

University of Waterloo
Indigenous Data Sovereignty Data Storage Analysis Report

Jeff Doctor - Impact Strategist

Agenda

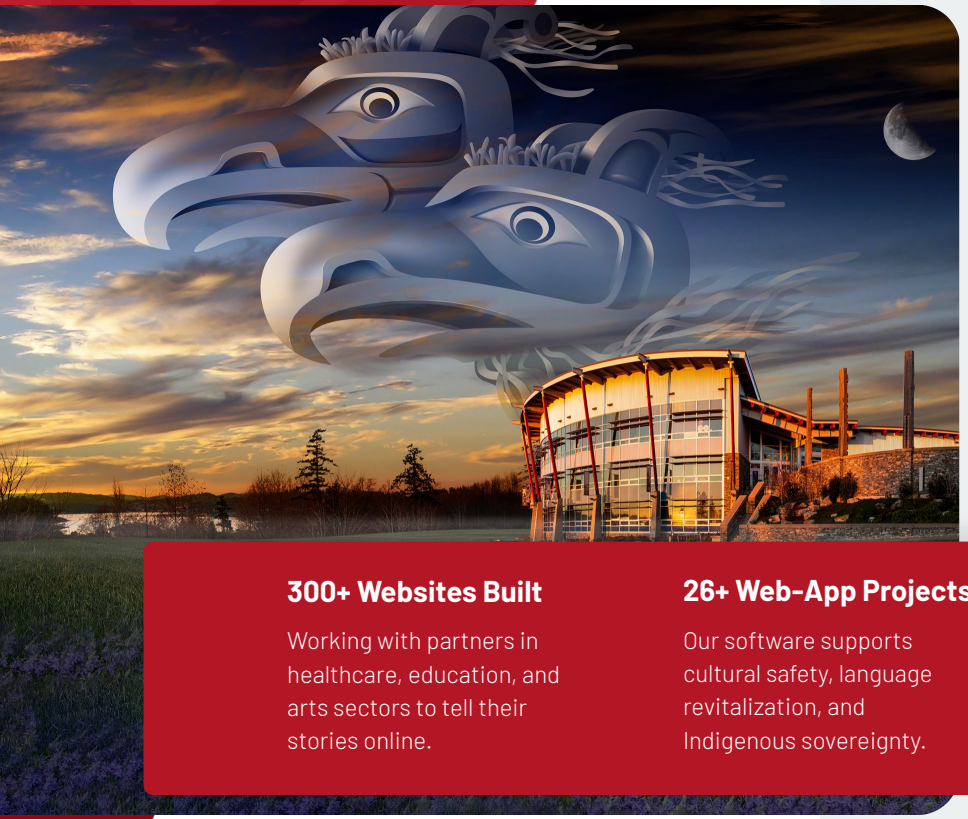
- Project Context
- Indigenous Data Sovereignty
- Report Findings
- Q & A
- Open Discussion



Animikii

Animikii

Indigenous Technology



Our Story

For 20 years Animiki has been a values-driven Indigenous technology company with products and services focused on web application development.

We recognize that technology itself is not a solution to the pressing issues we face, rather it must be guided by a system of values that promote mutual understanding and equity among diverse peoples.

300+ Websites Built

Working with partners in healthcare, education, and arts sectors to tell their stories online.

26+ Web-App Projects

Our software supports cultural safety, language revitalization, and Indigenous sovereignty.

35+ Pathfinding Reports

Our socio-technological software development approach provides transparency to our partners and holds us accountable to what we are creating with them.

“the term ‘research’ is inextricably linked to European imperialism and colonialism”

- Linda Tuhiwai Smith

Decolonizing Methodologies : Research and Indigenous Peoples

Responding to the 2021 Tri-Agency Research Data Management Policy:

- **“a distinctions-based approach is needed to ensure that the unique rights, interests and circumstances of the First Nations, Métis and Inuit are acknowledged, affirmed, and implemented.”**
  - Item 1: Review and analysis of current data storage / data repository options at UWaterloo from an Indigenous data sovereignty lens
  - Item 2: Research and recommendations on data storage options that support Indigenous data sovereignty
  - Item 3: Indigenous data storage governance protocols for UWaterloo-based data storage options



Collective Rights and Responsibilities

Adhering to UNDRIP, TRC, NIMMIWG

Indigenous Peoples — and the lands, waters, plants, animals, etc. they relate to — have intrinsic rights, not "stakes", that are integral to their unique cultures, identities, traditions, languages, and institutions. The **United Nations Declaration on the Rights of Indigenous Peoples** (UNDRIP) defines these intrinsic rights in further detail: **Articles 11.1 / 13.1 / 31.1**

We are also responding to the **Truth and Reconciliation Commission Calls to Action**: **43 / 92**

As well as Call for Justice **15.7** from the **National Inquiry into Missing and Murdered Indigenous Women and Girls**.

Indigenous Peoples' Rights in Data

Data For Governance	Governance of Data
Right to self-determination: the ability to organize and control data in relation to a collective identity	**Right to govern:** the right to lead and collaborate in the development and implementation of protocols and in decisions about access to data
Right to reclaim: the right to reclaim, retain, and preserve data, data labels, and data outputs that reflect Indigenous Peoples' identities, cultures, and relationships	**Right to define:** the right to define lifeways of knowing and being including how they are represented in data
Right to possess: the ability to exercise jurisdictional control over the ways that data flow/move/are queried	**Right to privacy:** the protection of collective identities and interests from undue attention, also including the possibility of requesting omission and/or erasure
Right to use: the ability of individuals and collectives to use data for their own purposes	**Right to know:** the ability to track the storage, use, and reuse of the data and who has had access to them
Right to consent: the expression of digital autonomy and the ability to assess risks and accept potential harms	**Right to association:** the recognition of provenance and terms of attribution
Right to refuse: the right to say "no" to certain uses of data	**Right to benefit:** the opportunity to benefit from the use of data and equitable benefit sharing from derivatives of data

Source: Maui Hudson et al., "Indigenous Peoples' Rights in Data: A Contribution toward Indigenous Research Sovereignty," *Frontiers in Research Metrics and Analytics* 8 (May 4, 2023), https://doi.org/10.3389/frma.2023.1173805.

What are Indigenous data?

> **"Indigenous data are data, information, and knowledge, in any format, that impact Indigenous Peoples, nations, and communities at the collective and individual levels; data about their resources and environments, data about them as Individuals, and data about them as collectives."**
>
> Stephanie Russo Carroll et al., "Operationalizing the CARE and FAIR Principles for Indigenous Data Futures," *Scientific Data* 8, no. 1 (April 16, 2021): 108, https://doi.org/10.1038/s41597-021-00892-0.

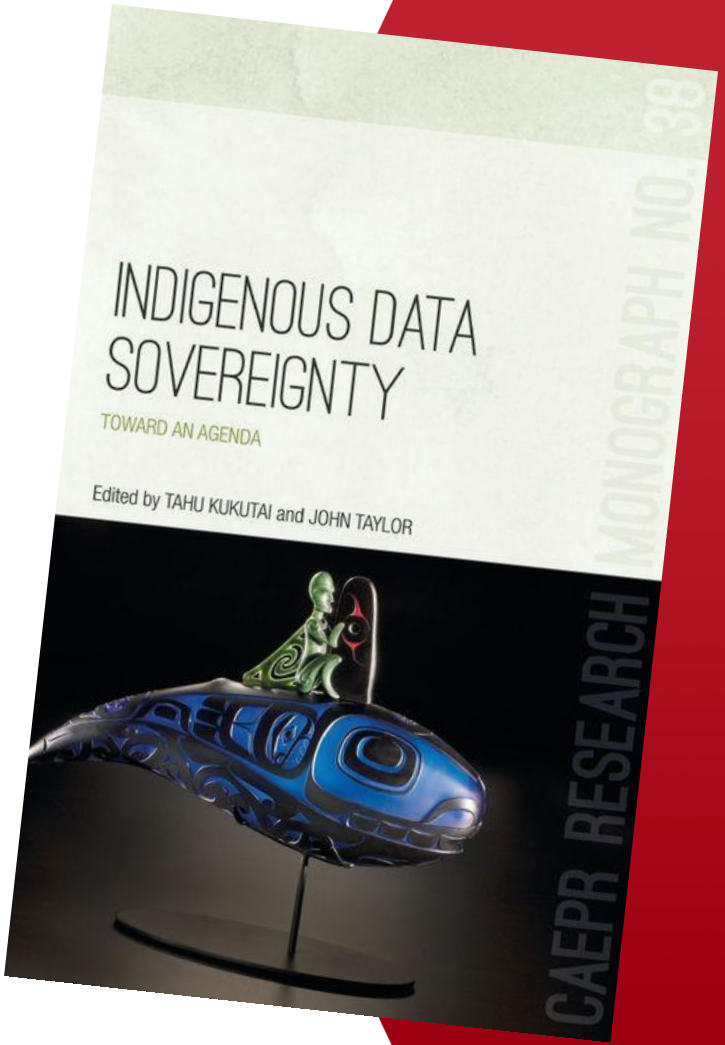
"Nothing About Us, Without Us"

- **Data about us**
  - Censuses, demographics, health surveys, birth/marriage/death records, maps, "research"
- **Data from us**
  - Oral histories, songs, arts, regalias, languages, ceremonies, teachings, creation stories
- **Data of our land**
  - Geographies, plants, medicines, animals, "all our relations", titles and obligations to these relations

Indigenous Data Sovereignty affirms the inherent right

of Indigenous Peoples to govern the collection, access, analysis, interpretation, management, dissemination, and reuse of data related to them.

– derived from *Tahu Kukutai and John Taylor, eds., Indigenous Data Sovereignty: Toward an Agenda*



#BeFAIRandCARE

From the Global Indigenous Data Alliance (gida-global.org)

FAIR: Findable, Accessible, Interoperable, Reusable (go-fair.org)
CARE: Collective Benefit, Authority to Control, Responsibility, Ethics

CARE Principle	Recommendation 1	Recommendation 2	Recommendation 3
Collective Benefit	For inclusive development and innovation	For improved governance and citizen engagement	For equitable outcomes
Authority to Control	Recognizing rights and interests	Data for governance	Governance of data
Responsibility	For positive relationships	For expanding capability and capacity	For Indigenous languages and worldviews
Ethics	For minimizing harm and maximizing benefit	For justice	For future use

OCAP® – For First Nations Data

- OCAP® is a registered trademark of the First Nations Information Governance Centre (FNIGC)
  - fnigc.ca/ocap-training
- Established in 1998 over the need to assert First Nations' agency over First Nations health research, mandate comes from the Assembly of First Nations

Ownership

"Refers to the relationship of First Nations to their cultural knowledge, data, and information. This principle states that a community or group owns information collectively in the same way that an individual owns his or her personal information."

Access

"Refers to the fact that First Nations must have access to information and data about themselves and their communities regardless of where it is held."



Control

"Affirms that First Nations, their communities, and representative bodies are within their rights to seek control over all aspects of research and information management processes that impact them."

Possession

"While ownership identifies the relationship between a people and their information in principle, possession or stewardship is more concrete: it refers to the physical control of data."

National Inuit Strategy on Research

Inuit Tapiriit Kanatami, *National Inuit Strategy on Research* (Ottawa, ON: Inuit Tapiriit Kanatami, 2018), https://www.itk.ca/wp-content/uploads/2020/10/ITK-National-Inuit-Strategy-on-Research.pdf.

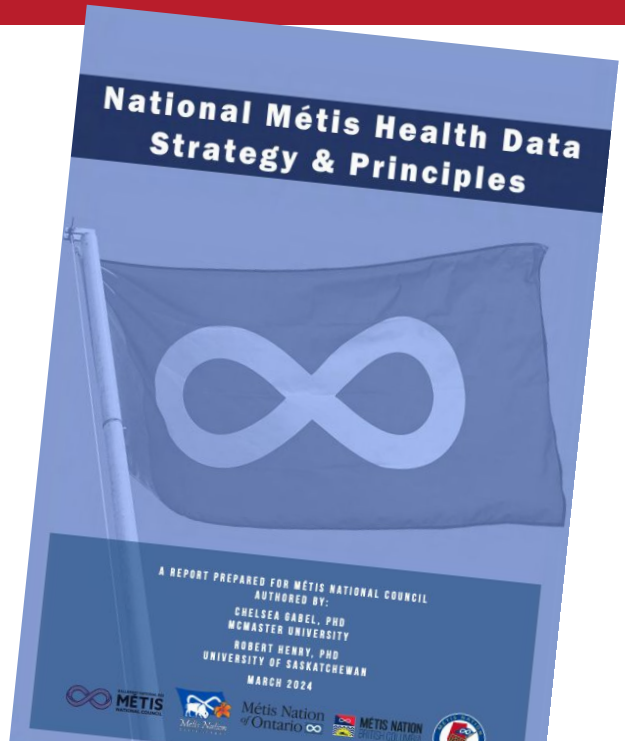
1. **Advance Inuit governance in research**
2. **Enhance the ethical conduct of research**
3. **Align funding with Inuit research priorities**
4. **Ensure Inuit access, ownership, and control over data and information**
5. **Build capacity for Inuit Nunangat research**



National Métis Health Data Strategy & Principles

Chelsea Gabel and Robert Henry, "National Métis Health Data Strategy & Principles" (Métis National Council, March 2024), https://macsphere.mcmaster.ca/bitstream/11375/30688/2/MNC_Final_Draft_Report_March_2024.pdf.

1. **Reciprocal relationships**
2. **Respect**
3. **Safe and inclusive environments**
4. **Diversity**
5. **Research ethics**
6. **Métis context**



Let's say we get our digital data back...

Where do we put them?

Internet Infrastructure

- Speed, bandwidth, reliability
- Geography, legal jurisdiction

Human "resources"

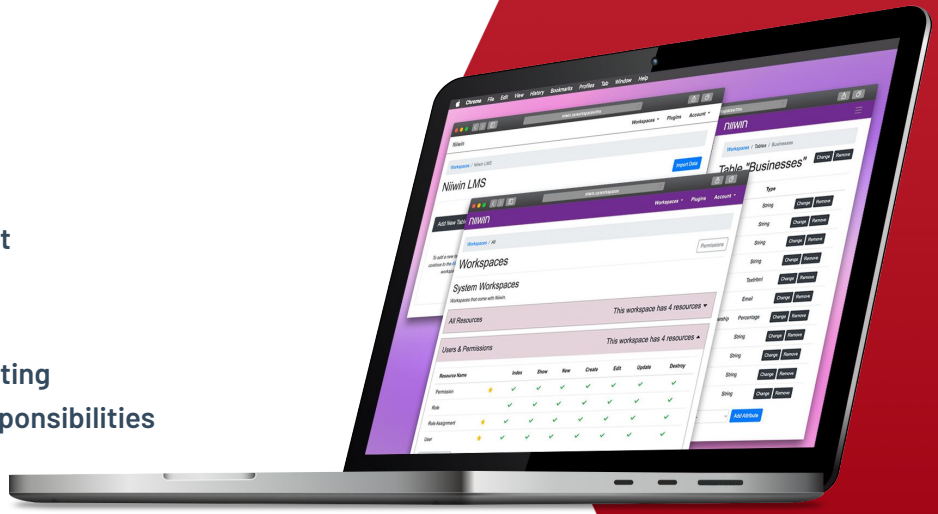
- Skills, wages, reliability
- Knowledge, cultural safety, trust

Hosting

- Trusting in Big Tech vs. self-hosting
- Legal jurisdictions, security responsibilities

Software

- Software-as-a-service (renting)
- Full licence or building it yourself (owning)





A Healthy Digital Ecosystem

Layers of the Digital Stack

- Devices
- Applications
- Data
- Tooling
- Soft Infrastructure
- Hard Infrastructure
- Resources

Item 1: Review and Analysis of Current Data Storage/Data Repository Options at UWaterloo from an Indigenous Data Sovereignty Lens

Category	Strengths	Weaknesses
Business Model	• Free to deposit and view • Open-source software • Developed and supported by non-profits • Diversity of funding sources	• Fees • Closed software • Developed, supported, and sold for-profits company • Risk of defunding
Governance	• Indigenous governance, collaboration, or involvement • Publicly available software development plans	• No Indigenous governance, collaboration, or involvement • No software development plans
Hosting	• Self-hosting available • Hosted within Canada • Choice of cloud provider	• Only cloud hosted • Hosted outside of Canada • Only hosted on Big Tech servers
File Upload & Handling	• Ability to embargo files	• No ability to embargo files
Account & User Info	• Flexible permissions and workspaces	• Limited permissions and workspaces

Adapted from: Derek Murphy and Julian Gautier, "A Comparative Review of Various Data Repositories," *A Comparative Review of Various Data Repositories* (blog), July 25, 2017, https://dataverse.org/blog/comparative-review-various-data-repositories.

Item 1: Review and Analysis of Current Data Storage/Data Repository Options at UWaterloo from an Indigenous Data Sovereignty Lens

Category	Strengths	Weaknesses
Metadata	• Support Traditional Knowledge and biocultural labels • Allows for Indigenous data provenance • Support for controlled vocabulary terms with URLs	• No support for Local Context Labels • Does not allow for Indigenous data provenance • Automatically generated URLs cannot be changed
Governance	• Indigenous governance, collaboration, or involvement • Publicly available software development plans	• No Indigenous governance, collaboration, or involvement • No software development plans
Terms & Licensing	• Users have to actively acknowledge conditions of use when downloading data • Terms of use and copyright • Ability for communities and individuals to withdraw consent	• Anyone can download data without any acknowledgement of conditions • Only used for open data (maximum reuse) • Data are open for everyone to be accessed indefinitely
Preservation	• Designed for long-term preservation	• No public data preservation plans or strategies

Adapted from: Derek Murphy and Julian Gautier, "A Comparative Review of Various Data Repositories," *A Comparative Review of Various Data Repositories* (blog), July 25, 2017, https://dataverse.org/blog/comparative-review-various-data-repositories.

Item 1: Review and Analysis of Current Data Storage/Data Repository Options at UWaterloo from an Indigenous Data Sovereignty Lens

Borealis (Dataverse)

- **Strengths**
  - Compatible with Local Contexts Hub
    - TK Labels
  - Open Source software
  - Designed for data security, reliability, and longevity
  - Discoverable and Accessible
- **Weaknesses**
  - Ownership and Access control
    - Limited metadata features
  - Lack of Indigenous involvement in governance and design
  - Doesn’t account for Indigenous data in the public domain that probably shouldn’t be
  - Core software developed by Harvard

Item 1: Review and Analysis of Current Data Storage/Data Repository Options at UWaterloo from an Indigenous Data Sovereignty Lens

Federated Research Data Repository (FRDR)

- **Strengths**
  - FAIR Compliant
  - Well documented
  - Designed for larger datasets
  - Discoverable and Accessible
- **Weaknesses**
  - Shares many of same weaknesses as Dataverse
  - Meant for public data
  - Relatively new (2021), proprietary software
  - Single source of funding and development (Digital Research Alliance of Canada)

Item 2: Research and Recommendations on Data Storage Options that Support Indigenous Data Sovereignty

UWaterloo Central Information & Systems Technology Services Supported Software

- **Microsoft Office365 Products**
  - **Strengths**
    - Familiar to most people
    - Integrated apps
    - Supported by IT
  - **Weaknesses**
    - Big Tech
    - SaaS and Enshittification
    - Generative “AI”
    - Expensive over time

Item 2: Research and Recommendations on Data Storage Options that Support Indigenous Data Sovereignty

UWaterloo Central Information & Systems Technology Services Supported Software

- **Bulk Storage and NetApp Snapmirroring**
  - **Strengths**
    - Easier than trying to self-host
    - Supported by professionals
    - Scalable computational power
    - Security and backups
  - **Weaknesses**
    - Affordable up front but may be expensive over time
    - Requires some IT knowledge (help available)

Item 2: Research and Recommendations on Data Storage Options that Support Indigenous Data Sovereignty

Recommended Data Storage Options That Support Indigenous Data Sovereignty

- **Mukurtu:** Digital cultural heritage CMS
  - **Strengths**
    - **Free and open-source project**
    - **Built with Indigenous collaboators**
    - **Proven Success**
    - **Can be self-hosted or with a hosting partner**
    - **Designed with IDSov in mind**
  - **Weaknesses**
    - **Core development in US, vulnerable to defunding**
    - **Obsolete Drupal 7 (upgrade plan in place)**
    - **Not for research per se, more for cultural heritage**
    - **Lacks feature parity with commercial tech**

Item 2: Research and Recommendations on Data Storage Options that Support Indigenous Data Sovereignty

Recommended Data Storage Options That Support Indigenous Data Sovereignty

- **Niiwin (Animikii):** Database / app builder / content management system
  - **Strengths**
    - **Web-app, no app store required**
    - **Custom data structure design**
    - **Self or cloud host**
    - **Flexible user role and workspaces**
    - **Designed to respect IDSov**
  - **Weaknesses**
    - **SaaS product not yet available to public**
    - **Sold on a custom software development basis**
    - **Developed by for-profit company (B Corp)**
    - **Not currently open source (TBD)**

Item 2: Research and Recommendations on Data Storage Options that Support Indigenous Data Sovereignty

Recommended Data Storage Options That Support Indigenous Data Sovereignty

- **Baserow:** Database / app builder
  - **Strengths**
    - **Open source, no-code database platform**
    - **Self or cloud host**
    - **GDPR, HIPAA, SOC 2 compliant**
    - **Further customizable with plugins**
    - **They can host free version**
  - **Weaknesses**
    - **Built and maintained by for-profit company**
    - **Based in Germany**
    - **Paid version bills per user**
    - **Not built with IDSov in mind**

Item 2: Research and Recommendations on Data Storage Options that Support Indigenous Data Sovereignty

Recommended Data Storage Options That Support Indigenous Data Sovereignty

- **Proton:** Email / VPN / Cloud storage / Password manager / Calendar
  - **Strengths**
    - **Collection of privacy-oriented apps**
    - **Has free version without ads**
    - **Enterprise options available**
    - **Swiss jurisdiction**
  - **Weaknesses**
    - **For-profit, SaaS model (largest shareholder is its own non-profit)**
    - **Not meant for self-hosting**
    - **Subscription costs add up over time**
    - **Not built with IDSov in mind**
    - **Recent PR controversy**

Item 3: Indigenous Data Storage Governance Protocols for UWaterloo-based Data Storage Options

Protocol 1. Be FAIR and CARE

- Much work has already been done to establish how to implement FAIR principles, but significant further resources need to be dedicated to augmenting them with the CARE Principles as well.

- Implementation of FAIR is relatively easy considering it's furthering the status quo, however implementing CARE is a much more difficult social, political, and technical challenge involving all steps of the data lifecycle and must be prioritized and resourced accordingly.

Source: Research Data Alliance International Indigenous Data Sovereignty Interest Group, "CARE Principles for Indigenous Data Governance" (The Global Indigenous Data Alliance, September 2019), https://static1.squarespace.com/static/5d3799de845604000199cd24/t/6397b363b502ff481fce6baf/1670886246948/CARE%2BPrinciples_One%2BPagers%2BFINAL_Oct_17_2019.pdf.

Item 3: Indigenous Data Storage Governance Protocols for UWaterloo-based Data Storage Options

Protocol 2. Evaluate policies, practices, relationships, and data infrastructure for alignment to the CARE Principles via the CARE Data Maturity Model

- Currently in development by some of the key researchers in the Indigenous data sovereignty movement, the CARE Data Maturity Model will provide criteria and indicators that help ensure alignment with the CARE Principles.

- Some of these have already been drafted and can be further refined to account for First Nations, Inuit, and Métis research needs.

Source: Riley Taitingfong et al., "CARE Data Maturity Model," The Collaboratory for Indigenous Data Governance, May 8, 2024, https://indigenousdatalab.org/care-data-maturity-model/.

Item 3: Indigenous Data Storage Governance Protocols for UWaterloo-based Data Storage Options

Protocol 3. Maintain situational awareness of Indigenous relations and rights

- This should occur at all levels of research, both strategic and policy, but also day-to-day tasks and interactions.

- Researchers have a responsibility to conduct ethical research involving Indigenous data, but the University must provide adequate support to ensure this is actually possible throughout the entire data lifecycle from collection to archiving.

Source: Ruth Duerr et al., "Earth and Environmental Science Repository Activities in Support of the CARE Principles," January 17, 2024, https://zenodo.org/records/10521041.

Item 3: Indigenous Data Storage Governance Protocols for UWaterloo-based Data Storage Options

Protocol 4. Engage in consistent outreach activities with appropriate Indigenous Peoples

- Given the legacy of colonial research practices, Indigenous Peoples have little reason to trust that Universities will not only treat them with respect, but adequately safeguard and archive their sensitive data.

- Just as museums have much work to do to rematriate Indigenous ancestors, Universities must also build trust with relevant Indigenous communities to ensure that their best interests are prioritized and protected both now and in the future.

Source: Ruth Duerr et al., "Earth and Environmental Science Repository Activities in Support of the CARE Principles," January 17, 2024, https://zenodo.org/records/10521041.

Item 3: Indigenous Data Storage Governance Protocols for UWaterloo-based Data Storage Options

Protocol 5. Improve repository protocols so they are compatible with Indigenous data sovereignty principles

- Research repositories must consider Indigenous Peoples as key rights-holders and develop protocols, policies, procedures, and software design that ensures these rights are respected at all parts of the archiving workflow.

- This requires a significant investment in resources and direct involvement of appropriate Indigenous people in positions of power and influence to ensure compliance with relevant Indigenous law and protocols.

Source: Ruth Duerr et al., "Earth and Environmental Science Repository Activities in Support of the CARE Principles," January 17, 2024, https://zenodo.org/records/10521041.

Item 3: Indigenous Data Storage Governance Protocols for UWaterloo-based Data Storage Options

Protocol 6. Ensure technology development and maintenance is adequately prioritized and resourced

- The vast majority of software developers and digital infrastructure engineers have not been trained to account for Indigenous Peoples, their rights, needs, and experiences as users.
- The genealogy of their disciplines, as well as digitization as a whole, comes from developing technologies that form the logistical backbone of colonization itself. Further, Indigenous Peoples are a relatively small demographic due to colonization, so their needs are not often given the priority needed to achieve basic equity.
- Fortunately, what's good for Indigenous Peoples is often good for the wider public, so investing adequate resources in digital equity will lead to better software and infrastructure in the long run.

Source: Ruth Duerr et al., "Earth and Environmental Science Repository Activities in Support of the CARE Principles," January 17, 2024, https://zenodo.org/records/10521041.

Item 3: Indigenous Data Storage Governance Protocols for UWaterloo-based Data Storage Options

Protocol 7. Prioritize using software designed to respect Indigenous data sovereignty

- Ideal software should have customizable user roles and permissions that consider and allow flexibility in Ownership, Control, Access, and Possession (OCAP®).

- Multiple hosting options should include a multi-user database that allows for real-time collaboration that makes it easier to work and share data with Indigenous rights-holders. Industry-standard and beyond data security practices and features should include off-site backups and disaster recovery plans.

- Further plans should include budgets for ongoing software development that includes business or economic model considerations that affect software usability, reliability, and long-term archiving. Indigenous people should be involved in all aspects of software design to ensure distinction-based solutions to their specific needs

Recommended Software Features

- **Customizable user roles and permissions that consider and allow flexibility in Ownership, Control, Access, and Possession (OCAP®)**
- **Cloud or self-hosted multi-user database that allows for Real-time Collaboration**
- **Industry-standard and beyond data security practices and features including Off-Site Backups and Disaster Recovery**
- **Plan for ongoing software development that includes social, economic, and political considerations**
- **Indigenous people, and Peoples, involved in distinctions-based software design that meets their specific needs**

Derived and updated from: Jeff Doctor et al., *#DataBack: Asserting & Supporting Indigenous Data Sovereignty* (Animikii Inc., 2023), https://animikii.com/databack, and Ian Capstick et al., *Truth Before "Indigenous AI" Understanding and Resisting Generative AI* (Animikii, 2024), https://animikii.com/databack/understanding-and-resisting-generative-ai.

Further Resources

- **First Nations Information Governance Centre (FNIGC)**
  - fnigc.ca
- **BC First Nations Regional Information Governance Centre**
  - bcfndgi.com
- **Global Indigenous Data Alliance (GIDA)**
  - gida-global.org
- **United States Indigenous Data Sovereignty Network**
  - usindigenousdatanetwork.org
- **Te Mana Raraunga - Māori Data Sovereignty Network**
  - temanararaunga.maori.nz
- **Maiam nayri Wingara Indigenous Data Sovereignty Collective**
  - maiamnayriwingara.org

Remember: technologies can empower, but they can also entrap

Just because something is old, that doesn't mean it's bad.

Just because something is new, that doesn't mean it's good.

Just because we *can* do something, that doesn't mean we should.



Nya:węh!

Jeff Doctor

Impact Strategist

jeff.doctor@animikii.com

(877) 589-2899 x111

@jeff.doctor

linkedin.com/in/jeff-doctor

See our News River for Indigenous-focused tech articles every week:
animikii.com/newsriver

Open Discussion

- What challenges do you have managing Indigenous data?
- Are there things that we shouldn't digitize?
- Is there anything else you'd like to ask?

