DEEP DIVE

Indigenous Data Sovereignty

Indigenous Data Sovereignty

At-a-Glance

- **Centers rights of Indigenous Peoples to control their Indigenous knowledge, practices, and related data** from, and about, their community.
- **Predicated on community data ownership and data governance.** Engagement needs to respect specific principles governing the community and co-create a bespoke approach.
- Suitable in geographies where Indigenous communities have been clearly identified, but a **lack of clear identification of Indigenous data** and communities is a challenge.

Indigenous data sovereignty (IDS) is the right of Indigenous Peoples (IP) to determine the means of governing their data, from whom it has been derived, or to whom it relates. Historically, Indigenous communities have been unrecognized, or unaccounted for, and the identification and recognition is very much an ongoing process. They are often not recognised as the legal right holders of their knowledge and data. IDS centers on Indigenous collective rights and authority over the management and governance of data about their people, communities, lands, lifeways, and natural resources.[1] This data is essential to make informed decisions to contribute to their future. In line with Carroll Rainie, et al., IDS provides a framework for: (1) the **stewardship of data**, asking "fundamental questions about ownership, representation, and control"; (2) the **right to control data** from and about IP, "articulating both individual and collective rights to data access and to privacy"; and (3) challenging dominant approaches to data ownership, licensing, and use, as to echo beyond Indigenous contexts, "**drawing attention to the power and post-colonial dynamics** within many data agendas."[2]

IDS is becoming an increasingly relevant topic of discussion, as limited opportunities for data sharing have focused attention on the protection of Indigenous rights and interests and the consequently unique structures of participation in data governance. For example, the British Columbia First Nations' Data Governance Initiative formalizes in its strategic framework a governance approach to Indigenous data and establishes an understanding of the vision, objectives, scope, and requirements.[3] It is led from the ground up, where First Nations are fully involved in leading, planning, and decision making.

The focus on IDS emerged as part of an effort of the Indigenous cultural and intellectual property rights discourse, research ethics communities, and principles under the U.N. Declaration on the Rights of Indigenous Peoples (UNDRIP).[4] With most Indigenous data held within non-Indigenous institutions and governments, opportunities for increasing control are connected to properly identifying Indigenous data. These objectives are critical to afford greater participation in data governance activities.[5]

1. Kukutai, T., & Taylor, J. (2016). Indigenous data sovereignty: Toward an agenda. ANU press.
2. Rainie, S.C., Kukutai, T., et al. (2019). Indigenous data sovereignty, State of Open Data, OD4D.
3. See their communications for more information.
4. Carroll, S., Rodriguez-Lonebear, D., & Martinez, A. (2019). Indigenous data governance: strategies from United States native nations. Data Science Journal, 18.
5. Rainie, S.C., Schultz, J.L., et al. (2017). Data as a strategic resource: Self-determination, governance, and the data challenge for Indigenous nations in the United States. International Indigenous Policy Journal, 8(2).

"Indigenous communities around the world have experienced the adverse consequences of being excluded from data, of having no say in how they will be measured, and of having their lived experience ignored," says Gwen Phillips, Ktunaxa Nation Indigenous data advocate.[6] IDS runs and exists parallel to Western practices. Data sovereignty refers to the understanding that data is subject to the laws of the nation within which it resides. IDS perceives data as subject to the laws of the nation or peoples from which it is collected.

Indigenous data refers to data in any format, by or about IPs, and that impacts lives at the collective and/or individual level. Above all, it should include Indigenous knowledge and practices, such as art, writings, dreams, oral traditions, etc. The way of thinking of data sharing, community and what constitutes data is different from Western perspectives. IDS challenges notions of data ownership, individual primacy, and governance that dominate Western thinking. While the discourse is nascent, it aims to respect the authority of Indigenous hierarchy structure and ensure that engagement with Indigenous data is done through an equitable and respectful platform. There are diverse challenges regarding Indigenous data. Each community has a wide range of tangible and intangible cultural features, knowledge, and data. The main bottlenecks of data ownership, responsibility, and common errors in data also exist in cultural materials.

In line with the International Labour Organization Convention 169 (ILO 169) and UNDRIP, data governance approaches should respect, include, and promote Indigenous issues in their work. The core principles of UNDRIP are self-determined development; respect for IPs' knowledge, cultures, and traditional practices that contribute to sustainable and equitable development; and free, prior, and informed consent (FPIC). FPIC is a specific right that pertains to IPs and allows them to give or withhold consent to a project that may affect them or their territories. Once People or communities have given their consent, they can withdraw it at any stage. FPIC enables IPs to negotiate the conditions under which the project will be designed, implemented, monitored, and evaluated. The Food and Agriculture Organization (FAO) has developed a policy on Indigenous and Tribal Peoples that ensures it makes all due efforts to respect, include, and promote Indigenous issues in relevant work.[7]

> **The CARE Principles for Indigenous Data Governance**[8] are designed to guide inclusion in data governance across data ecosystems. They create a **paradigm shift** from discussions to **value-based engagements and enhance equitable participation** for Indigenous communities in data use. The principles complement the data-centric nature of FAIR principles (findable, accessible, interoperable, and reusable). The aim is for data stewards and data users to include the CARE (and FAIR) principles when using Indigenous data. CARE informs the use and needs to be applied across all stages of a data lifecycle, likely in contexts that support polycentric, nested, or distributed approaches to governance. The CARE principles point data producers and repositories towards practices that consider the people and purpose for which data exist and are used. The principles focus on:

6. Global Partnership for Sustainable Development Data, Barbero, M., & Richards, K., (2022.) Participatory data governance: How small changes can lead to greater inclusion.
7. Reflected in FAO Environmental and Social Management Guidelines and the Guide to the Project Cycle.
8. Carroll, S.R., Garba, I., et al. (2020). The CARE principles for indigenous data governance.

1. Collective benefit for inclusivity and innovation, resulting in equal outcomes and better governance.
2. Authority to control and govern data, enhancing self-determination.
3. Responsibility to improve capacity development investments and capabilities.
4. Ethics to reduce harm, increase benefits, and enhance data availability for future use.

Some existing tools can be used to better identify and classify Indigenous data and devise respectful ways of working with data. Traditional Knowledge (TK) labels use existing local protocols for access and use of recorded heritage digitally circulating outside community contexts.[9] They are designed to identify and clarify which material has community-specific restrictions regarding access and use. These labels were developed with Indigenous People and clarify conditions for data sharing and engagement compatible with existing rules and protocols for using and sharing their data.

Culture Knowledge Management Systems (CKMS) can be used to preserve, organize, and resettle digital and cultural knowledge into communities. A CKMS consists of a configurable model for controlling a broad range of cultural information while observing protocols of culture and community expectations. One example is Mukurtu (MOOK-oo-too), which aims to empower communities to manage, share, narrate, and exchange their digital heritage in culturally relevant and ethically minded ways. Although not related to agriculture, the platform was built with an open, community-driven approach, to help build a platform that fosters relationships of respect and trust.

The BC Labels focus on accurate provenance, transparency, and integrity in research engagements. The BC Labels are digital markers that define community expectations and consent about appropriate use of collections and data. They connect data to people and environments over time. Ten BC Labels provide a practical application of Indigenous data governance principles to issues of access and benefit sharing for genetic resources.

Indigenous data sovereignty in agriculture

Identification of Indigenous farmers is challenging. It can often be difficult to distinguish Indigenous farmers from smallholder farmers, especially in countries with large Indigenous populations. However, Indigenous farmers are distinguished by their role as custodians of a systematic body of knowledge that results from the accumulation of experience, informal experiments, and understanding of their environment.[10]

Self-identification is usually the most important means for identifying Indigenous individuals, families, and communities, though outside actors should be aware that factors such as racism and other forms of bias or discrimination can discourage people of Indigenous ancestry from identifying themselves as such. Engaging with local, regional, and national Indigenous governance and representative bodies to identify Indigenous communities is an important step in working with such groups.

---

9. There are many other examples; Sq'éwlets People,
10. Dagne, T.W. (2021). Embracing the data revolution for development: A data justice framework for farm data in the context of African Indigenous farmers.

Agricultural data, a constituent of localized data, are a result of traditional knowledge in Indigenous practices. The activities in agriculture among Indigenous communities include engagements with farm data, with a lens on environmental and ecological views. Most global farm data is related to the physical, environmental, and ecological systems, but lack data on Indigenous knowledge and practices. In digital agriculture, the contribution of Indigenous knowledge systems to farm data is not adequately recognised.

There is limited information on IDS in the agriculture data governance arena. However, there are processes and enabling mechanisms promoted by IDS that provide a good model for engagement with Indigenous communities and data. Ensuring adherence to these principles would be important in any project that works with Indigenous data or people.

IDS can be compatible with insights on technical assistance, financing, and commercialization services. Some cooperatives and organizations focus on providing services to IP and supporting their practices. ANEI Coffee Cooperative, for instance, seeks to enhance the conditions of its Indigenous members and families (see [Case Study](#)). They provide for training and demonstrations by agronomists and environmental engineers on ecological practices to enhance productivity. They also aim to enhance access to credit to producers and provide a platform to market their coffee at reasonable prices.

IDS data can also explore localizing data to promote data justice. Data must be collected, stored, and processed in the jurisdiction of collection. The scope and stringency of localization can be worrisome, however, it provides a way to ensure data is located within the community and can help build data infrastructures and capacities within the community.

Why Indigenous data sovereignty and data governance approaches matter

Indigineous data sovereignty provide for a number of benefits, including:

- An alternative approach and principles to data governance and management for Indigenous, Tribal and local communities away from harmful dominant channels of power.
- A collective community model of data governance, stewardship, and control.
- The option for Indigenous and Tribal communities to assert their claims to data sovereignty, rooted in inherent and inalienable rights to self-determination.
- Meaningful involvement of IP in data governance to ensure data is used for purposes that support their beneficial outcomes. It underscores underlying intentions of IP, including the implementing processes.

Key components of success

- **Community data governance and collective ownership** is implied in IDS. IP possess the right to govern their data based on collective needs and beliefs. Accounting for a community model of data ownership is a central tenet. For this reason, there have been (research) agreements and mandates that fit the needs of specific communities. The Maori community developed a [Maori Data Audit Tool](#) to assess organizational readiness to work with Maori Data Sovereignty principles.[11]

- **Data use** influences data governance.[12] Norms depend on whether the user is Indigenous, who collects the data, whether it originated from multiple sources, and whether it can be presented from the perspective of the groups that contribute to the data. In Canada, First Nations Health Surveys are designed, implemented, and analyzed by First Nations, giving communities control in collection, access, and use of their data. Institute for Clinical Evaluative Sciences (ICES) worked with Indigenous partners to promote unique relationships covering data-governance and -sharing agreements.

- **Privacy and confidentiality** laws do not distinguish Indigenous data. However, attitudes and expectations of privacy may be different for IP. Definitions of private data might differ. In Western tradition, private information comprises personal information on, for example, finance or health. For IPs, it may comprise other activities like involvement in ceremonies, practices on hunting or gathering, as well as what is needed for the community. Defining data privacy is important and ensures data integrity and quality. This may be stewarded via a variety of mechanisms, including preservation by Indigenous entities, funding agencies, academia, or data repositories.

- **Data infrastructures** can support respectful use of Indigenous knowledge. ORCID and Local Contexts created a platform that enables research or use of Indigenous data, with the consent from the community.[13] When a research request is approved, the ORCID record will be updated, and researchers can access community data, where possible, and allow Indigenous knowledge to be used in innovative and respectful ways.

Challenges/pitfalls of Indigenous Data Sovereignty

- **The lack of access to data about communities poses major internal and external challenges.** External challenges include unreliable Indigenous identifiers, isolation of data for different sectors, unclear protocols for data sharing and access, and low levels of funding for data science skills and data infrastructure. Internal challenges include lack of access to hardware, software, connectivity, and data capacity.

- **While IDS theoretically allows for an user-centric model of participation with Indigenous data, there is very limited guidance available on how to implement such models.** Each IDS approach will need to be respectful to the context of specific principles governing the concerned community, and co-create a bespoke approach together with the community. This will require investment in time, processes, and money.

- **Most literature and practices arise out of experiences of Indigenous communities in Western societies,** i.e., Australia, New Zealand, U.S., and Canada. Low- and middle-income countries (LMICs) hold other challenges, such as limited identification of Indigenous communities. Networks, such as the Indigenous Peoples of Africa Coordinating Committee, look to bring more clarity to the perspectives of communities from LMICs.

---

11. Daly, A., Mann, M., & Devitt, S. K. (2019). Good data. Lulu.com.
12. Kukutai, T., & Taylor, J. (2016). Indigenous data sovereignty: Toward an agenda. ANU press.
13. Local Contexts also created the Local Contexts Hub, which allows institutions and researchers to generate notices and engage with Indigenous communities about the appropriate use of the TK & BC Labels.

Financial viability and sustainability

IDS requires significant investment from the government, the private sector, and development sector stakeholders. However, it is important to note that societies benefit from Indigenous knowledge in innumerable ways. Investing in a mechanism that provides fair value to the original creators cannot be measured through a business viability model.

> USAID's Policy on Promoting the Rights of Indigenous Peoples (PRO-IP) provides guidance to those looking to partner with Indigenous and local communities and design programs and activities that may affect IPs. It promotes more thoughtful and direct engagement of IPs in the design, implementation, and monitoring of projects, ensuring that outcomes are in line with the communities' self-determined objectives and that their challenges are addressed in the program cycle. "Meaningful engagement of IPs as partners in the development process is an essential part of conflict management and resolution, enhancing governance and human rights, reducing poverty, and sustainable environmental management."[14] While PRO-IP promotes local ownership and development approaches through partnership and co-creation with IPs and defines traditional knowledge and cultural considerations as a core element throughout the program design, the policy does not connect with IDS and data governance approaches. There could, therefore, be opportunities to make this connection more explicit in future iterations of the policy.

How can stakeholders create an enabling ecosystem for IDS?

Role of the government
- **Pass strict data protection laws** that protect the culture, data, and knowledge of IPs.
- **Co-develop data governance mechanisms with Indigenous communities.** These can act as guiding principles for engagement with the private and development sector.
- **Build communities' data capacities and provide direct access to data concerning their communities.**

Role of the private sector
- **Assume fiduciary responsibility** in dealing with data related to Indigenous communities.
- **Publicly provide information** on consultation processes used to engage with Indigenous communities.

Role of the development sector
- **Invest in tools that can help identify and protect indigenous knowledge and data.**
- **Promote research on IDS models that originate in LMICs.**
- **Provide guidance on engagement of the private sector with Indigenous communities,** with adherence to Indigenous, national, and international guidelines.

---

14. Policy on Promoting the Rights of Indigenous Peoples (PRO-IP).

The ANEI Indiginous Organic Coffee Cooperative in Colombia, with help from Ethos Agriculture, applied an approach where eco-centric values, bio-cultural identity, and Indigenous knowledge are woven into governance, value chain improvement, and conservation, in order to reclaim sustainability away from colonial, capitalistic paradigms or pre-defined governance activities. This involves aspects of Indigenous data governance and data cooperatives aiming to demonstrate a process for smallholder farmers and their communities to gain agency in defining their digital paths, data governance, and sustainability priorities and future goals. This "co-design" allows for an alignment of values to better coordinate resources. The case study gives insights into participatory co-design processes, eco-centricity, carbon trading, data co-ops, and Indigenous data governance.

Traditional Knowledge labels provide information and clarify specific responsibilities with regard to access and use of Indigenous knowledge, such as sacred or gender-specific material, and information especially used for outreach. TK labels enhance the dialogue between IPs and external users of cultural knowledge and traditional expressions. The labels provide opportunities for Indigenous groups to take control of their cultural heritage and provide awareness to users on incorporation of digital heritage items in a culturally sensitive way. Data includes digitized visual arts and recorded, written, and oral histories and stories. TK labels are a standard while operating any cultural heritage introduced by Indigenous communities. They also present a way to acknowledge ownership and use of work in the public domain. To use the information, users must be registered in the Local Contexts Hub, which provides community control via selection and delivery to facilities, data repositories, and other parties.

The Cadasta Foundation, Waatavaran, and Esri work together to create a fiduciary and communal data governance approach based on Indigenous and traditional knowledge governance with the aim to secure land and forest rights of Indigenous and local communities in West India. The process covers a stewardship of data model where the local community stewards and owns their data. The case study provides insights into data stewardship, communal data control and ownership, meaningful deliberative participation, and Indigenous data governance.

Mapeo is a free digital mapping toolset for communities to document, monitor, and map many types of data. Led by Digital Democracy, its main goal is to help communities to own their data, stating "[w]e're part of a larger movement that is fighting the monopolization of digital knowledge by a few Silicon Valley 'cloud' companies."[15] Mapeo's tools have been co-designed by Indigenous communities. Digital Democracy developed a set of principles called local-first software that lock autonomy into the data and code. Mapeo uses a P2P database, providing communities with a sovereign governance structure to keep ownership, management, and control over data. This means that the data collected will only exist on the device it is collected on or created with until shared with other participants or external actors. The data will never be stored on external servers or shared with app maintainers or others unless the owner chooses to do so. The individuals or communities generating the data are the data stewards, managing their own data and controlling who has access. "Once knowledge is distributed in this way, people are able to take cooperative ownership of the data they generate."[16]

---

15. McKelvey, K. (2020). Cooperative Ownership of Data without Blockchain.
16. Idem.