507

Keall, C. (2019). Lobby group objects to sensors in state houses, Kāinga Ora responds. Retrieved from https://www.nzherald.co.nz/business/lobby-group-objects-to-sensors-in-state-houses-kainga-ora-responds/MFXMPAQD66TVXFJDJJQPYCKW3Y/ Date accessed: 6/11/2020

Kukutai, T. (2019) Reflections on Indigenous sovereignty. *Journal of Indigenous Wellbeing: Te Mauri - Pimatisiwin, 4*(1): 3-5

Kukutai, T.,& Cormack, D. (2019)/Mana mōtuhake-ā-raraunga: Datafication and social science research in Aotearoa. *Kōtuitui:New Zealand Journal of Social Sciences,14*(2): 201-208. https://doi.org/10.1080/1177083X.2019.1648304

Kukutai, T., & Cormack, D. (2020). "Pushing the space" Data sovereignty and self-determination in Aotearoa NZ. In Walter, M., Kukutai, T., Carroll, S. R., & Rodriguez-Lonebear, D. (eds). *Indigenous Data Sovereignty and Policy.* Routledge: New York. pp. 21-36

Kukutai, T., Cormack, D.,& Cormack, C. (2020) *Whose data is it anyway?* Retrieved from https://thespinoff.co.nz/atea/21-08-2020/whose-data-is-it-anyway/ date accessed 13/11/2020

Kukutai, T., & Walter, M. (2017). Indigenous Statistics. In Liamputtong, P (Ed). *Handbook of Research Methods in Health Social Sciences.* https://doi.org/10/1007/978-981-10-2779-6_40-1

Ministry of Social Development (2015). Vulnerable Children Predictive Modelling. Retrieved from: https://www.msd.govt.nz/about-msd-and-our-work/publications-resources/research/predicitve-modelling/index.html Date accessed: 7/11/2020

Ministry of Social Development (2020). The Privacy, Human Rights and Ethics Framework. Retrieved from: https://www.msd.govt.nz/about-msd-and-our-work/work-programmes/initiatives/phrae/index.html Date accessed: 4/11/2020

Morton, J. (2020). NZ-developed AI can predict kids' hospital risk. NZ Herald Online. Retrieved from: https://www.nzherald.co.nz/nz/nz-developed-ai-can-predict-kids-hospital-risk/UNNQNQJNBTGRN3SBUYGCJS6ZB4/ Date accessed: 7/11/2020

Moses, C. (2020).The Integrated Data Infrastructure: Te Hanganga o ngā raraunga kōmitimiti ma te Whakaaetanga ā-Iwi. In Chen, A (ed). *Shouting zeros and ones: Digital technology, ethics and policy in New Zealand.* Bridget William Books: Wellington. Pp. 84-102

Ngomo,A.(December, 2019) Can biased people create unbiased algorithms? Retrieved from https://towardsdatascience.com/can-biased-people-create-unbiased-algorithms-a01813e0473a Date Accessed: 6/10/2020

nzRISK (2020). nzRISK website. Retrieved from: https://nzrisk.com/ Date accessed: 8/11/2020

Office of the Children's Commissioner (2020). Te Kuku O Te Manawa - Ka puta te riri, ka momori te ngakau, ka heke nga roimata mo toku pepi. Retrieved from: https://www.occ.org.nz/publications/reports/te-kuku-o-te-manawa/ Date accessed: 7/11/2020

Panian, Z. (2010) Some practical experiences in data governance. World Academy of Science, Engineering and Technology, 62.939-946. Retrieved from: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.190.6948&rep1&type=pdf

Pool, I. (2016). Colonialism's and postcolonialism's fellow traveller: the collection, use and misuse of data on indigenous people. In Kukutai, T., & Taylor, J. (eds) *Indigenous Data Sovereignty: Toward an Agenda.* ANU Press: Acton, ACT. pp. 57-78

Quince, K. (August 2020).Can Algorithms make decisions about our lives in an unbiased fashion? Retrieved from: https://www.stuff.co.nz/dominion-post/comment/columnists/122364803/can-algorithms-make-decisions-about-our-lives-in-an-unbiased-fashion Date Accessed: 6/10/2020

RNZ (2015). Child abuse risk study 'ethically flawed'. RNZ online. Retrieved from: https://www.rnz.co.nz/news/national/280069/child-abuse-risk-study-'ethically-flawed' Date accessed: 7/11/2020

Ross, M. (2020). The throat of Parata. In Elkington, B., Jackson, M., Kiddle, R., Mercier, O. R., Ross, M., Smeaton, J., & Thomas, A. (Contributors). *Imagining Decolonisation.* Bridget Williams Books: Wellington, NZ. pp 21-39

Smith, L. T. (2012). *Decolonizing methodologies: Research and indigenous peoples*. Zed Books Ltd..

Sporle, A., Hudson, M., & West, K. (2020). Indigenous data and policy in Aotearoa New Zealand. In Walter, M., Kukutai, T., Carroll, S. R., & Rodriguez-Lonebear, D. (eds). *Indigenous Data Sovereignty and Policy.* Routledge: New York. pp. 62-81

Stats NZ Tatauranga Aotearoa (2018). Algorithm Assessment Report. Retrieved from: https://data.govt.nz/assets/Uploads/Algorithm-Assessment-Report-Oct-2018.pdf Date accessed: 4/11/2020

Stats NZ Tatauranga Aotearoa (2020a). Submissions Summary: Draft Algorithm Charter. Retrieved from: https://data.govt.nz/assets/Uploads/Algorithms/Draft-Algorithm-Charter-Summary-of-Submissions-February-2020.pdf Date accessed: 5/11/2020

Stats NZ Tatauranga Aotearoa (2020b). Algorithm Charter for Aotearoa New Zealand. Retrieved from: https://data.govt.nz/assets/data-ethics/algorithm/Algorithm-Charter-2020_Final-English-1.pdf Date accessed: 5/11/2020

Stats NZ Tatauranga Aotearoa (2020c). He Tūtohi Hātepe (Algorithm Charter) Mō Aotearoa.Retrieved from: https://data.govt.nz/assets/data-ethics/algorithm/Algorithm-Charter-2020-Final-Te-Reo-Maori.pdf Date accessed: 5/11/2020

Stats NZ Tatauranga Aotearoa (2020d). Mana Ōrite Relationship Agreement. Retrieved from: https://www.stats.govt.nz/about-us/what-we-do/mana-orite-relationship-agreement/ Date accessed: 9/11/2020

Stats NZ Tatauranga Aotearoa (2020e). Aide Memoire to the Minister of Statistics: Meeting with the Data Iwi Leaders Group

Taiuru, K. (2019). Submission on Government algorithm transparency and accountability. Retrieved from: https://www.taiuru.maori.nz/submission-on-government-algorithm-transparency-and-accountability/ Date Accessed: 4/11/2020

Taiuru, K. (2020). Treaty of Waitangi/Te Tiriti and Māori Ethics Guidelines for: AI, Algorithms, Data and IOT. Retrieved from https://www.taiuru.maori.nz/tiritiethicalguide/ Date Accessed: 5/11/2020

Te Mana Raraunga (2016). Te Mana Raraunga - Māori Data Sovereignty Network Charter. Retrieved from: https://www.temanararaunga.maori.nz/tutohinga/ Date accessed: 9/11/2020

Te Mana Raraunga (2017). *Te Mana Raraunga Statement on Social Licence.* Retrieved from https://static1.squarespace.com/static/58e9b10f9de4bb8d1fb5ebbc/t/5924791cbf629a1367a65717/1495562526902/TMR+-+Statement+on+Social+and+Cultural+License.pdf

Te Mana Raraunga. (2018) *Principles of Māori Data Sovereignty.* Retrieved from https://www.temanararaunga.Māori.nz/nga-rauemi

Te Mana Raraunga (2019). Submission on the Review of the Draft Algorithm Charter. Retrieved from: https://static1.squarespace.com/static/58e9b10f9de4bb8d1fb5ebbc/t/5e79c0fb3ccc1d093689c06d/1585037565865/TMR+Submission+on+the+Algorithm+charter+Feb+2020.pdf Date Accessed: 5/11/2020

Te Mana Raraunga (2020). Te Mana Raraunga Statement on Department of Internal Affairs facial recognition system procurement. Retrieved from: https://static1.squarespace.com/static/58e9b10f9de4bb8d1fb5ebbc/t/5f861f5ac526656129d8e7df/1602625374065/Te+Mana+Raraunga+-+Statement+on+DIA+facial+recognition.pdf Date accessed: 9/11/2020

United Nations (2007). United Nations Declaration on the Rights of Undigenous Peoples. https://www.un.org/development/desa/indigenouspeoples/declaration-on-the-rights-of-indigenous-peoples.html Date accessed: 9/11/2020

Waitangi Tribunal (*undated*). A Guide to the Principles of the Treaty of Waitangi as expressed by the Courts and the Waitangi Tribunal. Retrieved from:

https://waitangitribunal.govt.nz/treaty-of-waitangi/principles-of-the-treaty/ Date accessed: 9/11/2020

Waitangi Tribunal, (2005). The Offender Assessment Policies Report - WAI 1024. Legislation Direct, Wellington, New Zealand. Retrieved from: https://forms.justice.govt.nz/search/Documents/WT/wt_DOC_68001752/Offender%20Assessment%20Policies.pdf Date Accessed: 4/11/2020

Waitangi Tribunal (2019). Hauora: Report on Stage One of the Health Services and Outcomes Kaupapa Inquiry (WAI2575). Retrieved from: https://www.waitangitribunal.govt.nz/inquiries/kaupapa-inquiries/health-services-and-outcomes-inquiry/ Date accessed: 8/11/2020

Walter, M. (2016). Data politics and Indigenous representation in Australian statistics. In Kukutai, T., & Taylor, J. (eds) *Indigenous Data Sovereignty: Toward an Agenda.* ANU Press: Acton, ACT. pp. 79-98

Walter, M. (2018). The voice of Indigenous data: Beyond the markers of disadvantage. *Griffith Review, 60* (2018): 256-263. https://search.informit.com.au/documentSummary;dn=586241932732209;res=IELAPA

Walter, M., & Carroll, S. R. (2020). Indigenous data sovereignty, governance and the links to Indigenous policy. In Walter, M., Kukutai, T., Carroll, S. R., & Rodriguez-Lonebear, D. (eds). *Indigenous Data Sovereignty and Policy.* Routledge: New York. pp. 1 - 20

Walter, M., & Suina, M. (2019). Indigenous data, indigenous methodologies and indigenous data sovereignty. *International Journal of Social Research Methodology, 22*(3): 233-243. https://doi.org/10.1080/13645579.2018.1531228

Wende K. (2007). A Model for data governance - Organizing accountabilities for data quality management (Paper 80). In ACIS 2007 Proceedings. Retrieved from: https://aisel.aisnet.org/acis2007/80/

West, K., Hudson, M. & Kukutai, T. (2020). Data ethics and data governance from a Māori worldview. In George, L., Tauri, J., & MacDonald, L. (eds). *Indigenous Research Ethics: Claiming Research Sovereignty Beyond Deficit and the Colonial Legacy Vol: 6.* Emerald Publishing Limited: Bingley, UK. pp 67-82

Whaanga, H. (2020). AI: A New (R)evolution or the New Colonizer for Indigenous Peoples. In: Indigenous Protocol and Artificial Intelligence Working Group. Indigenous Protocol and Artificial Intelligence Position Paper. Pp. 34-38. Retrieved from: https://spectrum.library.concordia.ca/986506/

Whare Hauora (2019). Open Letter: Kāinga Ora must stop their dodgy home sensor project. Retrieved from: https://our.actionstation.org.nz/petitions/stop-kainga-ora-s-dodgy-home-sensor-project Date accessed: 6/11/2020

Appendix A: Comments relating to Principles of Maori Data Sovereignty (Te Mana Raraunga: Submission on the Review of the *Draft Algorithm Charter*)

Rangatiratanga | Authority

Māori should be involved in governance and decision-making around government use of data, including which data are used in algorithmic decision-making or advanced data analytics. This should include recognition of data governance requirements for any data that is identifiable to a particular iwi or Māori collective.

Whakapapa | Relationships

Any new regulatory frameworks or accountability mechanisms in this space need to proactively consider and address the potential future impacts of government use of algorithms and advanced data analytics for Māori, and take steps to pre-empt potential future harm.

Whanaungatanga | Obligations

Both individual and collective rights in relation to data need to be recognised in government use of algorithms. A focus on individual privacy rights is not sufficient, and does not align with Indigenous data sovereignty or Indigenous rights under UNDRIP.

The government should be required to report regularly on the impacts on Māori of the implementation and operation of algorithms.

Kotahitanga | Collective benefits

Government agencies should be required to ensure that their use of algorithms provides real meaningful benefits for Māori, both at individual and collective levels.

Manaakitanga | Reciprocity

Free, prior and informed consent should be the underpinning principle and preferred approach to all government collection and use of data, in line with Indigenous rights.

Strong Māori governance should be embedded to ensure that government use of algorithms upholds the dignity of Māori and that algorithms and data analytics are not stigmatising or harmful to Māori collectives and/or individuals.

Kaitiakitanga | Guardianship

Māori should have control over deciding the protocols and policies around Māori data, including how Māori data are used in algorithms and advanced analytics.

Part B: Report on Expert Wānanga: Māori and Automated Decision Making

On the 21st of October 2020 an expert wānanga was held at Waipapa Marae to consider Māori perspectives on trust and automated decision making (ADM). The wānanga brought together Māori researchers, practitioners and community members, all with expertise in the tech sector. Representatives from the Digital Council, as well as Toi Aria and Brainbox, were present to engage in the dialogue also. The wānanga was organised in two parts; the first half was an opportunity for the representatives from the Digital Council, Toi Āria Massey University, and Brainbox to feed-back the findings of their research engagement, and for participants to ask questions and engage in dialogue. The second half was a discussion section where the experts were asked to engage with a series of questions.

The participants in the expert wānanga included; Miriame Barbarich, Kylie Reiri, Ben Ritchie, Paora Mato, Kepa Morgan, Elena Higginson, Atawhai Tibble, BenTairea, Rere Pope, Roman Mitch, Caleb Moses, Dan Walker, Sophie Tauwehe Tamati, Te Kahuratai Painting, and Petera Hudson. The following report presents a summary of the dialogue and engagement of the experts. This report begins with an overview of key points presented by representatives from the Department of Internal Affairs, Brainbox and Toi Āria, including associated questions raised by wānanga participants. These points provide a general context for the later dialogue that we presented in the “kaupapa kōrero” phase.

Part 1: Related Research and Context

Elena Higgins (DIA): Introduction to Digital Council Project

The Digital Council of Aotearoa New Zealand was established to advise the Government on how best to harness the capabilities of digital and data-driven technologies in a way that is both inclusive and representative of the communities they seek to serve. A key driver for the development and implementation of automated decision making technologies such as algorithms is to reduce the gaps that exist between Māori and non-Māori in social justice sectors. In order for ADM’s to be effective in their goal to ‘close the gaps’, more needs to be done to increase Māori trust in the system, as well as in ensuring the systems themselves are trustworthy.

Tom Barraclough and Curtis Barnes (Brainbox): Results of Literature Review: Trust and automated decision making

Trust was defined as *being vulnerable in the face of risk and interdependence while still expecting positive outcomes.* This is distinct from **trustworthiness** which is concerned with whether a system works as it should, or if it can be made to work as it should. Both trust and trustworthiness, as they are defined here are heavily dependent on the notion of **trust propensity**, which considers the factors which may shape a community's willingness to trust a system (for example historic exploitation).

Core recommendations (of literature review)

1. Some kinds of decisions are not amenable to automation
2. Don’t get hung up on kinds of algorithms
3. Profiling, risk prediction and forecasting are hazardous, but hazards can be managed

4. Focusing on safety and efficacy can lead to trust
5. It is difficult to extract 'decision-making' from other layers of an automated decision-making system
6. New Zealand needs to foster confidence in ADM among policy and technical communities

Expert dialogue and engagement on the above points:

1. "If Chief Executives are convinced of the reasonable reliability of a system, is ADM ok?"
2. At present, it seems that the legislative infrastructure surrounding data use and data driven technologies is currently reminiscent of the 'wild west' in that there is no legislation which deals specifically with the regulation of data. The controls that are in place in the digital space are limited and offer little by way of protections for Māori. These include:
    a. The algorithm charter - which operates as an opt-in, self-assessment model.
    b. The Privacy Act - which is limited in terms of its focus on privacy at the individual level (at the exclusion of collective interests).
    c. Chief Digital Officer - whose scope is relatively limited and somewhat ambiguous in the context of Māori digital and data interests

3. Technology and law move at drastically different rates. To create a legal regime that is fit for purpose is challenging because often by the time legislation is actually passed, it would be out of date.
4. Suggestion from one participant that there is potential for Māori to be actively engaged in policy setting by action, through the establishment of a Professional Organisation, similar to Engineering New Zealand, for example, which would allow/necessitate the development of a Code of Ethics.
    a. *"Code might be good, but the difficulty comes in where disparate pieces of code need to be connected. It's the connections where things go wrong"*
5. "*It's hard to trust ADM's when manual decision-making was never trustworthy"*
6. It is impossible to identify which systems are 'dangerous' without actually looking at them
7. Algorithms aren't the problem, the problems lie in the systems in which algorithms are operationalised. This comes back to the question of bias and the data driving the technologies. For some participants there was a strong assertion that we needed to shift away from this idea that bias is inherently bad - instead arguing that humans are laden with bias and that there is potential for bias to be used in a way to empower a Māori worldview.
    a. Associated question from the Māori literature review - should algorithms reflect the bias world that we live in, or should they be used to reflect the world we wish to create?
    b. Perhaps if we emphasise our biases so that we may see ourselves as Māori within it, trust may come as a result.
    c. On the point of bias it was noted that machines learn and our biases become theirs
8. There is an ethical question here for many Māori in the industry, where the role of enabling increased efficiency in decision-making can sometimes conflict with the notion that the objectives behind this efficiency in practice have the effect of further disadvantaging Māori and minority communities.
9. Question raised - What is the Government doing to demonstrate trustworthiness?
    a. There is a real problem when Māori are more likely to give data over to Facebook than to the Government - but is this because Facebook is considered more trustworthy?
10. Values- How do we create a values industry? That's where the gap is. Māori have self-defined values. 99% of algorithms are built overseas. How do we ensure those are reflective or accommodative of our values? Where is the place for Māori? How do we translate complex terminology that have different significance or nuance for different cultural understanding?
11. System design- The objective is to design systems with the aim of decolonising, re-humanising and diversifying this space. System design comes first and foremost, well before tech/data/inputs/algorithms.
12. Final point is that there is no Māori word for trust - what does this suggest about our worldview and concepts?

Anna Brown (Toi Āria: Design for Public Good, Massey University): Feedback on emerging themes

- **Public Engagement –** Thirteen engagement hui were carried out with participants from around Aotearoa including (in order engaged):
- Pacific youth leaders; Young people with care experience; Ethnic community leaders; Ethnic community youth; Women with migrant and refugee backgrounds; Whānau Ora Navigators; Members of the disability community (blind and vision impaired); Members of the Taxi community; Māori and Pacific youth; Members of the General Public.
- Using a scenario-based workshop process called 'Decisions and Algorithms — Where do you stand?' people we asked their comfort (trust and benefit) with the written scenarios (each developed with intentional ambiguity) in order to gauge insights and feedback.

Findings from the hui around what would make people more comfortable with ADM included:

- Start with people — not algorithms
- Start with what people need – not what the system needs
- Consider positive, people-based solutions first; Consider whether and how automated decision-making might relevantly support people-based solutions

Reasons for low trust in ADM systems:

- Lack of public information and understanding
- Data is a poor indicator of real human context
- Frequent emphasis on deficit-based data to inform decisions
- Lack of visibility, control and security of personal data
- Reinforcement of existing systemic inequities
- Mistrust of government agencies
- Mistrust of political influence on algorithmic policy

And if ADM approaches can contribute meaningfully the following should apply:

- Process of designing and overseeing algorithm must be led by people and communities most affected
- Define clear goals for equitable outcomes
- Balance deficit data with strengths-based inputs
- Temper ADM to overriding human decision-making; and
- Ensure transparency to people most affected.

Expert dialogue and engagement:

1. In the discussion of ADM, it is critical that there is a strong community voice. In saying this, we need to ensure that when we are engaging with communities that we value their time and energy by having done some preparation and study to engage in fruitful dialogue. It is not the role of the community to continually teach researchers who have not done the requisite background research.
2. We should not rely on algorithms or data to be a panacea for social problems. Transformative change happens at the point of human decision-making.
3. If the definition of trust provided by Brainbox is "a willingness to be vulnerable in the face of risk [...]", then it is unreasonable to expect that Māori would ever "trust" Government use of ADM. Therefore, whether it is possible for Māori to trust ADM is the wrong question to be asking.

4. We have the opportunity to create a Māori tech collective. We get to set the standard and share working models to grow this kaupapa in a collaborative way.
5. ADM- ‘Te whakapapa o te whakatau.’ Remnants of the old biased system will remain until it is intentionally removed. Without this, ADM will simply reproduce and reinforce existing systemic bias and inequity.
6. ‘Tātai hono ki te whakapapa’, what are the decision points connecting datasets within algorithms and where ADM is located within the decision making process?

Part 2: Kaupapa kōrero

The outcomes of engagement and dialogue are organised into two distinct, yet interrelated sections. First is the question of “trust” and what this looks like from a Māori worldview. The second section considers how a Māori approach to trust and ADM may be practically applied. Key questions for consideration included:

- What factors enhance Māori trust in the system?
- Where could the system shift power to enhance trust in Māori?
- What capacities need to be developed to support a more trustworthy system?

Trust

On the question of “trust” in relation to automated decision-making, it was noted that the English term does not have a fully commensurate kupu in te reo Māori. Instead, there are numerous related kupu that might be the focus of a similarly themed wānanga in relation to ADM. This raised the importance of being able to frame questions in ways that align with Māori concepts and values, allowing for discussion and debate within a te Ao Māori view and from the point of view of Māori interests as opposed to having the discussion framed in a way that is not easy to reconcile with Māori ways of doing and thinking about the relevant issues. The point was made that a wānanga for Māori perspectives in ADM ought really to be in te reo Māori only as that would properly ground the whakaaro.

A significant barrier to trust is that these systems are developed in such a way that the opportunity for Māori responses are primarily only reactionary. There is often a pattern of scandal and response. For example, the critical statements regarding the 2018 census or Tom Hemopo lodging a Waitangi Tribunal claim regarding the ROC*ROI algorithm. Some possible explanations for this pattern is that there is not sufficient transparent engagement with Māori regarding the proposed design and use of these systems and because the systems that are being supported by ADM were never designed with Māori at the centre. The automation of manual processes is flawed if the process that it is replacing already results in inequitable outcomes for Māori.

It is also possible for people to continue to use a system in situations of minimal trust if the value obtained from the service outweighs that lack of trust. For example, it may be uncomfortable to share certain kinds of data with Facebook and yet data might still be shared for the value provided. This is just to say that continued engagement with a service is not necessarily evidence of trust.

Notwithstanding that, there is a comment that one characteristic of trust for Māori is accountability to our people. Moreover, this accountability is built into the pou that matter to us. Accountability should be built into systems from our own values.

Another participant noted that many ideas that were offered in this wānanga could (and have) been offered years ago. This point speaks to the lack of trust in or legitimacy of Māori thought by the

Government regarding what processes and systems may or may not work for Māori. This also speaks in part to the issue that while there may be some distinctive characteristics of the technology (machine learning and expert systems), the current contexts for decision-making and the relationships between the decision-makers and those upon which decisions are made has not fundamentally changed.

Comments specifically on trust included that there was no current incentive for Māori to trust the systems in which ADM may be employed. Specific points that were mentioned at this point included the importance of an **alignment of values** between those who create, implement and maintain ADM systems and Māori groups; and transparency about ADM processes, which can be aided by collaboration. Further, good intentions on the part of collaborators is not enough to achieve this; it also requires the development of capability in technical ability but also capability in understanding Māori culture, Te Tiriti, and culturally safe engagement with meaningful partnership.

Mana-Mahi

There are of course a number of ways that we might organise the dialogue, in this instance the mana-mahi framework seems appropriate for the following reasons:

1. It is a model which is currently used successfully as an organising framework for the distribution of roles and responsibilities between TMR and the Data ILG in the data sovereignty space
2. It is a model which recognises the value of theory (mana) and practice (mahi) and the inter-relationship between the two.
3. Drawing on an existing model allows us to be able to start the mahi immediately without needing to create a model from scratch. It also has the added benefit of being a tested model, and has proven to be effective in its use to date.

Importantly,in consideration of the use of the mana-mahi framework, there was thought given to whether we would need to ‘stretch’ the principles of the framework for use in the ADM space. The outcome of this thinking is that while the framework has been designed for use in the Māori-Dsov space, the principles are suitably broad enough for application in the ADM kōrero from a Māori perspective.

Mana

Mana can be translated as power and authority. In the context of the mana-mahi framework, it represents governance considerations. The values that occur with mana include: whanaungatanga, rangatiratanga, and kotahitanga.

Whanaungatanga / Obligations

The Government has a duty to be transparent about the use (and proposed use) of ADM. This includes transparency regarding the outsourcing of development of these systems.

There should be a high level of māramatanga (understanding) regarding ADM, its use, development and future prospects. This should occur as a result of appropriate engagement with communities. Effective partnership and participation means that our conversations and contributions ought to be moving at the same speed as the technologies are being developed.

There is often a lack of transparency regarding the vendors involved in developing code or supplying systems to the Government. Information about the roles and responsibilities of the vendors ought to be made readily available.

There ought to be an audit process or design check to rate ADM systems according to criteria that would engender trust by Māori (which might include assessment of the various considerations that have been raised in the wānanga in relation to appropriate governance and operational activities and processes).

Data is power. The rights of Māori with respect to data retention and archiving requires improved regulation. There is a lack of control and a lack of security over data that Māori have rights and interests in. Analogous with building regulations, the regulations implemented for ADM might be stratified according to the degree of risk of harm involved. (I.e., some regulations might apply to basic structures, while larger buildings have additional regulations that need to be satisfied.) This model could help to keep models fit for purpose with different criteria depending on the impact of the system.

However, it was also noted during the wānanga that regulation cannot be the primary tool for ensuring safe ADM as it can take an extended period of time to be implemented and so is too slow to keep pace with technological change. Consequently, regulation ought to be supplemented with other tools like guidelines, which are more nimble. This will be discussed in greater detail later.

Rangatiratanga / Authority

A strong Māori regulatory body is required to support ongoing governance of ADM systems.

There is a suggestion of a digital iwi forum that would have the role of oversight of the building of ADM systems. This forum could also be empowered by a first right of refusal to build proposed ADM systems. Aside from (or alongside) a centralised model, a decentralised accountability model was also suggested. A decentralised model would allow for effective rangatiratanga at the level of iwi and hapū interests. (A centralised model, on the other hand, would require some kind of mandate from iwi and hapū.) Related to this is the idea that when you share power, you don’t lose power.

Māori in governance and in decision-making positions with respect to ADM are more likely to result in outcomes that avoid stigmatisation, inequitable outcomes, and other harms and that are more likely to deliver on possible benefits to Māori. A current lack of Māori empowered in these kinds of positions in Government diminishes trust in these systems.

There need to be frameworks that allow for meaningful consent of use of data and for potential withdrawal of consent. Consent is situational. For example, consent may be granted for data to be used for one purpose but then the existence of this data might prompt someone else to attempt to use it for a secondary purpose. The changing systems and environments in which algorithms are implemented may mean that the kinds of decisions that this technology originally supports might change as the interests of those who use and control the system changes (for example, a system originally developed to identify people with the aim of providing additional support might be turned to identifying individuals for whom punitive measures might be imposed). The repurposing of systems toward ends that they were not originally developed for changes the situational context in important ways. Māori communities ought to retain the mana to determine appropriate applications of ADM systems that impact them.

Kotahitanga / Collective benefits

Building Māori capability in the domain of ADM development, or in AI and machine learning more broadly, can provide benefit for Māori and potentially improve the likelihood that ADM systems are developed and implemented in ways that provide meaningful benefit to Māori.

- One comment regarding an emphasis on capability building is that this takes a long time to achieve. But the issues and prospects related to ADM are immanent, which means that strategies cannot be deferred into the future. So, other more immediately actionable ideas are required in addition to capability building.

Māori objectives/take ought to be discussed. Some related questions include: “what is the kaupapa? And whose kaupapa is it?” These questions and statements are related to a point that the design of ADM involves many questions, implicit and explicit assumptions, and sometimes competing goals. The way that problems are framed and the assumptions of those who create the system all inform the design of the resulting technology; the consequent outcomes of its operation; and which outcomes will be selected to be monitored. Te Ao Māori framing of the problems that proposed ADM systems are meant to solve is required in order to increase the potential for ADM systems to be created that provide benefit to Māori (rather than result in further inequities, as is the case with existing ways of framing the problems to be solved).

It is typical to consider bias in relation to models as something that ought to be eliminated. This fits a kind of conception of algorithm development as something that is either black or white. But often the activities involved in this kind of bias elimination project is really competition between stakeholders interests, and all stakeholders have biases. In order to reduce the ambiguity regarding these trade-offs, these biases ought to be made more transparent by recognising them and making them explicit. The prioritising of these values should be part of the design framework. Given that there are competing biases and interests from different groups, and given that trade-offs may be required in practice, there needs to be a design process that is transparent and allows for negotiation and possible consensus about which interests are to be traded off and for what end. In this way bias can be used to empower worldviews. There is precedent for a design process like this in the Mauri Model (http://mauriometer.org/).

As mentioned earlier, there is a lot of critical commentary that is reactive in nature: a pattern that includes a scandal regarding ADM followed by a response. This emphasis on trying to fix systems that have caused (or continue to cause) issues is a result of improper engagement. These systems should give importance to equitable outcomes, but while the lack of this is often a focus of criticism of ADM systems, equitable outcomes are not sufficient to address trust.

One practical step to assist with the above mahi would be the formation of a Māori professional group. This group could, for example, develop a code of ethics for members; provide co-design, development, implementation, and maintenance advice; collate and distribute tools like acceptance criteria or model test cases for safer development; create practice notes; provide mentoring opportunities, etc. to guide practitioners in safe practices. Here, principles can guide best practices, and the practices can be considered and applied in relation to various use cases. Additionally, this would enable for a form of self-regulation, which is more responsive to emerging technologies than developing and implementing legislation.

In addition, Māori forums and wānanga should be created and held to develop mātauranga regarding ADM. This might include the creation of a Māori values framework or tikanga in relation to ADM design, development, use and maintenance. Further, there should be provisions for data that has been defined, collected and analysed in te reo Māori.

Mahi

Whakapapa / Relationships

One question raised is: ***what is the whakapapa of a decision*?** This is a particularly rich question as it can lead to discussion of questions such as: whose data was involved in the creation of a model? Who decided which features would be in the data sets? Whose interests were involved in the capture of the data? (e.g., is it useful for the kinds of questions that iwi and hapū groups would like to answer)? Whose interests were involved in the design of the model? Which algorithms are used to train the model from the data? And so on.

Given that machine learning may be used to create models based on training algorithms on historical data, and given that in many cases the historical data in these systems contain inequitable outcomes and other potential characteristics that—when an algorithm is trained upon it—may lead to models employed in ways that exacerbate those inequities. Many Government systems capture "deficit data" that reflects the interests of centralised systems, and which is not necessarily useful for Māori collectives to achieve their aspirations. Recognition of the potential for historical data to create inequitable outcomes in ADM should demand automatic auditing and testing of these models. Related to this, the data used to train ADM have often not been collected according to Māori interests. Effective ADM systems should not just use data from existing systems *about* Māori. ("ADM on Pākehā data about Māori will just be another form of colonisation.") Māori mana motuhake data should be used for ADM--defined, collected and analysed by Māori--in order to be effective for Māori.

There are relationships within the Māori world view that can give rise to characteristics of systems that are non-human-centred and incorporate spirituality. Recognition and inclusion of these relationships (and associated obligations) in systems may improve trust.

Manaakitanga / Reciprocity

In order for Māori iwi / hapū / whānau organisations and individuals to be able to effectively make decisions about what might be appropriate or inappropriate uses of data in ADM, it is important that groups and individuals know what data about us exists. But it is not adequate simply to have open data, the goal of active transparency and collaboration ought to be an integral component. Effective use of data can be powerful, but simply leaving data openly available is insufficient for empowered utilisation of that data.

The framing of the question for this wānanga is: what are Māori perspectives on trust in relation to ADM? But one issue is that there is little in the way of reciprocal trust relationships. There are numerous cases where feedback from Māori about the negative impacts of proposed programmes—within which these systems will be implemented—have been ignored or downplayed. This history of not trusting Māori views on what will or will not work for Māori undermines a relationship that upholds the mana of all parties. There needs to be engagement with Māori groups and individuals that upholds the mana of all parties.

There must be meaningful participation and collaborative partnership by Māori in algorithm governance and development. Moreover, Government organisations should "share the space" with Māori organisations. One comment was: "when you share the power, you don't lose the power."

Kaitiakitanga / Guardianship

Mainstream discussions of algorithms represent them as somewhat abstract entities, but this representation does not hold from a te Ao Māori point of view. The data employed in many ADM systems is from people, and so has a mauri. The algorithms that run on that data have been crafted with particular biases in mind to learn in particular ways to answer certain kinds of questions (consider, for example, the development of tailored natural language processing algorithms for bilingual purposes). The selection of specific datasets and features upon which the algorithms are trained is an exercise in curation and engineering of the data. The machines upon which these models are trained and executed have a materiality (even in the cloud), grounded in the sources of electricity used to power them and the hardware upon which they run. These models then ought to be viewed as both living—requiring kaitiakitanga—and relational—implicated in a network of obligations and relationships that need to be appropriately maintained.

'You don't leave a carving alone in the rain'

Reflecting on the need for ongoing care and maintenance of algorithms, one participant made an analogous comparison between the kaitiakitanga of algorithms and pou whakairo. The sentiment brought forward in their reflections was that there is an expectation that when a pou whakairo has been completed, it is then gifted - the gift though is not simply an exchange of the pou itself from one to another, but the recipient is also gifted with the responsibility of care. The idea being that the whakairo will not be 'left alone in the rain' where it might deteriorate, the kaitiaki is responsible for implementing adequate measures to ensure that the pou is protected so that it may be strong enough to stand exposed to the seasonal changes. In this sense, it was suggested that algorithms require a 'kaitiaki' who would be responsible for ensuring that the system itself can withstand broader structural changes and continue to be fit for purpose.

Central to this analogy is a recognition that Mahi Toi, in te ao Māori, have associated tikanga to guide process, and that if we understand coding as an artform, then we can see that tikanga should also be part of the process in ADM development. There is also a foundational belief here that mahi toi as both a process and an outcome is relational, and that the outcome, or in the case of the analogy, the pou, cannot be divorced or alienated from its context.

Recommendations

Insights emerging from the expert wānanga informed the recommendations outlined below which are intended to give effect to Te Tiriti o Waitangi and promote greater transparency about the use of ADM as it affects Māori communities.

1. Build Māori Data and Digital Capacity within both Māori communities and across networks of Māori practitioners.
2. Develop robust equity assessment protocols for algorithms.
3. Ensure meaningful Māori participation in Institutional algorithm self-assessment processes.
4. Support collaborative partnership in project governance and the development and use of algorithms.
5. Create a Māori values framework and tikanga guidelines to support ADM design, development, use and maintenance.
6. Explore Te Ao Māori use-cases involving te reo Māori, tikanga Māori, and mātauranga Māori in ADM

Action was seen by the Māori experts as time critical because ADM systems are impacting Māori communities now.