



The State of Open Data

Histories and Horizons

Edited by Tim Davies, Stephen B. Walker,
Mor Rubinstein, and Fernando Perini

The State of Open Data

Histories and Horizons

[This page is intentionally blank]



The State of Open Data

Histories and Horizons

African Minds
Cape Town

International Development Research Centre
Ottawa • Amman • Montevideo • Nairobi • New Delhi

Open Data for Development (OD4D) Network

HOW TO CITE THIS VOLUME

Davies, T., Walker, S., Rubinstein, M., & Perini, F. (Eds.). (2019). *The State of Open Data: Histories and Horizons*. Cape Town and Ottawa: African Minds and International Development Research Centre.

First published in 2019 by African Minds and the International Development Research Centre (IDRC).

African Minds
4 Eccleston Place
Somerset West, 7130
Cape Town, South Africa
www.africanminds.org.za

A co-publication with
International Development Research Centre
PO Box 8500, Ottawa, ON, K1G 3H9, Canada
www.idrc.ca / info@idrc.ca

© Contributors 2019. Licensed under the Creative Commons Attribution 4.0 International licence (<http://creativecommons.org/licenses/by/4.0/>).



The research presented in this publication was carried out with the aid of the Open Data for Development (OD4D) Network and a grant from the International Development Research Centre, Ottawa, Canada. The views expressed herein do not necessarily represent those of IDRC or its Board of Governors.

ISBNs:
Print edition 978-1-928331-95-7
eBook edition (IDRC): 978-1-55250-612-7
ePub edition: 978-1-928331-96-4

Orders:
African Minds
4 Eccleston Place, Somerset West, 7130, Cape Town, South Africa
info@africanminds.org.za
www.africanminds.org.za

For orders outside Africa:
African Books Collective
PO Box 721, Oxford OX1 9EN, UK
orders@africanbookscollective.com

021

Indigenous data sovereignty

Stephanie Carroll Rainie, Tahu Kukutai, Maggie Walter,
Oscar Luis Figueroa-Rodríguez, Jennifer Walker, and Per Axelsson

send copy

Key points

- Indigenous Data Sovereignty (IDS) has emerged as an important topic over the last three years, raising fundamental questions about assumptions of ownership, representation, and control in open data communities.
- IDS refers to the right of Indigenous peoples to control data from and about their communities and lands, articulating both individual and collective rights to data access and to privacy.
- Ideas from IDS provide a challenge to dominant discourses in open data, questioning current approaches to data ownership, licensing, and use in ways that resonate beyond Indigenous contexts, drawing attention to the power and post-colonial dynamics within many data agendas.
- Growing IDS networks are working to shape open data principles to better respect the rights of Indigenous peoples.

Introduction

"Open data in the context of Indigenous peoples is a double-edged sword."

Open data is a site of tension for Indigenous peoples. Open data provides opportunities for sustainable development according to Indigenous aspirations, yet also sits at the nexus of current and historic data challenges as a result of colonisation, bias, and a lack of knowledge of Indigenous rights. Indigenous data sovereignty (IDS) provides a framework for maximising the benefit of open data for Indigenous peoples and other users of Indigenous data and for affecting the stewardship of all data. Open data communities often assume many binaries, including a single government actor (nation-states), that data is open or not, and that open data is useful data

(devoid of biases and relevance issues). In the context of Indigenous peoples, there are clear challenges for the mainstream open data movement around these binaries, as well as paths forward to assure the protection of Indigenous rights and data for development.

IDS refers to the right of Indigenous peoples to govern the collection, ownership, and application of data about Indigenous communities, peoples, lands, and resources. Indigenous data is defined here as data in a wide variety of formats inclusive of digital data and data as knowledge and information. It encompasses data, information, and knowledge about Indigenous individuals, collectives, entities, lifeways, cultures, lands, and resources.^{2,3,4,5} Under IDS, the data governance rights of Indigenous nations apply regardless of where the data is held or by whom. This includes the right to the generation of the data that Indigenous peoples require to support nation rebuilding and governance. IDS concerns itself with binary digital data (e.g. scientific, administrative, corporate), as well as information and knowledge, meaning a somewhat broader scope than normally considered by the open data movement. However, all too often researchers, agency staff, and others digitise Indigenous knowledge and information and enter it into open data arenas without the express permission of Indigenous peoples.⁶ While these acts may be well-intentioned, the result is the co-opting of Indigenous knowledge and the removal of Indigenous peoples from data governance processes. Therefore, IDS also comprises the entitlement to determine how Indigenous data is governed and stewarded, referred to as Indigenous data governance (IDG). IDS covers both data for governance and IDG.^{7,8}

Over the past three years, during the first iteration of the Open Data Charter (ODC),⁹ IDS became a global movement. The initial establishment phase, beginning in 2015, was primarily focused on raising awareness of IDS within Indigenous nations and nation-state data entities. As of 2018, IDG principles and protocols are now being conceptualised and operationalised across nation-states and across the broad terrain of Indigenous data realities. These actions notwithstanding, in the open data community, there remains a general lack of knowledge or understanding of IDS.

To date, open data policy and discussions have largely been framed around the needs and interests of nation-states and of open data advocates and users with minimal Indigenous engagement. This is unacceptable from an IDS perspective because of the potential conflicts between open data goals and the aspirations of Indigenous nations and peoples. For example, the ODC states that “Open data is digital data that is made available with the technical and legal characteristics necessary for it to be freely used, reused, and redistributed by anyone, anytime, anywhere.”¹⁰ This is a lofty goal, but the objective is in direct tension with the rights of Indigenous peoples to govern their data, including the right to decide what is shared or withheld, likely resulting from the ODC being developed without the involvement of Indigenous peoples. Articulating this fundamental tension, and how it can be addressed through Indigenous data governance and stewardship mechanisms, underpins this chapter.

Indigenous peoples

The United Nations (UN) estimated in 2009 that there were approximately 370 million Indigenous peoples living in 90 countries, with up to 5 000 different Indigenous cultures around

Rainie, S.C., Rodríguez-Lonebear, D., & Martinez, A. (2017). *Policy brief: Data governance for native nation rebuilding*. (version 2). Tucson: Native Nations Institute, University of Arizona. http://nni.arizona.edu/application/files/8415/0007/5708/Policy_Brief_Data_Governance_for_Native_Nation_Rebuilding_Version_2.pdf

Rainie, S.C., Rodríguez-Lonebear, D., & Martinez, A. (2017). *Policy brief: Indigenous data sovereignty in the United States*. Tucson: Native Nations Institute, University of Arizona. http://nni.arizona.edu/application/files/1715/1579/8037/Policy_Brief_Indigenous_Data_Sovereignty_in_the_United_States.pdf

Rainie, S.C., Schultz, J.L., Briggs, E., Riggs, P., & Palmanteer-Holder, N.L. (2017). Data as a strategic resource: Self-determination and the data challenge for United States Native Nation and Tribes. *International Indigenous Policy Journal*, 8(2). <http://ir.lib.uwo.ca/iipj/vol8/iss2/1>. DOI: 10.18584/iipj.2017.8.2.1

Walter, M., Lovett, R., Bodkin Andrews, G., & Lee, V. (2018). *Indigenous data sovereignty. Briefing Paper 1. Miaim nayri Wingara Data Sovereignty Group and the Australian Indigenous Governance Institute*. <https://static1.squarespace.com/static/5b3043afb40b9d20411f3512/t/5b70e7742b6a28f3a0e14683/1534125946810/Indigenous+Data+Sovereignty+Summit+June+2018+Briefing+Paper.pdf>

About the authors

Stephanie Carroll Rainie is Ahtna (Alaska, US), and is Assistant Professor, Public Health Policy and Management, Mel and Enid Zuckerman College of Public Health; Assistant Research Professor, Udall Center for Studies in Public Policy; Associate Director, Native Nations Institute; and Co-Director, Center for Indigenous Environmental Health Research at the University of Arizona. She is the co-founder of the US Indigenous Data Sovereignty Network. You can follow Stephanie on Twitter at <https://www.twitter.com/scbegonias>.

Tahu Kukutai is from the Ngāti Tiipa, Ngāti Kinohaku, and Te Aupōuri tribes of Aotearoa/ New Zealand and is Professor of Demography at the University of Waikato. She is a founding member of the Te Mana Raraunga Māori Data Sovereignty Network. You can follow Tahu on Twitter at <http://www.twitter.com/thkukutai> and learn more about her work at <https://www.temanararaunga.maori.nz/>.

Maggie Walter is Palawa (Tasmanian Aboriginal) and Professor of Sociology at the University of Tasmania, Australia. She is a founding member of the Maiaim nayri Wingara Indigenous Data Sovereignty Collective. You can follow Maggie on Twitter at <http://www.twitter.com/IDSovOz>.

Oscar Luis Figueroa-Rodríguez is Associated Professor Researcher with the Rural Development Studies Program at the Colegio de Postgraduados Campus Montecillo, Mexico.

Jennifer Walker is a member of the Six Nations of the Grand River and holds a Canada Research Chair in Indigenous Health at Laurentian University. You can follow Jennifer on Twitter at <http://www.twitter.com/jencandlish>.

Per Axelsson is an Associate Professor at the Centre for Sámi Research and the Department of Historical, Philosophical and Religious Studies at Umeå University, Sweden.

The ODC sits in a powerful seat to advocate for IDS and changes to data stewardship, as well as to facilitate investments in building Indigenous data capacity and capability. The ODC, now in a revision phase, must recognise IDS in its next iteration. Such recognition also should include the importance of IDG and the stewardship of Indigenous data by others, in partnership with Indigenous peoples. IDS can be seen as an anathema to open data, but acknowledgment of IDS and engagement with Indigenous peoples supports ethical open data that allows for development aligned with and benefitting Indigenous aspirations. Adoption and implementation of the principles of IDG emerging from the RDA IDS Group would strengthen the ODC. Engagement with Indigenous peoples during the next round of ODC revisions must include the IDS networks, but also Indigenous leaders, scholars, and community members. Particular care should be taken to include Indigenous rights holders from the Global South, including Africa and Central and South America. The existing IDS networks provide a launching point for establishing such relationships.

When creating open data stewardship policies and practices to make data open, nation-states, researchers, civil society, and others must abide by the rights of Indigenous nations to govern data. This requires engagement with the data's rights holders beyond mere consultation or advice. Indigenous peoples have the right to decide what is shared or withheld, ultimately affecting how others steward open data. Relationships are key and necessitate that open data actors reach out to Indigenous peoples and not just assume their involvement. While tension exists between IDS and open data, multiple paths forward exist as opportunities to diversify data types and improve sources, stewardship, access, and data quality.

Further reading

Kukutai, T. & Taylor, J. (Eds.). (2016). *Indigenous data sovereignty: Toward an agenda*. Canberra: Australian National University Press. See the editors' Introduction and Chapter 7.

Kukutai, T. & Walter, M. (2015). Indigenising statistics: Meeting in the recognition space. *Statistical Journal of the IAOS*, 31(2), 317–326.

Miaim nayri Wingara Indigenous Data Sovereignty Collective and the Australian Indigenous Governance Institute. (2018). Indigenous data sovereignty communique. *Indigenous Data Sovereignty Summit*, 20 June 2018, Canberra. <http://www.aigi.com.au/wp-content/uploads/2018/07/Communique-Indigenous-Data-Sovereignty-Summit.pdf>

National Congress of American Indians. (2018). *Resolution KAN-18-011: Support of US Indigenous data sovereignty and inclusion of tribes in the development of tribal data governance principles*. http://www.ncai.org/attachments/Resolution_gbuJbEHWpkOgcwCICRtgMJHMsUNofqYvuMSnzLFzOdxBlMIRjij_KAN-18-011%20Final.pdf

the world.¹¹ Concern for many years about the rights of Indigenous peoples led to the establishment of the United Nations Permanent Forum on Indigenous Issues (UNPFII) in 2000.¹² This body is mandated to deal with Indigenous issues related to economic and social development, culture, education, health, and rights. The UNPFII defines Indigenous peoples as those who are

*inheritors and practitioners of unique cultures and ways of relating to people and the environment. They have retained social, cultural, economic and political characteristics that are distinct from those of the dominant societies in which they live. Despite their cultural differences, Indigenous peoples from around the world share common problems related to the protection of their rights as distinct peoples.*¹³

As a means of setting a minimum standard to protect the rights of Indigenous peoples, the UN has developed the *UN Declaration on the Rights of Indigenous Peoples* (UNDRIP).¹⁴ Adopted by the General Assembly of the United Nations in September 2007, it develops a cohesive set of Indigenous peoples' rights through 46 articles.¹⁵ Article 18, in particular, is relevant to Indigenous data rights, stating that Indigenous peoples have the right to participate in decision-making in matters which would affect their rights in accordance with their own procedures. It is also important to note that in the context of Indigenous data, the UNDRIP specifically addresses the collective rights of Indigenous peoples.

Indigenous peoples and data

Data collection: Invisibility and bias

Indigenous nations need data about their citizens, communities, lands, resources, and culture to make informed decisions. Yet few official statistics agencies, researchers, and data collectors make any meaningful concession to Indigenous rights in relation to Indigenous data. Despite being the rights holders in relation to data about them or for them, Indigenous peoples across nation-states remain peripheral to the channels of power through which consequential decisions about Indigenous statistics are made.¹⁶ This marginalisation continues within open data discussions, the open data community, and the ODC itself.

There are also numerous contexts in which Indigenous peoples are invisible in their national statistics systems. A recent study showed that in the global 2010 Census round (2005–2014), under half of the countries that encompassed at least one Indigenous people actually included an Indigenous identifier on their census form.¹⁷ This is particularly a challenge in the Global South in regions such as Africa where Indigenous peoples are not counted or recognised.

In nations where Indigenous data is collected, federal, state, and local governments, as well as researchers, primarily collect data with limited input from the Indigenous nations, communities,

and individuals described. The result is that Indigenous nations rely on external data that largely fails to reflect community needs, priorities, and self-conceptions. This data imbalance threatens self-determination, limits informed policy decisions, and restricts progress toward Indigenous aspirations for healthy, sustainable communities.^{18,19,20} Likewise, reliance on this data by researchers and governments limits the robustness of data-driven research and the validity of policy decisions. The IDS movement has been critical to fostering discussions and actions to improve the quality and relevance of official statistics data to Indigenous peoples and other policy-makers.

Australia – Closing the gap

The Indigenous peoples of Australia are comprised of two separate groups, Australian Aborigines and people from the Torres Strait Islands. Torres Strait Islander people make up about 10% of the total. While conventionally combined for statistical purposes, the two populations are not homogeneous and have significant demographic, social, and cultural differences. Regardless of population, the collection of Indigenous data in Australia has a fraught history. Until amended by referendum in 1967, Section 127 of the Australian Constitution specifically excluded the “aboriginal race” from official population figures.²¹

The first Census of Population and Housing inclusion of Indigenous peoples was in 1971. The reliable collection of Indigenous data has developed only slowly since then and not always in positive directions. The importance of Indigenous data is increasingly recognised within official statistical agencies, policy areas, and public and private administrative entities, yet there remains a fixed focus on what the government wants to know about Aboriginal and Torres Strait Islander people and very little on what Aboriginal and Torres Strait Islander people need to know.

There is also very little true engagement with Aboriginal and Torres Strait Islander people on what data is collected, why that data is collected, or the data needs of Indigenous Australians. Not coincidentally, the primary Indigenous policy which identifies the *Closing the Gap* targets developed by national and state governments is now being refreshed. After ten years, no socioeconomic or health gaps have been closed and only a minority of targets have shown either absolute or relative improvement in that period.²²

As a way of advancing Indigenous data governance in Australia, an Indigenous Data Sovereignty Summit for Aboriginal and Torres Strait Islander leaders across academia, peak bodies, and community organisations was held in June 2018 as a partnership between the Maim nayri Wingari Indigenous Data Sovereignty Group and the Australian Indigenous Governance Institute, with the outcomes delivered in a Summit Communique.²³

Sweden – Absence of data

The European Union General Data Protection Regulation (GDPR) replaced the Personal Data Act in Sweden in 2018 but maintained that the processing of data that reveals ethnicity or race is prohibited. This regulation was first introduced in the Data Act (1973:289) and has severely impaired discussions in relation to Indigenous data.

Sweden is an example of a welfare state where health equity and equality are advanced, and where epidemiology and health statistics are cutting-edge, but also where laws surrounding data have resulted in Sweden being unable to provide any significant data to understand the health and social well-being of the Sámi, its Indigenous people. This circumstance also means that there is little official data produced by the state that Sámi people can take ownership of.

Indigenous data is mainly produced by researchers and is guarded by Swedish ethical protocols that do not take Sámi ownership or control into account.²⁴ However, things have recently been moving forward due to increased pressure from Sámi society and non-governmental organisations (NGOs) that have called for a truth and reconciliation commission, ethical guidelines for Sámi research, a consultation order for Sámi issues, and enhancement of the Sámi Parliament's role in the collection of data. Together with the implementation of a Nordic Saami Convention, this may force the Swedish state to open up for discussions on ownership and governance of data.

Data access, use, and interpretation

There are multiple problems with the ways in which state agencies have collected, stored, analysed, and disseminated data about Indigenous peoples and their lands and resources, which have been well documented across a range of contexts and timeframes.^{25,26,27,28,29} The information that Indigenous nations have access to is often unreliable, inaccurate, irrelevant, and fraught by a long-standing mistrust of data and data systems by Indigenous peoples.³⁰ Furthermore, statistics about Indigenous peoples often perpetuate a narrative of inequality, creating a dominant portrait of Indigenous peoples as defined by their statistically measured disparity, deprivation, disadvantage, dysfunction, and difference.³¹ Data infrastructures are designed based on cultural assumptions that can lead to the systematic misrepresentation of Indigenous peoples (e.g. not allowing for the entry of names that do not conform to the dominant cultures naming conventions).³² Further, conceptualisations of open data purely as digital data produce an area ripe for knowledge co-optation and the theft of Indigenous knowledge as, for example, in cases where researchers or others who collect Indigenous knowledge about the environment (as opposed to digital data) digitise that knowledge and then share it openly without consent or oversight from Indigenous peoples.³³

Indigenous peoples and nations experience a number of challenges in accessing data. These difficulties are driven both by internal and external environments.^{34,35,36} External challenges include inconsistent Indigenous identifiers, the siloing of data by sector, laborious or unclear data

sharing and access protocols, low investments in Indigenous peoples' data science skills, and Indigenous nations' data infrastructures. Internally, for Indigenous nations and peoples, there are difficulties in accessing and paying for programmes to build data capacity and a general lack of access to digital hardware, software, connectivity, and funding for issue-specific capacities.

The inconsistencies and inadequacies of, and lack of access to, existing Indigenous data systems have led to researchers, data repositories, and data service operations being increasingly aware of the need to understand IDS. At the same time, few researchers, governments, or organisations are aware of the appropriate processes. Thus, the aim is for governments, researchers, funders, and others to choose data governance and stewardship mechanisms that better align with Indigenous rights and aspirations; to improve data quality, access, and value; and to invest in building data capacity and infrastructure.

Mexico – Data access

Mexico's Indigenous population (7.4 million) represents about 6% of the total population in the country. Most of Mexico's Indigenous peoples live under marginalised conditions (87.6%). Only 15% report having access to a computer, and only 10% report having access to the internet. This alone represents a big challenge for open data availability within Mexican Indigenous communities, despite the existence of a well-established government open data platform.³⁷ Furthermore, even if accessed, data interpretation and use is limited or irrelevant since there are limited capacities within Indigenous communities to make sense of data available from open data platforms.

The average education level for Indigenous people in Mexico is five years of school attendance, representing less than the completion of elementary school. Nevertheless, as a first step toward recognising IDS, an Indigenous consultation law has been drafted which will establish the right for Indigenous groups to be consulted regarding, among other topics, their natural resources and land use. Such consultation is to be prior to action, free, and informed. This last point implies the right to access information presented in a fashion understandable and relevant to Indigenous communities in advance of consultation.

Data ownership and appropriation

IDS derives from inherent sovereignty and finds its genesis in the oral traditions of Indigenous peoples and community roles and responsibilities.³⁸ The collection, storage, sharing, and use of data have been a strong part of Indigenous cultural knowledge throughout history³⁹ with data storage taking diverse forms, including art, painting, written records, oral traditions, and stories. Open data communities often solely consider data as digital. There is often a focus on quantitative information at the expense of qualitative information based on the lived experience. IDS constitutes a challenge to this narrow conception of data both by protecting knowledge and information that may be taken and digitised and by underscoring that there are other ways of knowing. IDS is then also a broader critique of the turn to digital data in governmental and societal "ways of knowing".

IDS also refers to a collective right to data.^{40,41} While individuals may hold data and have data rights, Indigenous peoples as collectives (or nations) have the right to govern the data about their peoples, lands, and resources. This represents several implications for open data and big data: conceptualising collective rights for data linkage, sharing, and use; protecting data used to describe or compare Indigenous nations; and exploring collective rights of Indigenous nations to privacy and confidentiality. While these and other issue areas have been identified in relation to Indigenous peoples' collective rights to data, the IDG and data stewardship mechanisms and legal strategies have yet to be fully realised. However, some international standards, particularly those set by the UN, support Indigenous peoples' collective rights to govern data.

As previously noted, Article 18 of the UNDRIP stipulates the right to participate in decision-making in matters affecting Indigenous rights and to maintain and develop Indigenous decision-making institutions. As Kukutai and Taylor argue,⁴² the UNDRIP raises urgent questions about the proper role of state machinery in gathering statistics on Indigenous peoples. The UNDRIP also clearly sets out the rights of Indigenous peoples related to data about them. As per Articles 3, 5, and 23, Indigenous peoples have the right to self-determination, inclusive of the right to control and determine what for them as Indigenous peoples constitutes economic, social, and cultural development. The functional planning and implementation of that development is reliant on data – data that Indigenous peoples currently do not have access to or does not exist.

The UN Sustainable Development Goals (SDGs) provide an example of engagement to support Indigenous rights and development according to Indigenous aspirations. The “Transforming Our World: The 2030 Agenda for Sustainable Development” resolution⁴³ refers to Indigenous peoples six times, underscoring the need for the participation of Indigenous peoples at the country level and calling for disaggregated data on Indigenous status on Indigenous peoples' terms, aligning with the UNDRIP and other human rights standards. Since the open data of nation-states plays a key role in tracking progress toward the SDGs, the engagement of Indigenous peoples and respect for Indigenous rights must be fundamental components of this process, as well as the open data principles of the ODC and practices.

In some cases related to the preservation and/or exploitation of natural resources, Indigenous rights are established in legal international protocols. Article 26 of The Cartagena Protocol on Biosafety regarding socioeconomic considerations, for example, establishes that parties “may take into account ... the impact of living modified organisms on the conservation and sustainable use of biological diversity, especially with regard to the value of biological diversity to Indigenous and local communities”.⁴⁴ With regard to the Nagoya Protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilisation, Article 12 on traditional knowledge associated with genetic resources states that parties “shall in accordance with domestic law take into consideration Indigenous and local communities' customary laws, community protocols and procedures, as applicable, with respect to traditional knowledge associated with genetic resources”.⁴⁵

Nagoya Protocol – Who benefits from data?

Benefits stemming from access to data should be mutual among all actors involved, particularly data related to traditional knowledge from Indigenous communities. The Nagoya Protocol⁴⁶ serves as an important reference regarding three main issues:

1. **The importance of establishing proper mechanisms and legal grounds to achieve mutual benefits from data use.** Article 7 establishes that “In accordance with domestic law, each Party shall take measures, as appropriate, with the aim of ensuring that traditional knowledge associated with genetic resources that is held by Indigenous and local communities is accessed with the prior and informed consent or approval and involvement of these Indigenous and local communities, and that mutually agreed terms have been established” (p. 7).
2. **The relevance of raising awareness about data exploitation** regarding traditional knowledge in relation to genetic resources through education and training (capacity building) about data access, interpretation, and use. Article 21 paragraph (g) establishes the need to take measures to raise awareness regarding “education and training of users and providers of genetic resources and traditional knowledge associated with genetic resources about their access and benefit-sharing obligations” (p. 16). All this should be developed in a culturally sensitive fashion.
3. **The importance of maintaining an awareness of the delicate position of the least developed countries** which are inhabited by a number of Indigenous communities that generally have increased difficulty with accessing open data sources.

These issues illustrate the importance of properly establishing common legal and operational grounds regarding data exploitation of any kind in relation to Indigenous knowledge and resources, including not only genetic resources but also cultural, demographic, and other types.

As political entities, Indigenous peoples and nations are more than mere “stakeholders” in data ecosystems.⁴⁷ They have the right to control data about their peoples, lands, and resources. That right is the fundamental difference in the relationship between Indigenous peoples and Indigenous data and other stakeholders’ relationships with Indigenous data. Data stakeholders include nation-states and other governments, researchers, NGOs, Indigenous organisations, funders, and IDS networks. These stakeholders have diverse interests in Indigenous data and, at times, are in situations to govern or steward Indigenous data.

Increases in knowledge and awareness of IDS in Aotearoa/New Zealand, Australia, Canada, Sweden, and the United States (US) are occurring at multiple levels and among a variety of rights holders and stakeholders (e.g. Indigenous nations and peoples, nation-state and local governments, non-profits/NGOs, researchers). Although IDS is not currently an open data concern for most nation-states, incremental efforts to improve the standing of Indigenous peoples and nations in relation to open data are occurring.



Canada – Data governance and informing the nation-to-nation relationship

Canada's Open Government Plan with respect to Indigenous data has been evolving. The plan has gone from promoting programme access for First Nations people to recognition of the developing nation-to-nation relationship between Indigenous nations in Canada, including over 600 First Nations, Metis Nations, Inuit, and the federal Crown. Canada has made a commitment to reflect this renewed relationship in open government planning. It is expected that Canada's Open Government Plan will continue to evolve as these nations assert data sovereignty. Across Canada, the conversation about open data also involves provincial and territorial governments. For example, the Ktunaxa Nation asserts data sovereignty and will work with both the governments of Canada and British Columbia to determine the parameters for opening data related to their Nation or to their people.

Aotearoa/New Zealand – Embedding Māori data sovereignty across the government

Aotearoa/New Zealand is one of the world's most advanced digital nations.⁴⁸ Data is seen as a key national strategic asset, and several policy and legislative initiatives are underway to facilitate easier data sharing and linkage. The Integrated Data Infrastructure (IDI), a world-leading research database, contains de-identified data from more than 50 surveys and administrative datasets across the state, research, and NGO sectors. There is also increasing interest in how a "social licence" can enable more flexible data sharing without individual consent, and the government recently adopted the ODC. Missing, however, are robust, innovative models of data governance and ethics, value creation, and benefit-sharing.

Māori have well-tested "tikanga" (ethics, processes, principles) around the protection and sharing of knowledge for collective benefit that can be readily adapted to digital data environments. Māori Data Sovereignty (MDS) advocates are developing a number of tikanga-based solutions, including models of Māori/iwi (tribal) data governance for the IDI and wider government ecosystem, a "cultural license" as the "social licence" alternative for community acceptability of data use, and a Māori Data Audit Tool to assess organisational readiness to incorporate MDS principles. The potential benefits of embedding MDS principles across government data ecosystems extend beyond Māori to include the wider Aotearoa/New Zealand public.

Challenging the dominant data discourse

As the scale and scope of Indigenous peoples' economic, social, and cultural development accelerates, the demand for data is increasing. Indigenous nations also are seeking methods to protect and control their proprietary information, especially data stewarded by other

governments, non-profits, and researchers. At the same time, non-tribal entities and individuals stewarding and using Indigenous data increasingly recognise the need to protect information about Indigenous nations and peoples. Often, this data collection and usage exists in a vacuum with little to no guidance from others, given the unique circumstances of Indigenous peoples' data. These are issues not just of IDS but also of IDG. These issues are crystallised when set against key principles and goals set out in the ODC.⁴⁹ This section addresses the ODC, discussing assumptions which underlie the movement, including assumptions around democracy and citizenship that are highly problematic for Indigenous peoples. It then presents a path forward for engaging with Indigenous peoples around open data at the nation-state and international levels through existing and nascent IDS networks.

Principle 1: Open by default

As the name suggests, this principle supports the open use of data as the norm, shifting the onus onto governments to justify why and when data should be kept closed (e.g. for security or data protection reasons). In general, the argument behind the open data movements has been that with data in the open, discussions to improve interpretation can then take place. This assumes particular positions of privilege in relation to access to public debate, but also in how that data is interpreted. The problem here is that very few governments have supported or incorporated IDG principles into the values and principles that guide open data practices. A lack of IDG principles results in the absence of processes aligned with Indigenous rights to guide decision-making for how and why Indigenous data should be shared and in the ethical use of that data. More fundamentally, opening Indigenous data by default bypasses entirely the rights of Indigenous peoples to decide what, if any, of their data should be shared, let alone issues of ownership. In the absence of such basic decision-making ability, there is a heightened risk of data misinterpretation and misuse.

While the ODC recognises the need to protect individual privacy and to adhere to domestic laws, it is silent on the issue of collective privacy. Yet this is a crucial factor for Indigenous peoples, especially with the move toward algorithmic decision-making and data mining. National privacy laws are increasingly being revised to strengthen the protection of personal data privacy and impose penalties for data breaches, but they have yet to grapple with more complex issues of collective privacy, and, thus, they offer minimal protection. Likewise, open data licensing regimes do not generally accommodate collective rights, and reuse often relies on use of permissive licensing with few restrictions. These Eurocentric conceptualisations of privacy and licensing challenge IDS collective rights to data and are also problematic in many other societies with imbalances in power, such as those with a post-colonial context.

Principle 5: Improved governance and citizen engagement

As data becomes available through open data portals, more sectors and individuals gain access. However, for Indigenous peoples, one of the key challenges to open data is the risk of interpreting results in the absence of historical, cultural, political, and social contexts. This risk could even further marginalise and stigmatise Indigenous peoples, ostensibly the opposite of what is



intended. There will be unintended and pejorative consequences for Indigenous peoples if open data movements do not acknowledge the bias and values inherent in all data.

The choice of what is counted, the ways that people are categorised and grouped, and the methods of data collection all reflect decisions made by people who may not have the cultural or contextual knowledge to interpret data fairly. For example, findings may show that the prevalence of chronic disease is higher for Indigenous people living in cities than for those who are living in more remote traditional lands and communities. The data could be used to justify decreased resources to communities for chronic disease management; however, a key underlying reason for the difference may be the pre-existing lack of resources in the communities that force people to relocate to seek care.

As per inherent sovereignty rights underscored by UNDRIP, and recognition by nation-states, particularly Aotearoa/New Zealand, Canada, and the US, Indigenous peoples have the right to control the data about them. Additionally, Indigenous data histories are plagued by the misuse and misinterpretation of Indigenous data,^{50,51} sometimes intentionally, sometimes not. Without Indigenous data ownership and control, data can inadvertently result in a perpetuation of marginalisation. But while results may be inadvertent, a lack of knowledge does not exonerate those who cause it. Only through engagement with, and active control over, data by Indigenous peoples can the inadvertent wrongs be righted, and, ultimately, the benefits of data be realised for governance.

Principle 6: Inclusive development and innovation

This principle recognises that open data can “help to identify social and economic challenges, and monitor and deliver sustainable development programs. Open data can also help meet global challenges such as poverty, hunger, climate change, and inequality.” These goals are indeed laudable and highly relevant for Indigenous peoples who, through processes of colonisation and colonialism, incur systematic social, economic, and political disadvantage in most, if not all, national contexts.⁵² For many decades, Indigenous peoples have often been the targets of policy interventions aimed at improving their socioeconomic position, but they have rarely been asked, let alone empowered, to contribute their own solutions. It is not uncommon for Indigenous values to clash with the development goals of national governments and, at times, even intergovernmental organisations and international NGOs, particularly in relation to environmental stewardship. The principle of inclusive development and innovation needs to be tempered with an appreciation of the fraught relationship that many Indigenous communities have experienced in the name of development, democracy, and citizenship, as well as an awareness of the systemic barriers that continue to make it challenging for Indigenous peoples to take leadership of solutions which support their own aspirations (including data-driven solutions).

Developing a new discourse

To date, three national-level IDS networks exist: Te Mana Raraunga Māori Data Sovereignty Network,⁵³ the United States Indigenous Data Sovereignty Network (USIDSN),⁵⁴ and the Maiam

nayri Wingara Aboriginal and Torres Strait Islander Data Sovereignty Collective in Australia.⁵⁵ Similar initiatives are underway elsewhere, including in Scandinavia, where there is not yet an established network, but two Sámi research centres (Umeå University and UiT – The Arctic University of Norway) have recently started discussions on establishing a network. And while the use of IDS terminology is relatively new, the First Nations Information Governance Centre (FNIGC) in Canada has been a leading voice for the rights of Indigenous peoples in relation to their data for over two decades.⁵⁶

The First Nations principles of OCAP* (standing for data Ownership, Control, Access, and Possession) are a set of standards that establish how First Nations data should be collected, protected, used, or shared and are now the *de facto* standard for how to conduct research with Canadian First Nations.⁵⁷ Building off the strong history of Indigenous rights in relation to data, a network focused on IDS is being incubated in Canada.

Currently, these networks are engaging in an informal, and somewhat ad hoc fashion, to share information and strategies, hold joint events, and collaborate on research. In the last three years alone, this spirit of collaboration has produced seven events, ten joint panel/workshop initiatives, and a co-edited book, *Indigenous data sovereignty: Toward an agenda*.⁵⁸ Freely available online, the book was downloaded more than 8 000 times in its first year, reflecting the very high level of international interest in IDS.

In 2017, the founders of the existing networks joined to create the International Indigenous Data Sovereignty Interest Group at the Research Data Alliance (RDA IDS Group).⁵⁹ This group is committed to expanding the IDS discussion beyond North America and Australasia to include the Global South, notably South America and Africa. The RDA IDS Group also supports the drafting of principles for the governance of Indigenous data for adoption and implementation by international research organisations.

IDS network engagement with other data actors has occurred both at the nation-state and international levels across a range of data topics, including privacy, ethics, research data, big data, open data, and many others. Open data is just one aspect of the larger data and Indigenous data sovereignty discussions. Open data agents and actors must recognise that Indigenous peoples and networks, which have already experienced bias, disregard, and limited investment in data capability and capacity, are operating across a range of topics. Of the almost unlimited amount of work still to be undertaken, open data is but one aspect.

The Canadian government has recognised the need to collaborate more closely with Indigenous peoples with respect to open data. First Nations have had input into Canadian open government discussions, and Canada's 4th Plan on Open Government 2018-20 includes Indigenous engagement activities.^{60,61} Members of the RDA IDS Group also have participated in the third and fourth International Open Data Conferences, hosting a side event, the International Indigenous Open Data Summit, at the 2016 gathering in Madrid. The summit drew over 25 participants from Australia, Aotearoa/New Zealand, US, Africa, Taiwan, and other countries. Presentations and discussions focused on three elements: 1) how Indigenous peoples were or were not engaging with open data and why; 2) building connections and community at the intersection of IDS and open data; and 3) identifying the tensions between IDS and open data (and potential paths forward to ameliorate the tensions, while supporting useful data to meet development goals). In addition, members from Te Mana Raraunga and Maïam nayri Wingara



Indigenous Data Sovereignty collective presented at the 2018 UN Special Rapporteur on the Right to Privacy Consultation on Open Data-Big Data workshop held in Sydney.

Despite these efforts, there are ongoing resource and infrastructure constraints to advancing the shared goals and aspirations of IDS partners, including connecting with and expanding the IDS dialogue beyond the wealthy colonial settler states of North America and Australasia. Given the global scope of IDS, it is critical that the colonial and oppressive exclusion of Indigenous peoples in the Global South is not reproduced in IDS discourse and advocacy. To that end, a more robust and coherent international collaboration is needed to achieve impactful outcomes at the intersection of IDS, IDG, and open data. Currently there are no funders or investors driving activity at the nexus of these topics.

The RDA IDS Group and the nation-state network activities highlight three potential steps forward for the open data community in relation to Indigenous data and peoples. First, the necessity of engaging with Indigenous peoples, not merely in a consultative way, but rather as partners and knowledge holders at the table to inform how to steward Indigenous data. Second, the IDS networks need to provide a way forward for engagement with Indigenous peoples. The networks offer pre-existing contacts for non-Indigenous entities to begin working with in order to insert an Indigenous voice and vision into existing open data principles and practices. These networks can also connect non-Indigenous data actors with Indigenous leaders and communities. Finally, there needs to be progress related to research, exemplified by the desire of RDA to adopt and implement principles for the governance/stewardship of Indigenous data for researchers, and the opportunity for work on administrative open data to incorporate IDS.

United States – Generating principles of Indigenous data governance

The University of California Los Angeles (UCLA), the Native Nations Institute⁶² at the University of Arizona, and the United States Indigenous Data Sovereignty Network, with funding from UCLA and the Stewart L. Udall and Morris K. Udall Foundation, hosted an “Indigenous Policy Forum: The Governance of Indigenous Data” in May 2017. The forum fostered discussion on IDS across four stakeholder groups, including tribal leaders, scholars, federal government officials, and non-profit organisations, as well as staff. The dialogue focused on drafting principles of IDG for use by tribes and other entities that govern and steward Indigenous data. Draft principles included recognition of inherent sovereignty and the right to self-determination. With respect to data, this has many implications for: control and access; the protection of Indigenous peoples and their data via ethics that ensure equal explanatory power of Indigenous ways of knowing and equitable outcomes; a focus on intergenerational collective well-being, understanding that data needs to align with Indigenous values for collective well-being across generations; the importance of relationships to the governance of Indigenous data, including respect, responsibility, and reciprocity between Indigenous peoples/nations and other stakeholders; and data governance and stewardship that serves to honour Indigenous knowledge, asserting that such knowledge is of the peoples and includes relationships to the non-human world. Since gathering at UCLA, the draft principles have been discussed at a number of events in order to finalise the principles and a format for sharing.

International Indigenous collaboration – Creating broad principles of Indigenous data governance for policy-makers, researchers, and others

The RDA IDS group, leveraging its “network of IDS networks”, has been drafting principles of IDG for adoption and implementation by scientific organisations, international policy entities, and others. Conceptualised as a set of five to seven key words, each principle would have a brief descriptor statement and then one to three paragraphs placing the principle in context for the intended audience (e.g. researchers). The principles will be accompanied by a companion set of use cases to present the principles in practice. While the principles are currently being drafted, they are envisioned to be broad concepts akin to those in the Te Mana Raraunga Māori Data Sovereignty Network Charter, put forth in the Maiam nayri Wingara Indigenous Data Sovereignty Collective Communique, and recognised in the First Nations principles of OCAP⁸.

Conclusions

The concept of open data that is free to use, reuse, and share is laudable. But as we have described in this chapter, open data principles are in direct tension with IDS and the rights of Indigenous peoples to govern their data. This chapter has articulated, via a description of the history and current state of Indigenous data, case studies, and recommendations, that the path forward to addressing this fundamental tension between IDS and open data is through engagement with Indigenous peoples, both in the drafting of the next round of the ODC and the myriad open data contexts, and the inclusion of IDS and IDG principles within the ODC and in how open data is stewarded. All open data actors have a role in this path forward, including funders, national statistical offices, those building data infrastructures, and sector-specific communities like agricultural or environmental data groups.

Additionally, 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. Such efforts could range from projects to increase Indigenous community data science capacity to encouraging engagement between nation-states and IDS networks in order to create open data policies around the stewarding of data in accordance with IDS and IDG.

This chapter also describes how government is more multi-layered than current open data governance assumes, and that data for governance requires a number of sources. Indigenous nations are political entities, and, as such, are another government actor in the open data world. This challenges the open data binary with one government actor, the nation-state. In addition, the history of Indigenous data and IDS illustrates that the data needed for governance (for Indigenous nations or other nation-states) requires information from more sources and perspectives than are currently available. As a result, IDS calls for more nuanced approaches to data than open data binaries often assume. Thus, the stewardship of open data arises as a key area of action requiring engagement with Indigenous peoples.