

# A scan of Indigenous data in an institutional data repository: protocol

January 8, 2025

**Authors:** Alicia Zuniga, Senior Assistant Librarian, California State University - Sacramento (CSUS); Wanda Marsolek, Data Curation Librarian, University of Minnesota (UMN); and Shanda Hunt, Research Data Services Education & Outreach Librarian, University of Minnesota (UMN)

## Background

The [Data Repository for the University of Minnesota](#) (DRUM) is an open access (fully available to the public) institutional data repository. DRUM accepts Indigenous data into the repository utilizing informal practices (i.e., requests for data use agreement between researcher and Tribal Agency) and no formal policies. As a first step toward incorporating the [CARE Principles](#) into DRUM's workflow and policies, we proposed a project in partnership with [Data Services Continuing Professional Education](#) (DSCPE), a ten-week experience that connected Alicia Zuniga (CSUS) with Wanda Marsolek and Shanda Hunt (UMN).

The DSCPE project - which occurred September through November 2024 - included several vital activities performed by Zuniga which led to the following outputs: 1) list of search terms related to Indigenous data in Minnesota; 2) spreadsheet of those datasets that captures the level of Indigenous consent that was given to share the data publicly; 3) resource list of policies, practices, and documents related to Indigenous data; 4) recommendations for DRUM based on findings and the DSCPE experience; and 5) map of this process so that others, including DRUM staff, can implement and build on the important groundwork established by Zuniga.

This document is outputs #1 and #5 from above, authored by Zuniga and mentored by Marsolek and Hunt.

## Protocol

The first step was to create a list of potential search terms related to Indigenous data in Minnesota. In addition to collating a list of general Indigenous search terms (e.g., Indigenous, Tribe, etc.) and terms related to Minnesota (e.g., Grand Portage, Red Lake, etc.), Zuniga looked to Indigenous databases like [Native Health Database](#) and [Plateau Peoples' Web Portal](#) to explore additional metadata terms being used and added those that were relevant to the list. It's important to note that when Zuniga compiled keywords specific to Minnesota, she looked for Tribal terms that went beyond the 11 federally recognized Tribes, finding [this website](#) from the Mendota Mdewakanton Dakota Tribal Community useful. She then used these 61 search terms

to systematically search DRUM for Indigenous datasets [See Appendix A]. She then populated a spreadsheet of the found datasets, noting specific characteristics for each.

This protocol documents the process of populating the spreadsheet for the audit of Indigenous data located in DRUM. The headers below correspond to the column headers in the audit document, and should be completed in the order listed. The spreadsheet itself is not shared publicly as it will not be useful to anyone outside of DRUM staff.

## 1. Title

The title of the DRUM record.

## 2. DRUM link

The link to the DRUM record. Use the handle (located below the thumbnail photo) rather than the URL from the web browser. This ensures that DRUM staff can quickly navigate from the spreadsheet to other curatorial documents for further investigation as needed.

## 3. Description

A description of the data included in the DRUM record. This can usually be copied and pasted from either the Abstract or Description section in the DRUM record.

## 4. Keyword searched

The keyword searched in DRUM that produced the record being entered.

## 5. Indigenous data

Whether or not the data is about Indigenous people, land or resources. Context may be found in the description, funder information, and/or the supplemental files (specifically the ReadMe file). Sometimes the keyword will be mentioned in the record or the supplemental file but does not indicate that it is Indigenous data (e.g. You search “Chippewa” and the only mention of Chippewa in the record is a variable for the collection site like “Chippewa County” or “Chippewa Lake”).

Enter Y, N, or ? in this column.

- Y: Yes, data involves Indigenous people, land or resources.
- N: No, data does not involve Indigenous people, land or resources.
- ?: It is not clear whether or not the data involves Indigenous people, land or resources and requires further investigation or elevation.

This will be evaluated on a case by case basis. For example, look at these two examples of DRUM records whose subjects are both wild rice, an Indigenous resource:

- 1: [Whole Genome Assembly and Annotation of Northern Wild Rice \(\*Zizania palustris\* L.\), a North American Grain](#)
- 2: [Minnesota Pollution Control Agency Wild Rice Sulfate Standard Mesocosm Study: Water, Sediment, and Porewater "Synoptic" Sampling 2013 and 2015](#)

Dataset 1 is whole genome sequencing data of Northern Wild Rice. This dataset was surfaced by a keyword search for “tribe” in DRUM. In biology, a tribe is a taxonomic rank above genus, and below family and subfamily. The word “tribe” is used in reference to the wild rice tribe in this dataset, not in reference to a tribal nation. For this reason, you would mark N in this column since it is only data about the wild rice genome.

Dataset 2 is a sampling of 30 experimental, controlled outdoor environments for chemical parameters to understand the conditions in which wild rice thrive. At first glance this dataset appears to not have any tie to Indigenous people, land or resources since the experiments were done in controlled outdoor environments at the UMN Duluth campus, but the ReadMe file specifies that funding was received from “the Fond du Lac and Grand Portage Bands of Lake Superior Chippewa with Band funds and water quality funds provided by the Environmental Protection Agency.” The DRUM record funding section lists the [NSF grant EAR-0949962](#), which is publicly available to look at. The grant states “LacCore staff are engaged in education and outreach including an ongoing collaborative project aimed at study of the paleoenvironments of wild rice lakes in northern Minnesota with tribal college students and teachers from the Fond du Lac Band of Lake Superior Chippewa and Fond du Lac Tribal and Community College.” It clearly involves contributions from and collaboration with Indigenous people, so you would mark Y in this column.

## 6. If no Indigenous data, notes:

If the answer is no (N) in the previous column, note why the data is not considered Indigenous data (e.g. "Chippewa" listed as Lake name, geographic location of data collection).

If the answer is no (N) in the previous column, do not proceed to fill in the subsequent columns, unless there is something worth noting in the Notes column.

If the answer is unknown (?) in the previous column, make note of the circumstances that make it unknown in the Notes column.

## 7. Data permissions

The data permissions that are assigned to the dataset. This can be located in the License section of the DRUM record, and sometimes in the ReadMe file. Occasionally different answers will be listed in each of these locations for the same record. Note both

in this column and make note of the discrepancy in the Notes column. If there are no obvious data permissions stated, note this in this column.

## 8. Link to data permissions

If there are data permissions stated, copy and paste the URL where the permissions are stated. Sometimes this can be more than one location/URL, in which case separate the URLs with a semicolon (;). Most of the time they will be listed in the ReadMe file or the DRUM record.

## 9. Indigenous consent confirmed

Whether or not there is evidence that a person representing Tribal interests gave permission to share the data publicly. The “evidence” may be in a curator’s email, logged as part of the private curation record (i.e., the “backend” of the repository), or in a data use agreement between the University and a specific Tribe.

Enter Y, N, or ? in this column.

- Y: Yes, there is evidence of Indigenous consent for public data sharing.
- N: No, there is no evidence of Indigenous consent for public data sharing.
- ?: It is not clear whether there is evidence of Indigenous consent for public data sharing.

## 10. Notes

Any other useful information. Examples include:

- Missing ReadMe file
- Locked supplementary file
- Different data permissions listed in License section of DRUM record and in ReadMe file.
- Listed funder or data description suggests data is Indigenous, but there is no specific data permissions discussing data sovereignty.

As we move forward with this initiative and determine the best path forward, we may add variables to this spreadsheet. For example, through relationship building with Indigenous partners, we may be able to retroactively apply Indigenous metadata, such as TK Labels, to the DRUM record; apply reparative descriptions; or even contact Tribes represented in the data to ensure they’re aware of its existence – there is a precedent set, as all of these actions have been implemented by other repositories. We would want to document these actions in the audit spreadsheet as well.

## Appendix A. Keywords used to search DRUM for Indigenous data

Indigenous
Native American
American Indian
Indians, North American
Tribal
Tribe
Bois Forte Band of Chippewa
Red Lake Nation
White Earth Nation
Leech Lake Band of Ojibwe
Fond du Lac
Fond du Lac Band of Lake Superior Chippewa
Grand Portage
Grand Portage Band of Lake Superior Chippewa
Mille Lacs
Mille Lacs Band of Ojibwe
Mille Lacs Band of Chippewa
Upper Sioux Community
Lower Sioux Indian Community
Prairie Island Indian Community
Shakopee Mdewakanton Sioux Community
Chippewa
Dakota
Ojibwe
Mendota Mdewakanton
Ni-Mi-Win Ojibways
Rice Lake Band of Mississippi Ojibwe
Sandy Lake Band
Mille Lacs Band of Ojibwe
Snake and Knife Rivers Band
St. Croix Chippewa

Kah-Bay-Kah-Nong
Gabekanaang Anishinaabeg
Warroad Chippewa
Kettle River Band
ceremonial
clan
sacred
unceded land
ceded territory
ceded lands
reservations
Bureau of Indian Affairs
US Fish and Wildlife Services Tribal Wildlife Grant Program
Great Lakes Indian Fish and Wildlife Commission
Tribal Entity
EPA Region 5 Tribal Operations Caucus
United Tribes of Michigan
Minnesota Tribal Environmental Group
Tribal Nation
Native Tribes
Indian Child Welfare Act
Center for Regional and Tribal Child Welfare Studies
Native+
Indian Country
Minnesota Indian Affairs Council
Tribal Nations Education Committee
Minnesota Department of Health Native Health Directors
Traditional medicine
Shamanism
Minnesota Chippewa