Indigenous Data Sovereignty & Open Data; Tensions and Resolutions

Tk Cram

The Author: *Tangitū ki te moana; Maungaharuru ki uta; Mōhaka te awa; Ko Ngāti Pāhauwera te iwi; Ko Tk Cram ahau.*

The Institute: *Pratt Institute is situated on Lenapehoking, the traditional and unceded homeland of the Lenape people past, present, and future.*

Introduction

In the age of big data an array of questions are being asked as to how our data is created, controlled, and distributed. Two movements that have manifested to answer these questions are Indigenous Data Sovereignty (IDS) and the Open Data Movement (ODM). IDS examines these questions through an indigenous perspective, focusing on the reclamation of data about, with, or by indigenous peoples. ODM, on the other hand, believes that data should be freely available without restrictions. While both of these movements come with lofty ambitions, the ways they approach data may not inherently be harmonious. The goal of this paper is to examine what tensions exist between these two movements, as well as what steps can be taken to have them coexist and potentially even support one another.

This paper will be divided into three parts. First, we will give a background on both of the movements. This includes the basic principles of each, their constituents, and their general relationship to one another. Second, we will review the literature around what tensions exist between the two movements. This will be done by examining some potential issues around the concept of openness, followed by factors about the data itself. Lastly, we will analyse what work has been done, and what still needs to be done, in order to resolve the tensions outlined in the previous section. This will cover current frameworks, best practices, and current examples, with a particular emphasis placed on what practical steps can be taken.

Background

Both IDS and ODM are emergent fields of study. While each has a strong theoretical framework, the precise implementations and best practices are still being developed. One such area is the way these two movements interact with one another. As more and more government entities adopt open data policies (Open Data Charter, n.d.), and IDS continues to develop amongst its stakeholders, this interaction is likely to become critical. In order to avoid potential conflict, steps should be taken to identify potential problem areas, as well as appropriate solutions. The goal of this paper is to consolidate what work has been done to date and highlight key practical steps that can be taken to not only mitigate conflict, but also to allow each movement to best achieve its stated goals. Simply stated:

What possible, practical resolutions can be found to the tensions between Indigenous Data Sovereignty and the Open Data Movement?

In order to best answer this question, a variety of literature will be collected on both IDS and ODM. This will include full publications, policy briefings, and grey literature. These will be reviewed and laid out in a manner consistent with what is outlined above. If gaps are identified within the literature they will be outlined at the end of the paper.

For literature on IDS, several networks exist to provide resources. At the time of writing these include the Global Indigenous Data Alliance, as well as regional networks in Aotearoa/New Zealand (Te Mana Raraunga), Australia (Maiam nary Wingara), and the USA (The United States Indigenous Data sovereignty Network). Of particular note amongst the literature is the edited collection "Indigenous Data Sovereignty: Towards an Agenda" (Kukutai & Taylor (Ed.) 2016), which will act as our primary source, as well as the United Nations Declaration on the Right of Indigenous Peoples (UNDRIP), on which much of IDS was founded (Kukutai & Taylor, 2016).

Unlike IDS, ODM is more decentralised in its literature. This creates variance amongst the philosophies and implementations of open data, depending on both region and sector. "The State of Open Data" (Davies et al. (Ed.) 2019) is an edited collection of works that will serve as our primary source on ODM, as it consolidates a variety of different perspectives in an attempt to unify the narrative. The Open Data Charter (ODC) also attempts to establish a consistent framework within the movement. (Open Data Charter, n.d)

Due to the scope of the paper, discussion and critique of either movement movement in isolation will be limited. In addition, discussion on government policies, and treaties will be limited to specific examples.

Indigenous Data Sovereignty

Indigenous Peoples

An exact definition of what constitutes an indigenous people would be contentious at best and prejudiced at worst. This is further complicated by government policies, such as 'blood quantum' (Kukutai & Taylor, 2016), that seek to draw absolute and quantifiable boundaries on membership to indigenous groups. Ultimately this led to UNDRIP adopting no formal definitions for either of these concepts, instead relying primarily on self identification. However, during the United Nations early engagement with indigenous peoples, Jose Martinez Cobo (in Davis, 2016) attempted to formalise a definition, reproduced in part below to allow some insight into the peoples discussed in this paper:

> "Indigenous communities, peoples and nations are those which, having a historical continuity with pre-invasion and pre-colonial societies that developed on their territories, consider themselves distinct from other sectors of the societies now prevailing on those territories, or parts of them. They form at present non-dominant sectors of society and are determined to preserve, develop and transmit to future generations their ancestral territories, and their ethnic identity, as the basis of their continued existence as peoples, in accordance with their own cultural patterns, social institutions and legal system."

With neither a precise definition nor complete data [see below] estimates on global indigenous populations can vary significantly. In 2009, the UN publication, "The state of the worlds indigenous peoples", estimated there were "more than 370 million [indigenous peoples] in some 90 countries". Unfortunately most of the research, including this

paper, is being done from the perspective of the CANZUS states (Canada, Australia, New Zealand, and the USA) with the Global South unrepresented in indigenous affairs. (Kukutai & Taylor, 2016).)

Data Sovereignty

From cloud computing to foreign surveillance, questions around ownership and control of data have become crucial in the modern era. Data Sovereignty is a movement that seeks to address these questions, ultimately asserting that data is subject to the laws of the nation in which it is collected. The first branch, now referred to as National Data Sovereignty (Irion, 2013), deals with the way national governments and multinational corporations deal with both personal and administrative data. But this paints an incomplete picture, as Kukutai & Taylor (2016) point out "Missing from those conversations have been the inherent and inalienable rights and interests of indigenous peoples relating to the collection, ownership and application of data about their people, lifeways and territories." Thus the second branch, Indigenous Data Sovereignty, was born.

Indigenous Data Sovereignty

At its heart, Indigenous Data Sovereignty is that idea that "indigenous peoples have inherent and inalienable rights relating to the collection, ownership and application of data about them, and about their lifeways and territories." (Te Mana Raraunga, n.d)

While indigenous peoples the world over have been fighting for their rights since they were first lost, IDS traces its formulation back to two key documents from the United Nations. The first, the inaugural sessions of the United Nations Permanent Forum on Indigenous Issues in 2002 and 2003, recognised that "a key challenge faced by national and international bodies is the lack of disaggregated data on indigenous peoples" (Tauli-Corpuz, 2016)

The second, the United Nations Declaration of the Rights of Indigenous Peoples (UNDRIP), was adopted by the UN general assembly. While the entire declaration is of importance to the movement, of specific note are articles 3-5 which covers indigenous rights to self-determination, 18-20 which covers the rights to participation and representation, and 31-32 which covers the rights to cultural and strategic development (UN, 2007).

In 2015, UNDRIP was leveraged in the 'Data sovereignty for indigenous peoples: current practice and future needs' workshop held in Canberra, Australia. The workshop was attended by representatives of the CANZUS states with the goal of of developing an IDS agenda. Drawing on this workshop, the edited collection "Indigenous Data Sovereignty: Towards an Agenda" was produced as the "first to engage with the topic of data sovereignty from an indigenous standpoint" (Kukutai & Taylor, 2016).

A few months after this meeting, attendees from Aotearoa held the inaugural Māori Data Sovereignty (MDS) meeting in Hopuhopu, Aotearoa from which emerged Te Mana Raraunga MDS Network . Per their charter, their purpose is to "enable Māori Data Sovereignty and to advance Māori aspirations for collective and individual wellbeing", particularly in regards to data. (Te Mana Raraunga, n.d) Since then similar networks have emerged including the "US Indigenous Data Sovereignty Network (USIDSN)" in the United States and "Maiam nayri Wingara" in Australia.

Each of these networks has developed their own internal principles designed for their local framework, as well as creating policy documents. In 2019 the networks came together, along with other stakeholders. This meeting had two key outcomes (1) the formation of the Global Indigenous Data Alliance (GIDA) and (2)

the endorsement of the CARE principles (Collective Benefit; Authority to Control; Responsibility; Ethics) to supplement open data's FAIR principles (GIDA, n.d.)

Open Data Movement

Open Data

An early modern conception of open data practices was presented by Robert King Merton in 1942. These Mertonian Norms were "four sets of institutional imperatives taken to comprise the ethos of modern science." The first of these imperatives, communism, formulates "The substantive findings of science are a product of social collaboration and are assigned to the community." (Merton, 1973)

It was in this spirit of collaboration that scientists undertook the Human Genome Project (HGP). The project, which began in 1990 (HGP, n.d), required scientists from across the world to sequence individual pieces of human DNA so that collectively they would be able to create a complete model. In order to facilitate this the Bermuda Principles were created in 1996. These included "that all human genomic sequence information, generated by centres funded for large-scale human sequencing, should be freely available and in the public domain in order to encourage research and development and to maximise its benefit to society." (HGP, 1997) It was on the back of these principles that a complete genome sequence was created in 2021. (NCBI, 2021)

Facilitated on the same collaborative principles that allowed the HGP, various implementations of open data came to be in the fields of art and technology. Creative Commons "gives every person and organization in the world a free, simple, and standardized way to grant copyright permissions for creative and academic works".(Creative Commons , n.d.) Open Source "is software that can be freely accessed, used, changed, and shared (in modified or unmodified form) by anyone". (Open Source, n.d)

In the wake of the explosion of data available in the information era, in 2016, 'Scientific Data' published the FAIR principles (Findability, Accessibility, Inter-operability, and Reuse of digital assets) by Wilkinson et al. to better facilitate the flow of data. "The principles emphasise machine-actionability because humans increasingly rely on computational support to deal with data as a result of the increase in volume, complexity, and creation speed of data."

Open Government Data

2007 was the year that Open Data made the transition into public data. In the lead up to the 2008 US presidential election, 30 open government advocates gathered in Sebastopol, California to define open public data in the hopes of having it adopted by the candidates, and thus the government. Their reasoning being that "open data promotes increased civil discourse, improved public welfare, and a more efficient use of public resources.". In all, eight principles were set out in order for government data to be considered open: Complete; Primary; Timely; Accessible; Machine Readable; Non-discriminatory; Non-proprietary; and Licence free. (opengovdata, 2007; data.gov 2013)

And they succeeded. In 2009, sitting president Barack Obama signing the 'memorandum on transparency and open government' which stated that government should be transparent, participatory, and collaborative. This paved the way for the data.gov website, which provides access to "government data [...] made available in open, machine-readable formats" (data.gov, n.d.)

The Open Government Partnership Global Summit, which took place in Mexico, refined the principles created in Sebastopol in order the create the

Open Data Charter (ODC).(Gurin, 2015) The ODC mandates that government data must be (1) Open by default (2) Timely and comprehensive (3) Accessible and usable (4) Comparable and interoperable (5) For Improved Governance and Citizen Engagement and (6) For Inclusive Development and Innovation. (Open Data Charter, n.d) At the time of writing, 25 countries and 64 local governments have signed the charter, including Aotearoa, Australia, and Canada (Open Data Charter, n.d.)

Open Administrative Data

Government data is roughly divided into two categories: Research data and administrative data. Research data is generated by any research that receives government funding. For example in the US, 44% of all research funding came from the government. (Mervis, 2017) While this data is important to the ODM, the majority of this data goes beyond the scope of this paper.

Instead the focus will be on administrative data. This is data collected by the government either directly or provided by non-governmental organisations (NGOs). Most often this is statistical information about the citizens that reside within the government's territory, including census data, tax information, or profiles created by governmental services.

Administrative data is vital in not just the operations of the governing body, but also for monitoring and evaluating the wellbeing of its constituents (Wilson & Cram, 2018). This is part of what makes it invaluable to IDS, as article 23 of UNDRIP states "Indigenous peoples have the right to determine and develop priorities and strategies for exercising their right to development", (UN, 2007). which would be otherwise impossible without this information.

Tensions

The literature on IDS and OD can be roughly split into two categories. The first category is regarding ownership, control, and openness of the data. The second category pertains to the data itself, what's collected, how it's collected, and some of the built in biases. For this review we will examine both categories by assuming issues in the other have been resolved. That is to say, we will first review the openness of the data assuming the data is good, and secondly we will review the data itself assuming issues of openness are resolved.

About Openness

"high-modernist ideology can be detected in the relatively uncritical embrace of Big Data technologies and privileging of Open Data policies required for these technologies now sweeping Western nations, including the CANZUS countries." (Scott, 1998, in Walter & Carrol 2021)

The first principle of the Open Data Charter is that data should be open by default. "Open by default […] says that there should be a presumption of publication for all. Governments need to justify data that's kept closed" (Open Data Charter, n.d.). However the exact conceptualisation of "openness" leaves room for significant interpretability. From an indigenous perspective, "opening Indigenous data by default bypasses entirely the rights of Indigenous peoples to decide what, if any, of their data should be shared […] In the absence of such basic decision-making ability, there is a heightened risk of data misinterpretation and misuse." (Rainie et al., 2019)

A key framework is the First Nation principles of OCAP, which stands for ownership, control, access, and possession. What follows is an exploration of the challenges that open data might face while adhering to the first three of those principles, with open possession

able to explored in future writings. (First Nations Information Governance Center, 2014).

Open Ownership

Indigenous and colonial conceptions of ownership have differed significantly since they first came into contact. While colonial powers had an individualistic sense of ownership, indigenous communities also had a strong sense of collective ownership. This often led to conflict, with one such example being in 1870, when then New Zealand Minister of Justice wrote on the detribalisation of Māori "It was hoped that by the individualisation of titles to land […] they would lose their communistic character." (In Pool, 2016)

This is further complicated by differing concepts of what could potentially be owned. Continuing with the Māori example, the treaty that was signed between them and the colonial settlers, Te Tiriti o Waitangi, guarantees Māori "*te tino rangatiratanga o […] o ratou taonga katoa*", which can be interpreted as 'chieftainship of all their treasures'. While 'treasures' might be understood as material wealth in colonial terms, Māori worldview expands this to encompass immaterial things such as knowledge.

Fast forward to the present, and with "data is the new oil" (Humby, in Charles 2013), we see an increased discussion on data ownership. One perspective from the Open Knowledge Foundation says "Open data and content can be freely used, modified, and shared by anyone for any purpose" (Open Definition, n.d.). While this approach seems synergistic with the "communistic character" of the Māori peoples, this runs directly counter to Principle 6.3 of the Principles of Māori Data Sovereignty which states "Māori shall decide which Māori data shall be controlled (tapu) or open (noa) access." (Te Mana Raraunga, 2018)

Open Control

In the summer of 1990, two professors from the University of Arizona came to the Havasupai people, who reside below the rim of the grand canyon. A diabetes epidemic had struck the tribe, and the researchers had come to collect blood samples to see if genetics could help explain it. Over the next four years, roughly 100 tribe members gave blood samples, but no link was found. It wasn't until 2003 that the tribe found out these results, and this was also when they discovered that their samples had been used in some two dozen studies, ranging on topics from human migration to inbreeding. The tribe was hurt by the revelation, and issued a banishment order on all UoA employees. (Harmon, 2010)

This is but one example of the ways indigenous peoples have lost control of their data. Too often is data collected about them, by researchers or by governments, and then those very people lose all say in what becomes of it. This data can be used to contradict tribal knowledge, portray indigenous peoples in a negative light, or simply exert suzerainty over them.

Open Access

Returning to the ODC, principle 3 states that data should be Accessible and Usable. In order for someone to access the data, they must not only know how to interpret it, but also know how to operate the machinery within which it exists. This is a problem area for indigenous peoples, who are traditionally left behind across all areas of education. As Davies (et al.) writes "there has also been growing recognition that, in conditions of unequal access to the skills and resources to work with data, or with wider existing patterns of exclusion and disadvantage, opening up data may not always lead to desirable outcomes."

A pertinent example of this is in Mexico, where, of the 7.4 million indigenous peoples, only 15% report access to a computer, and 10% have internet access. In such a situation, the indigenous people have almost no mechanisms by which their data could be accessed. And those that do may not have the data literacy to act on them. (Raine et al. 2019)

About Data

"With their limited scope, aggregate format, deficit focus and decontextualized framework, this joint data/policy narrative cannot, and does not, yield meaningful portraits of the embodied realities of Indigenous lives" (Walter & Carrol, 2021)

What if we assume that all of the above issues were resolved, and that it's possible to have true openness of the data while maintaining an indigenous framework. The next area to investigate should be the exact nature of the data that is being released into the world.

Open data is fundamentally a privileged concept. Indigenous peoples have a tumultuous relationship with data, often being subjects of research rather than participants in it. Without being fully involved in the research process, this can lead to the data being severely lacking. This runs the gamut from being decontextualised, to being implicitly or explicitly biased, to not existing in the first place. What follows is a breakdown of how each of these topics affect the overall quality of the data.

Data context

While data is often stripped of context, that is not to say it's devoid of it. This can be problematic even when those collecting the information and those giving it exist within the same cultural framework. This lack of understanding can result in the data not painting a complete picture and can lead to conclusions being drawn that don't match the true state of things. When data is taken about indigenous peoples and made to fit in a western framework, these problems are exasperated.

When examining any indicator, it's vital to ask whether it's "simplification, or reductionism" (Davis et al., 2012, in Morphy 2016). In Arnhem Land, Australia the local indigenous people, the Yolngu, were asked on a survey to identify their religious beliefs. The problem came when they were only allowed to select a single response. As one participant noted "My beliefs are traditional, but my religion is [Christian denomination]". This lead Morphy to state "Because of the lumping of 'traditional beliefs' into the same category as religions such as Christianity, the prevalence of the former is consistently underreported" (ibid.)

Likewise, there is often a dearth of indicators that provide an accurate depiction of indigenous lives. For instance the concept of household in western contexts often imagines a single nuclear family, the indigenous reality is more complicated, with larger and more dynamic family groupings. (Walter, 2016) Often indigenous peoples will identify more with the land on which they reside rather than a specific household, but this isn't reflected in the data. Walter writes: "there is an absence of indicators concerning the nature and extent of connection to place."

Data bias

The rhetoric of racial inferiority may seem antiquated to modern society, but this is not to say that racism is behind us. Instead, it has been usurped by the "New Racism" (Walter, 2016). This new racism is founded not on genetic predispositions, but instead rooted in culture. While the narrative may be different, ultimately the idea is the same: that racial minorities rightfully belong to a lower caste.

A key tool in the myth of cultural inferiority is supposedly 'unbiased' data. After all, data is simply

facts and figures, how could they be biased? A regularly repeated phrase online is "Despite making up only 13% of the population, blacks make up 52% of crimes." (Know Your Meme, 2019). While correlation in no way makes causation, this overrepresentation is deeply troubling. As Walter and Carrol (2021) put it "Data do not make themselves".

This overrepresentation in certain statistics is a problem for indigenous peoples. Australian researcher Morphy (2016) summarised the topics of these statistics as the "5 Ds" of indigenous data. That is, disparity, deprivation, disadvantage, dysfunction and difference. While individually innocuous, these topics make up the vast majority of statistics on indigenous peoples that "serves to cement a 'deficit data–problematic people' correlation"

Data absence

In order for indigenous peoples to work with the data about them, that data needs to exist. Unfortunately this is not always the case. This is typically either due to the data never being collected or to it being anonymised beyond the point of use.

Data can be anonymised in two ways. The first is by removing directly identifiable information, such as names or government ID numbers, and the second is by removing inferentially identifiable information. In short, if the data is about a sufficiently small population then it may be possible to identify an individual based on a series of other data points. Often it is tribal regions that might not meet the size requirements, meaning "American Indian tribes are policymaking bodies currently operating without accurate and reliable data that are or can be disaggregated at levels that facilitate sound tribal policy." (Rodriguez-Lonebear, 2016)

Other times, it may not even be possible to disaggregate the data due to the nature of its collection. On example is in Sweden where "the processing of data that reveals ethnicity or race is prohibited." (Rainie et al, 2019) This poses problems for research regarding the Sami, the local indigenous peoples, where there is little to no official data being produced.

Lastly there are cases where indigenous peoples are not only excluded statistically, but politically as well. "This is particularly a challenge in the Global South in regions such as Africa where Indigenous peoples are not counted or recognised." (Ibid.)

Resolutions

With some of the issues faced at the intersection of IDS and ODM outlined, what follows is an examination of some potential solutions. First we will examine this from an indigenous perspective, calling out what steps can be taken to help align the reality of open data with the utopia it presents. Next we will examine what the ODM might do in order to adapt its principles to create a more inclusive framework for indigenous peoples. Lastly we will examine how the two movements might come together to continue to work this out.

An IDS perspective

"control by indigenous peoples over developments affecting them and their lands, territories and resources will enable them to […] promote their development in accordance with their aspirations and needs" (UN, 2007)

Open data provides a great egalitarian promise: that the world can be better with greater access to data. Unfortunately, as we've seen, this promise is only as good as the data itself. In order to fully realise this promise, there is work to be done to improve the quality of the data. Many of the things that IDS is already fighting for are things that would seek to improve the data. Perhaps the biggest thing that can be done is to involve indigenous peoples at every level of

the data, from participant to researcher, and from steward to user.

Indigenous Participants

As has been shown, a significant amount of data is being collected about indigenous peoples without their involvement. In Australia, efforts to enforce truancy were made without consultation with the community as to why it was occurring in the first place. This lead to a major effort being pushed, and ultimately failing, because it was treating a symptom and not the root cause (Walter & Carrol, 2021). There are many such cases where a similar story plays out, and where it all could have been avoided if it had been done in consultation with the indigenous populations.

To help with this, the Australian Bureau of Statistics (ABS) has implemented the Indigenous Community Engagement Strategy (ICES). "For example, before the 2011 census, the Australian Statistician championed the need to expand the ICES, [which] resulted in good response rates to the census and an improved range of data for Aboriginal and Torres Strait Islander peoples" (Jelfs, 2016)

Indigenous Researchers

The need for diversity cannot be understated. Māori make up 17% of Aotearoa's population, but less than 5% of the academic workforce. American Indians were "completely absent from the top 50 departments in mathematics, mechanical engineering, economics, political science and sociology." (Gewin, 2021) This means that data is being collected by researchers foreign to the community they're working in, which can lead to fundamental differences and misunderstandings of the community, their knowledge, and their traditions. In the Havasupai story above, for example, the lead researcher maintains that what they did was ethically sound despite outcry from the tribe (Harmon, 2010).

Efforts are being made made across the CANZUS states to increase indigenous participation in academia. In Aotearoa a "Performance Based Research Fund will reward work by Maori researchers at 2.5 times the rate of non-Maori academics" (Gerritsen, 2021) in order to boost scholarship. Not to say it will be as simple as funding the issue, as [someone] writes "It's a real challenge, however, to walk in two worlds. There are Western ways of learning, but you also have to maintain and protect Indigenous cultural integrity." (Gewin, 2021)

Indigenous Stewardships

There is no magic bullet for the way government can better handle indigenous data. What there is, however, are a variety of guiding principles developed by IDS groups in the CANZUS states that can be adopted.

Some work is already being done on this. The NZ Integrated Data Infrastructure already has in place the Data Ethics and Privacy Assessment (DEPA), based on Ngā Tikanga Paihere framework that "draws on 10 tikanga (Te Ao Māori - Māori world concepts) to help you establish goals, boundaries, and principles that guide and inform your data practice." (Stats NZ, n.d.)

This extends beyond national governments as well. The ODC itself is being continuously revised to better meet the needs of a diverse population. The Open Government Partnership already states a need to be "intentional, strategic and ambitious in bringing the perspectives of women and gender groups into open government." (Open Government Partnership, n.d). An expansion of this to include indigenous peoples is far from unreasonable.

Indigenous Users

Greater digital accessibility is of equal concern for IDS and ODM, and serves to benefit both. As Rainie et al. (2019) put it "funder commitments are

needed in order to support increased scholarship, action, and education about the issues at the intersection of open data and IDS and to bring Indigenous peoples into the conversations around open data."

A higher degree of data literacy amongst indigenous populations would allow them to better articulate their needs, and better communicate with movements such as Open Data.

Further, principles 5 and 6 of the ODC state that open data is "For improved government and citizen engagement" and "For inclusive development and innovation". Neither of these things can come to fruition if the data is inaccessible to a proportion of the population.

An ODM perspective

"Policy-makers are starting to understand that their role is not just to release data, but also to play an active role in governing data infrastructure and use" (Rainie et al. 2019)

One of the key issues with the current iteration of the ODM is it operates in very simple ideas. That government is a single homogenous entity, that there is a one to one relationship between an individual and data, that data is either open or its not, or that knowledge only exists in specific forms. These may have been sufficient when the movement was young, but in order to progress further we must realise that there are a myriad of nuances and grey areas.

Grey government

In many cases ODM assumes that government data only has a relationship with a single entity, i.e. the nation state. IDS has challenged this assumption by emphasising the role of indigenous nations as political entities, backed up by UNDRIP. This means that it is not at the sole discretion of the nation state to decide what data is and isn't open, and allow indigenous groups a seat at the table (Rainie et al. 2019).

The Canadian Government has already acknowledged this need. First Nations Peoples have had direct input on Canada's 4th plan of open government. "The Government of Canada will engage directly with First Nations, Inuit and Métis rights holders and stakeholders to explore an approach to reconciliation and open government, in the spirit of building relationships of trust and mutual respect." (Government of Canada, 2018)

Grey individuals

Much like the assumption of a single government actor, there is also the assumption of a single data actor. That is, data rights and privacy are only applicable to an individual. IDS challenges this idea by showing that often the collective rights are also an important topic to consider as "control and use of the data have predictable risks and benefits, at an individual as well as a collective level." (Jansen, 2016)

This is especially important in discussion who has ownership and control of data. "questions arise around the collective rights of *iwi* to unit-record access […] only culturally sensitive data might be seen as sovereign for *iwi"* (Kukutai & Taylor, 2016), ODM needs to consider not just its own framework, but also the legislative framework for collective rights

Grey openness

Open Data is not an all or nothing movement. There are options between the government hoarding data liking a dragon and swinging the barn doors open for all to see.

One example of such a middle ground can be seen in the New Zealand Integrated Data Infrastructure (IDI). (Stats NZ, n.d.) Under this system, the government collects, cleans, and maintains data

from across the public sector that anyone can request access to. However, in order to be granted access an adequate reason is needed, and the research must be well defined. Although there are critics of this approach from open data purists who say it's simply redefining the problem, it strives to maintain the spirit of openness while minimising possible harm.

Grey data

Data can take many forms, but in western society there is a strong bias towards the written word, and especially things that can be codified. There needs to be an expansion of the way that data is handled, and what is considered valid. "It is essential to recognise that, before contact with imperial powers, indigenous peoples had their own vibrant, meaningful bodies of data" (Pool, 2016)

Recently the merits of indigenous knowledge have been gaining recognition. Traditional medicines, once written off as local superstitions, are being examined to reveal positive impacts on health and wellbeing. Unfortunately not everything fits in nicely with a western knowledge framework. This includes Māori *whakapapa* (genealogy) or the totem poles found in the American North West. (Ibid.)

Coming together

The above recommendations are only a starting point. Both IDS and ODM have evolved significantly over the last decade, and it remains to be seen what shape they will take in a decade more. This means there is still much more work to be done. And it cannot be done alone. Stakeholders from both IDS and ODM need to come together to find the best way to navigate these issues, and hopefully reach resolutions that not only suit the needs of both, but allow both to thrive.

<u>Conclusion</u>

Both the Indigenous Data Sovereignty and Open Data movements present lofty ideals. The former being concerned with the data rights of indigenous peoples the world over, while the latter believes that all data should be available and accessible for everyone. For all these two movements have alike, such as the want to better people's lives, there is a tension between them when it comes to the data itself.

These tensions come in flavours. The first is the way indigenous rights over their data is threatened by it all becoming open. These issues include how traditional knowledge is treated, who owns the data, how external forces might use the data, and the lack of indigenous operators to guide the data. IDS also raises concerns about the data itself, both in its absence, the explicit and implicit bias of the data that does exist, and it's decontextualisation.

Several steps are suggested to help these movements coexist. Indigenous populations needs to be involved at every step of the process, from data acquisition, to stewardship, to increasing the number of data professionals. ODM needs to expand their framework to better represent the nuances that are found in the real world, including multiple government actors, collective data rights, and a more nuanced view of what constitutes openness.

This writing only makes up a small piece of the work to be done on this topic. Both IDS and ODM are burgeoning fields that continue to iterate and adapt. What's important is for both groups to be able to sit down together and decide how they want to proceed.

Future Work

Indigenous Data Possession

Missing from the discussion on openness was the way it relates to the possession principle of OCAP. Possession is the physical manifestation of control. This leads back to the larger Data Sovereignty movement, and the discussion on data centres and cloud storage. In the literature review for this paper, no significant writing was discovered on the topic, and it should be explored further.

Expanding the Framework

Zooming Out

Much of the work to date on ODM has come from the west, and that narrows down to the CANZUS region when talking about IDS. If we're to construct an inclusive framework, as the ODC principles state, more work needs to be done in other regions.

Several countries in the east have open data policies, including Japan and Taiwan (Open Data Index, n.d.). Two Scandinavian universities have begun research into Sami data sovereignty. (Rainie et al, 2019) But when it comes to the Global South little to no work is being done on either front. This is especially pertinent in countries such as Brazil, where government corruption and indigenous exploitation are rampant.

Zooming In

With the multitude of indigenous peoples being that IDS looks to represent, it's important to remember that indigenous peoples are not a homogenous group. Work needs to be done on individual data sovereignty movements. This means that while networks such as Te Mana Rauraunga and Maiam nary Wingara will overlap significantly in their ideology, Māori and Aboriginal & Torres Straight Islander Data Sovereignty should not be identical.

Corporate Data

This paper has been primarily dealt with the role of government data, which historically has been the most significant data resource. But as we move forward the reality is we need to enter a discussion on the wealth of data that's being assembled by private entities, and the role they play as neocolonial enterprises. This is anathema to both movements.

In an ideal scenario, ODM and IDS would be able to directly influence the tech giants such as Meta or Google. More likely any change will need to come from government regulation. A strong framework is needed for the relationship between personal & collective data and private enterprise.

Bibliography

Arthur, C. (2013, August, 23). Tech giants may be huge, but nothing matches big data. *The Guardian*. https://www.theguardian.com/technology/2013/aug/23/tech-giants-data

Creative Commons. From creativecommons.org

Daly, A., Devitt, S. K., & Mann, M. (Ed.) (2019). *Good data*. Institute of Network Cultures.

Data.gov. From data.gov

Davies, T. (Ed.) (2019). The state of Open Data Histories and Horizons. African Minds.

Davies, T., Walker, S. B., Rubinstein, M., & Perini, F. (2018). Conclusion and Remarks. In Davies, T., Walker, S. B., Rubinstein, M., & Perini, F. (Ed.) *The State of Open Data*. African Minds

Davis, M. (2016). Data and the United Nations Declaration on the Rights of Indigenous Peoples . In T.

Kukutai & J. Taylor (Ed.), *Indigenous data sovereignty: toward an agenda* (pp. 25-38). ANU Press.

Gerritsen, J. (2021, July, 05). Hiring more Māori researchers 'needs its own dedicated funding'. *Radio New Zealand*. https://www.rnz.co.nz/news/national/446190/hiring-more-maori-researchers-needs-its-own-dedicated-funding

Gewin, V. (2021, January, 12). How to include Indigenous researchers and their knowledge. *Nature*. https://www.nature.com/articles/d41586-021-00022-1

GIDA. From https://www.gida-global.org/

Government of Canada (2018). Canada's 2018-2020 National Action Plan on Open Government. From www.opengovpartnership.org/wp-content/uploads/2019/01/Canada_Action-Plan_2018-2020_EN.pdf

Harmon, A. (2010, April, 21). Indian Tribe Wins Fight to Limit Research of Its DNA. *New York Times*. https://www.nytimes.com/2010/04/22/us/22dna.html

Human Genome Project. Policies on Release of Human Genomic Sequence Data Bermuda-Quality Sequence. From https://web.ornl.gov/sci/techresources/Human_Genome/research/bermuda.shtml

Irion, K. (2013) Government Cloud Computing and National Data Sovereignty. In Policy & Internet (pp. 40-71)

Jansen, R. (2016). Indigenous Data Sovereignty: a Māori health Perspective. In T. Kukutai & J. Taylor (Ed.), *Indigenous data sovereignty: toward an agenda* (pp. 193-212). ANU Press.

Jelfs, P. (2016). The Australian Bureau of Statistics' Aboriginal and Torres Strait Islander enumeration and engagement strategies: challenges and future options. In T. Kukutai & J. Taylor (Ed.), *Indigenous data sovereignty: toward an agenda* (pp. 275-290). ANU Press.

Jurin, J. (2015, October, 27). The Open Data Charter: A Roadmap for Using a Global Resource. *Huffington Post*. https://www.huffpost.com/entry/the-open-data-charter-a-r_b_8391470

Know Your Meme (2019). *Despite Being Only 13 Percent of the Population*. From https://knowyourmeme.com/memes/despite-being-only-13-percent-of-the-population

Kukutai, T. & Taylor, J.. (2016). Data sovereignty for indigenous peoples: current practice and future needs. In T. Kukutai & J. Taylor (Ed.), *Indigenous data sovereignty: toward an agenda* (pp. 1-24). ANU Press.

Merton, R. K. (1973) [1942], "The Normative Structure of Science", in Merton, Robert K. (ed.), *The Sociology of Science: Theoretical and Empirical Investigations*, Chicago: University of Chicago Press, pp. 267–278.

Mervis, J. (2017, March, 09). Data check: U.S. government share of basic research funding falls below 50%. *Science*. https://www.science.org/content/article/data-check-us-government-share-basic-research-funding-falls-below-50

Morphy, F. (2016). Indigenising demographic categories: a prolegomenon to indigenous data sovereignty . In T. Kukutai & J. Taylor (Ed.), *Indigenous data sovereignty: toward an agenda* (pp. 99-116). ANU Press.

NCBI (2021). CHM13 T2T v1.1 From www.ncbi.nlm.nih.gov/assembly/GCA_009914755.3

Obama, B. (2009). Transparency and Open Government. From obamawhitehouse.archives.gov/the-press-office/transparency-and-open-government

Open Data Charter. From opendatacharter.net

Open Definition. From opendefinition.org

Open Government Data. (2007). Request for Comments. From https://public.resource.org/8_principles.html

Open Government Partnership. From www.opengovpartnership.org

Open Source. *Open Source Initiative*. From https://opensource.org

Pool I. (2016). Colonialism's and postcolonialism's fellow traveller: the collection, use and misuse of data on indigenous people. In T. Kukutai & J. Taylor (Ed.), *Indigenous data sovereignty: toward an agenda* (pp. 57-77). ANU Press.

Rainie, S., Kukutai, T., Walter, M., Figueroa-Rodríguez, O. L, Walker, J., & Axelsson, P. (2018). Indigenous Data Sovereignty. In Davies, T., Walker, S. B., Rubinstein, M., & Perini, F. (Ed.) *The State of Open Data.* (Pp. 300-319). African Minds

Stats NZ. *Integrated Data Infrastructure.* From https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/

Tauli-Corpuz, V. (2016). Preface. In T. Kukutai & J. Taylor (Ed.), *Indigenous data sovereignty: toward an agenda*. ANU Press.

Taylor, J., & Kukutai, T. (Ed.) (2016). Indigenous data sovereignty toward an agenda. Australian National University Press.

Te Mana Raraunga (2018) Principles of Māori Data Sovereignty. From static1.squarespace.com/static/58e9b10f9de4bb8d1fb5ebbc/t/5bda208b4ae237cd89ee16e9/1541021836126/TMR+Māori+Data+Sovereignty+Principles+Oct+2018.pdf

Te Mana Raraunga. From www.temanararaunga.maori.nz/

UN General Assembly, *United Nations Declaration on the Rights of Indigenous Peoples : resolution / adopted by the General Assembly*, 2 October 2007, A/RES/61/295, available at: https://www.refworld.org/docid/471355a82.html

Walter, M. (2016). Data politics and Indigenous representation in Australian statistics. In T. Kukutai & J. Taylor (Ed.), *Indigenous data sovereignty: toward an agenda* (pp. 79-98). ANU Press.

Walter, M. (Ed.) (2021). Indigenous data sovereignty and policy. Routledge, Taylor & Francis Group.

Walter, M., & Carroll, S. (2021) Indigenous Data Sovereignty, governance and the link to Indigenous policy. In Walter, M, et al. (Ed). I*ndigenous Data Sovereignty and Policy. (pp. 1-20).* Routledge

Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3,** 160018 (2016). https://doi.org/10.1038/sdata.2016.18

Wilson, M. & Cram, F. (2018). Predictive Modelling. In Tolich, M., & Davidson, C. (Eds.). *Social science research in New Zealand : An introduction*. Auckland University Press