



Evaluation of FAIR Principles in Indigenous Water- Climate-Environment (WCE) Data Repositories

RESEARCH PAPER

PARISA SARZAEIM

GRACE BULLTAIL

**Author affiliations can be found in the back matter of this article*

ubiquity press

ABSTRACT

Data-driven research on water resource planning and management of US reservation lands faces substantial challenges owing to a lack of accessible, updated, and integrated scientific data. Although several Indigenous studies have concentrated on water rights and policy, comparatively few have addressed the technical aspects of water resource planning and management of these lands. The goal of the study reported in this paper was to critically evaluate the discoverability of Water-Climate-Environment (WCE) data within Indigenous data repositories useful for data-driven water resource research. This study reviewed the implementation levels of Findability, Accessibility, Interoperability, and Reusability, known as FAIR guiding principles, in these scientific datasets for use by researchers. To accomplish this goal, we developed a systematic web-based search framework to identify active Indigenous-led WCE data repositories, evaluated them using FAIR principles, and offered recommendations for data utility enhancement. Our search yielded two publicly accessible data repositories: the Native Climate and Native Land Information System. The evaluation highlighted multiple key areas for the FAIRification of these data repositories, including assigning unique and persistent identifiers, enriching metadata, reformatting both data and metadata, registering data in different search engines, and specifying data use terms. These FAIRifications can enhance data discoverability and usability in data-driven water resource research for the interests of Native communities. Although this study is limited to Indigenous-led publicly accessible WCE data sources, we recommend conducting FAIR assessments of additional datasets, including those contained in larger databases or managed by non-tribal entities, as an important direction for future research.

CORRESPONDING AUTHOR:

Parisa Sarzaeim

Nelson Institute for
Environmental Studies,
University of Wisconsin-
Madison, Madison,
WI 53706, US

sarzaeim@wisc.edu

KEYWORDS:

Data-driven Water Resource
Research; FAIR Guiding
Principles; Indigenous Data
Repositories; Water-Climate-
Environment Data

TO CITE THIS ARTICLE:

Sarzaeim, P. and Bulltail, G.
(2025). Evaluation of FAIR
Principles in Indigenous
Water-Climate-Environment
(WCE) Data Repositories.
Data Science Journal, 24: 31,
pp. 1–16. DOI: <https://doi.org/10.5334/dsj-2025-031>

For thousands of years, Indigenous knowledge has contributed to sustainable water, land, and natural resource conservation on reservation lands (Hoagland, 2016). Modern environmental and ecological issues have become increasingly ubiquitous owing to global warming, hydroclimatological extremes, and population growth. As these issues become increasingly widespread, there is a growing need for data-driven research, in tandem with water rights and policy studies, to highlight the specific risks Native communities face. Gamble *et al.* (2016) and Fillmore and Singletary (2019) explained that underserved groups, including Indigenous communities, are most vulnerable to climatic changes. Consequently, there is an urgent need for tribes to prepare for and respond to climatic change impacts and risk exposure in their water systems, natural resources, and ecosystems (Cozzetto *et al.*, 2013).

Water resource management of reservation lands is a complex issue because of the involvement of different regulatory systems at the federal, state, and tribal levels (Royster, 2011). A simple review of the existing Indigenous scientific literature on water resources reveals a strong focus on topics related to Indigenous nations' sovereignty rights and self-determination in water resource protection (Bark *et al.*, 2012; Bulltail, 2024; Wilson *et al.*, 2021), land and natural resource policies (Lertzman and Vredenburg, 2005; Bulltail and Walter, 2020), environmental and ecological data control and sovereignty (Jennings *et al.*, 2023; Reid *et al.*, 2024), and tribal water rights and governance (Cohn *et al.*, 2019; Craig, 2022). Although numerous studies have focused on the policy, legal, and governance aspects of water resources, comparatively few have conducted data-driven and quantitative case studies of reservation lands.

Previous studies have underscored prominent water issues in tribal lands, including climate change and hydroclimatic extremes (Cozzetto *et al.*, 2013; McNeeley, 2017), water contamination (Chief, Meadow and Whyte, 2016), ecosystem and natural resource degradation (Doyle, Redsteer and Eggers, 2013; Gautam, Chief and Smith, 2013), and lack of access to safe and clean piped water (Murphy *et al.*, 2009; VanDerslice, 2011). These issues underscore the need for reliable databases that can support data-driven studies for developing constructive resolutions. However, numerous Indigenous and non-Indigenous researchers interested in addressing these issues have been constrained by the lack of accessible, high-quality, and consistent long- and short-term Water–Climate–Environment (WCE) databases (Chief *et al.*, 2016; Sarzaeim and Bulltail, 2023). To address this knowledge gap, we aimed to evaluate the availability, discoverability, and (re)usability of WCE data provided by Indigenous data repositories, with the goal of empowering scientists to address sustainability barriers in water resource management on US reservations.

Williamson *et al.* (2023) clarified that unavailable and inaccessible data coupled with poor data management are notable barriers to advancing environmental research on reservation lands. Data scarcity related to reservation lands may also represent one of the main barriers to the active engagement and involvement of Indigenous and local communities in climate research (David-Chavez and Gavin, 2018). To address these challenges, establishing efficient data system management that fosters the democratization of spatiotemporal WCE data collection, validation, documentation, acquisition, verification, and authentication is crucial (Sarzaeim *et al.*, 2023a), as this approach can enhance data accessibility and support data-driven research in Indigenous contexts. Although the rapid growth of digital libraries and web-based data repositories has led to the generation of numerous data packages (Yue *et al.*, 2012), accessible, well-documented, and well-structured datasets are needed to support efficient natural and water resource management on reservation lands. These data resources will equip researchers, scientists, stakeholders, and decision-makers with tools required to advance sustainable development goals.

Scientists have developed several data principles to enhance the discoverability and accessibility of data, thereby improving the infrastructure of scientific data management systems. Wilkinson *et al.* (2016) stated that while efficient data management is not a standalone goal, it is a leading requirement in digital democracy and knowledge equity. They introduced guiding principles that promote data Findability, Accessibility, Interoperability, and Reusability (FAIR) for both humans and machines to benefit science and researchers. The FAIR principles serve as a guide for data publishers to support data retrieval and reuse by other researchers. Additionally, in the Indigenous context, Carroll *et al.* (2021) developed the CARE data principles for Indigenous data

governance and data protection as a complement to the FAIR principles. The CARE principles support Collective benefits, Authority to control, Responsibility, and Ethics to foster the benefits of data use and protect data, with a focus on the well-being of Indigenous communities at every step of the scientific data lifecycle. The FAIR and CARE principles offer complementary frameworks that collectively maximize the value of data by ensuring the appropriate and responsible (re)use, particularly in the context of Indigenous data (Carroll et al., 2021). In this approach, development of FAIR Indigenous data is the primary step.

It appears that FAIRification practices have been examined in several scientific domains, including biodiversity science (Lannom, Koureas and Hardisty, 2020), geoscience (Kinkade and Shepherd, 2021), agricultural engineering (Ali and Dahlhus, 2022), and aquatic systems (Bayer et al., 2023). However, a FAIRification framework for Indigenous WCE data that can build a strong foundation for data-driven water research has not been evaluated. Thus, this study aims to assess the level of FAIR principles implemented in Indigenous-led WCE data repositories and propose a specific scientific and community-driven standard FAIRification framework. In this study, Indigenous-led WCE data repositories refer to data repositories that are generated, collected, resourced, or supported by Indigenous communities. To accomplish this goal, we defined three objectives as follows: (1) to explore and review the Indigenous-led WCE data repositories that are publicly available on the Web, (2) to evaluate the data repositories concerning each principle and sub-principle of FAIR, and (3) to recommend a list of FAIRification actions to support the reproducibility of Indigenous WCE data repositories identified in objective 1, for use in water resource research. Figure 1 illustrates the workflow of this study.

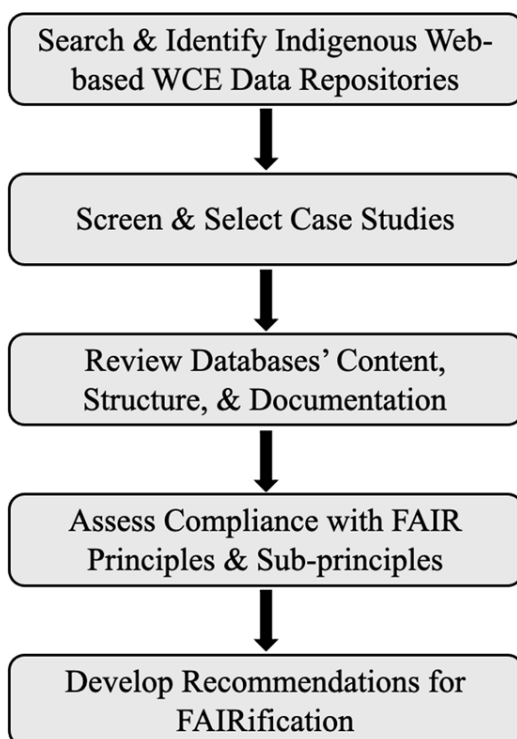


Figure 1 Overview of the key stages of this study.

FAIR DATA PRINCIPLES

Advancements in cyberinfrastructure technologies for measuring, collecting, recording, monitoring, storing, archiving, reporting, sharing, and visualizing have generated increasing data across disciplines that require an efficient, rapid-responsive, and user-friendly data management system (Jeong et al., 2019; Roche et al., 2015; Sun et al., 2019; Yang et al., 2010). The inclination toward interdisciplinary studies is another critical stimulus for increasing the effectiveness of multidimensional database management for a broader range of researchers across domains.

In response to large and complex volumes of data, scientific data in any discipline is recommended to follow FAIR principles. The FAIR principles support a basic standard of acceptable data management by data producers/publishers and benefit the entire academic/scholar community regarding knowledge, discovery, and science advancement by ensuring

data can be effectively found, accessed, easily integrated, and reused by other researchers (Go-fair.org). Each FAIR primary principle consists of sub-principles developed by Wilkinson *et al.* (2016) (Figure 2):

- (a) Findability requires that (meta)data (i.e., both data and metadata) be easily found by both humans and machines and consists of four sub-principles, including F1: (Meta)data are assigned a globally unique and persistent identifier, F2: Data are described with rich metadata, F3: (Meta)data clearly and explicitly include the identifier of the data they describe, and F4: (Meta)data are registered or indexed in searchable resources.
- (b) Accessibility refers to the requirement that once users find the desired data, they should be able to understand how to retrieve the data; this term consists of two sub-principles, including A1: (Meta)data are retrievable by their identifier using a standardized communication protocol that should be open, free, and universally implementable (A1.1) and should allow for authentication/authorization procedures if necessary (A1.2), and A2: Metadata should be accessible, even when the data are no longer available.
- (c) Interoperability requires that the (meta)data should be capable of interoperating with other applications/workflows for analysis, storage, and processing and consists of three sub-principles, including I1: (Meta)data should use a formal, accessible, shared, and broadly applicable language for knowledge representation, I2: (Meta)data should use vocabularies that follow FAIR principles, and I3: (Meta)data should include qualified references to other (meta)data.
- (d) Reusability refers to (meta)data that must be comprehensively described and documented to ensure that they can be replicated and/or combined in different settings; this function consists of one sub-principle, R1: (Meta)data should be richly described by a plurality of accurate and relevant attributes; in particular, (meta)data should be released with a clear and accessible data usage license (R1.1), (meta)data should be associated with detailed provenance (R1.2), and (meta)data should meet domain-relevant community standards (R1.3).

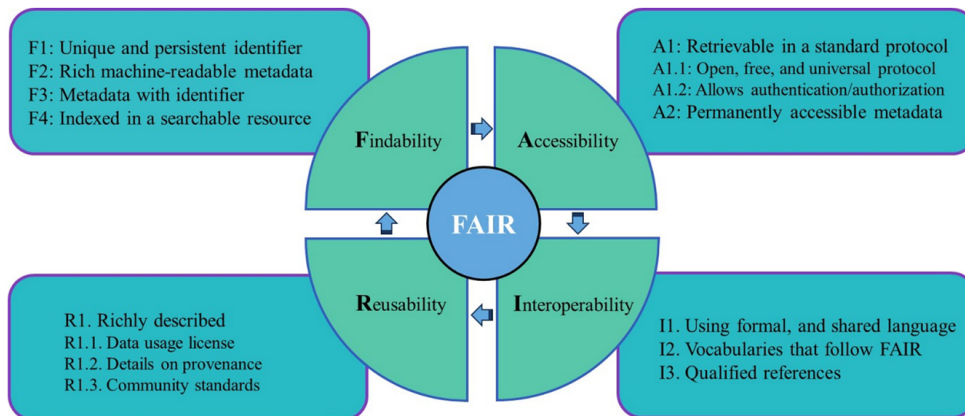


Figure 2 Overview of the FAIR principles and their corresponding sub-principles.

METHODS AND DATA

SEARCH DESIGN

The evaluation of the FAIR implementation level in Indigenous WCE data repositories began by identifying them on the Web (Figure 3). We systematically searched the data using the Google Search Engine in the “Incognito” mode, as recommended by Sarzaeim *et al.* (2023b), on Google Chrome to prevent interference from cookies and browsing history when finding relevant entries. We used the following strings on Google Search: (‘water’ OR ‘climate’ OR ‘environment’ OR ‘land’) AND (‘Indigenous’ OR ‘Tribal’ OR ‘Native’) AND (‘data’ OR ‘database’ OR ‘information’).

In our last search in February 2024, Google returned over two billion results. Because the search engines ranked the results based on relevance, we screened the first 20 entries for further evaluation. We sought possible databases consisting of digital, numerical, and geospatial data on water quality/quantity variables, climate/weather time series, and ecological and land-use

datasets located within the reservation lands. We limited our screening strategy to the first 20 entries after confirming that no further active WCE data beyond this point were found for inclusion as case studies in our research.

For the selection of the relevant web entries, we reviewed the “Home Page” or “About Us” pages of each result and filtered out the data not focused on the US. Next, we categorized the results into two groups: (1) Indigenous and (2) non-Indigenous data repositories. Indigenous data repositories specifically focus on data from tribal/reservation lands, whereas non-Indigenous data repositories are broader and may include Indigenous data as a part of a more general dataset. Given the study’s scope, we excluded the “non-Indigenous” data repositories. After applying the filters, two primary case studies remained for detailed examination: (a) Native Climate (NC) and (b) Native Land Information System (NLIS). Upon the initial review of the NC and NLIS data repositories, we identified overall information related to the scope of the data, including the data types, variables, and developers/sponsors, as outlined below:

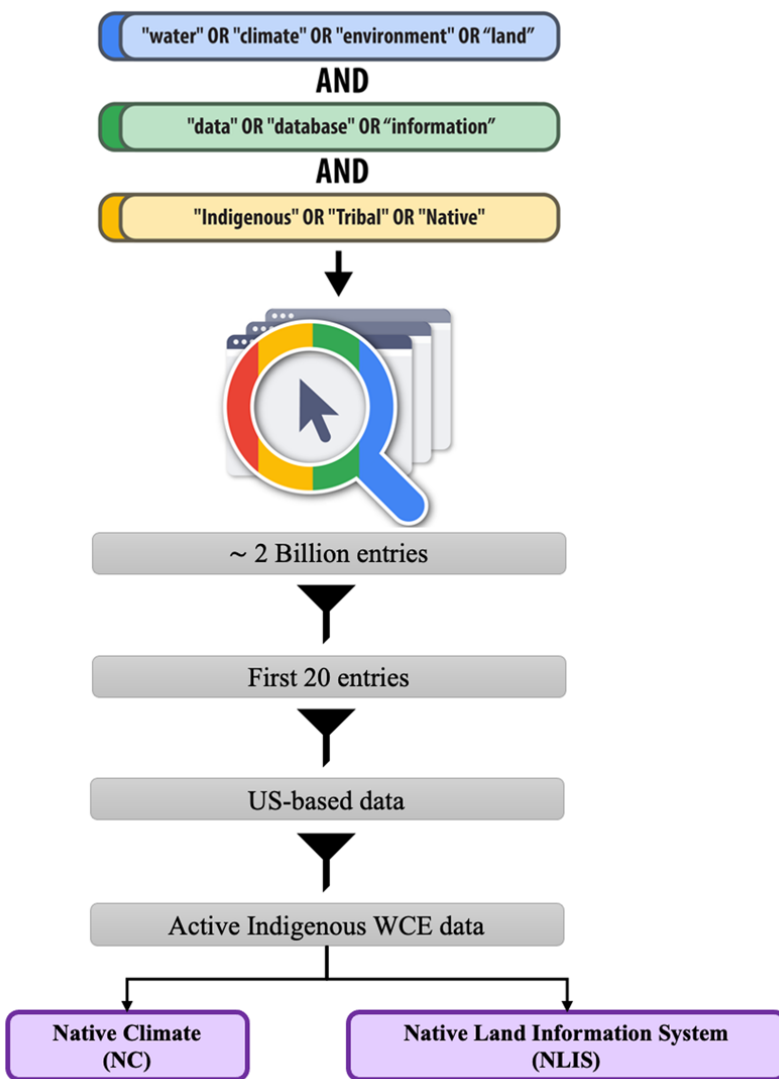


Figure 3 Systematic search flowchart for the active Indigenous Water-Climate-Environment (WCE) data repositories using the Google Search Engine on Google Chrome.

NATIVE CLIMATE (NC)

The Native Climate (NC) project was established to support climate change adaptation within tribal communities by rebuilding connections between Indigenous knowledge and Western scientific data (<https://native-climate.com>). Led by the Desert Research Institute, University of Nevada-Reno Tribal Extension, the Montana Climate Office at the University of Montana, the University of Arizona Tribal Extension, and the USDA Climate Hub, the project is funded by the United States Department of Agriculture-National Institute for Food and Agriculture (USDA-NIFA) and Native Waters on Arid Lands. The NC project provides access to multiple historical and

projected climate and weather variables, primarily focusing on temperature and precipitation time series for US reservation lands.

In the NC data repository, under the 'Climate Data/Projections' tab, the data folders directory can be accessed via the 'DOWNLOAD AND ACCESS THE DATA' button. The data were systematically structured and distributed across 633 zip folders, each corresponding to a distinct tribally controlled region developed by one of 35 Tribal Colleges/Universities (TCUs) responsible for dataset development, based on the last modified date in July 2024. Each .zip folder was then downloaded and extracted. Within each folder, the dataset was stored in an Excel file with the format of '.xlsx' and accompanied by a metadata file in ".pdf" format. The metadata provides details on the data attributes and includes various time series plots for climate projections under different representative concentration pathways for greenhouse gas emissions, covering the projection period from 2015 to 2100, as well as historical simulations from 1950 to 2014. [Table 1](#) presents the key features of the NC data repository.

NATIVE LAND INFORMATION SYSTEM (NLIS)

The Native Land Information System (NLIS) is facilitated by the Native Land Advocacy Project based in Fort Collins, Colorado, to support Indigenous communities' sovereign rights to protect their own data, land, and resources (<https://nativeland.info>). In the NLIS data repository, under the 'Data Tools/Dashboards' tab, a diverse collection of data is embedded in separate dashboards for Native demographics, agricultural systems, land use and cover, soil properties and capabilities, water resources, environmental risks, and energy and mineral data. Each dashboard presents the data in various formats, including maps, plots, tables, and images. Additionally, each dashboard is accompanied by descriptive metadata that provide contextual information, statistical summaries, and notes on the limitations or potential shortcomings of the presented data.

In this study, we selected five dashboards relevant to water resource research, including (a) the 'Surface Waters on Native Lands' dashboard, featuring four tabs of 'Map', 'Derived Mean Monthly Flow Graph', 'Measured Mean Annual Flow Table', and 'Lake Storage', (b) the 'Ground Water on US Native Lands' dashboard, featuring three tabs of 'Map', 'Aquifer Data Table', and 'Well Data Table', (c) the 'Environmental Risks for US Native Lands' dashboard, featuring four tabs of 'Environmental Maps', 'Timeline', 'Metadata', and 'Raw Data', (d) the 'Land Covers on Native Lands in the Coterminous US' dashboard, featuring three tabs of 'Map + Pie', 'Change Over Time', and 'Table Explorer', and (e) 'Precipitation Projections and Winter Trends on US Native Lands', featuring three tabs of 'Precipitation Projection Map', 'Snow-to-Precipitation Map', and 'Data Table'. This selection strategy was pivotal, as it covered a broad spectrum of climatological, hydrological, and environmental parameters and variables and was essential for comprehensively documenting water resource dynamics and interactions on Native lands. [Table 1](#) outlines the features of the NLIS dashboards used in this study.

FAIR ASSESSMENT

In the following steps, we evaluated the (meta)data of NC and NLIS for compliance with all of the FAIR principles and sub-principles following the FAIR metrics framework formulated by Wilkinson *et al.* ([Go-fair.org](https://go-fair.org); 2018). This qualitative framework is a questionnaire-based approach originally designed to evaluate the utility of the FAIR metrics themselves, rather than to evaluate different data resources. However, for evaluating the NC and NLIS (meta)data, we designed a questionnaire focused on answering this question, 'What must be provided?' (see Supplementary File 1: Appendix. FAIR Assessment Questionnaire). We employed this questionnaire to evaluate NC and NLIS by comparing the elements that were provided or missing in each data repository with the requirements for FAIR compliance. We formulated our responses, represented in the Results section, using the explanations, meanings, and examples for each FAIR principle and sub-principle provided on [Go-fair.org](https://go-fair.org). This approach allowed the authors of this work to perform a detailed assessment of each individual digital resource's compliance with FAIR principles/sub-principles to identify the strengths and limitations and suggest a specific set of FAIRifications to improve data discoverability, dataset formatting, metadata structure, access protocols, and (meta)data interpretability and comprehensibility.

DATA REPOSITORY	DATA TYPES	VARIABLES	DEVELOPERS	SPONSORS	WEBSITE
Native Climate (NC)	Agricultural climate variables	Annual precipitation, spring precipitation, summer precipitation, fall precipitation, winter precipitation, average precipitation on wet days, average precipitation on wet days (trace), maximum 3-day precipitation, number of wet days, number of wet days (trace), number of dry days, number of dry days (trace), day of first snow, average temperature, number of days ≥ 100 F, number of days with heat index hazard, growing degree days, frost free days, first day of growing season, last day of growing season, length of growing season, average surface wind speed, number of days with average wind speed ≥ 20 mph	Desert Research Institute, University of Nevada-Reno Tribal Extension, Montana Climate Office, University of Arizona Tribal Extension, and USDA Climate Hub	US Department of Agriculture-National Institute of Food and Agriculture, Native Waters on Arid Lands	native-climate.com
Native Land Information System (NLIS)	Surface water, groundwater, hydroclimatic variables, environmental risks and land cover variables	Mean monthly and annual surface flow, lake average storage and capacity, aquifer characteristics, wells characteristics and water table, Palmer Drought Severity Index (PDSI), Extreme Heat Days (EHD), precipitation, Wildfire Hazard Potential (WHP), land cover types and acres, projected mean annual precipitation, change of snow-to-precipitation ratio	Native Land Advocacy Project	Indian Land Tenure Foundation, Native American Agriculture Fund, Village Earth	nativeland.info

Table 1 Summary of the NC and NLIS data repositories' features.

RESULTS

FINDABILITY

F1: The NC and NLIS do not explicitly assign unique and persistent identifiers to each dataset they support.

An identifier is defined as 'a globally unique name for a resource that can be used to retrieve a description of the resource or to identify an entity like a file or dataset without dereferencing the identifier' ([Ananthakrishnan et al., 2020](#)), such as the Digital Object Identifier (DOI).

Given the rapid global increase in dataset searches, assigning a unique, persistent, resolvable, and thoroughly described identifier that supports both human and machine readability facilitates the discovery process, enables citations in scientific publications, and forms a fundamental part of an efficient data management system ([Ananthakrishnan et al., 2020](#); [Berber and Yahyapour, 2017](#)).

F2: The NC and NLIS contain descriptive metadata but not in a consistent, machine-readable format.

In the NC, the metadata file was released in both a '.pdf' format and within the first sheet of the '.xlsx' data file, labeled 'Metadata', if the user downloaded the data zip folder for a desired reservation. The NLIS data repository lacks a section explicitly labeled as 'metadata' in all the selected dashboards, except for the 'Environmental Risks for US Native Lands' dashboard, where a brief plain text description is provided.

Zhu and Cole ([2022](#)) argued that formats such as PDF are not designed for machine reading or software processing. Similarly, Awad *et al.* ([2020](#)) criticized Excel spreadsheets as a low-quality format for data, as they are proprietary, the sheets may be interconnected, and they may contain hidden contents and visual effects. Instead, non-proprietary machine-readable formats, such as JSON-DL, XML, and RDF, should be used, as they can be indexed by search engines more effectively and consequently improve data discoverability ([Nair and Jeevan, 2004](#); [Wilkinson et al., 2018](#)).

Regarding metadata richness, our observations showed that in NC, both the PDF and '.xlsx' metadata provided information on data provenance, data generation methods, temporal coverage, and contact email addresses. The PDF metadata were more explanatory and included more details on methodologies, data provenance, variable definitions and units, and data sponsors. Additionally, the PDF metadata included a wide range of variable time series plots.

In NLIS, the 'Metadata' tab of the 'Environmental Risks for US Native Lands' dashboard provides an overview of the methods, data provenance, spatiotemporal coverage, and the

main references for data development. In the other dashboards, text-based metadata are provided on the same data pages explaining the methods, importance of the datasets, and data limitations.

Metadata plays a crucial role in fostering user understanding of the data, including relevant processes, techniques, tools, and instruments used for data generation, projection, collection, validation, and curation, along with other data features like purpose and applications. Details about space and time coverage and resolution are particularly essential for spatiotemporal datasets that are employed in water resource research (Ge et al., 2019). Rich, highly explanatory, complete, and thoroughly documented metadata embedded in a machine-readable format substantially simplify data interpretations for users and help save time and data analysis resources.

F3: In both NC and NLIS, metadata do not include identifiers.

As described in the evaluation of the F1 sub-principle, the NC and NLIS data repositories do not assign unique and persistent identifiers to each dataset; consequently, the metadata do not include such identifiers. The metadata should explicitly contain a data identifier to improve the searchability, indexing, and discoverability of the data through search engines and to enable the data discovery from the metadata directly.

F4: The NC and NLIS data repositories are indexed in a searchable resource.

As confirmed by our discovery of these data repositories during the web-based search process outlined in the methodology, both NC and NLIS are indexed in the Google Search engine’s database (see Table 2).

DATA REPOSITORY	FINDABILITY
NC	<p>F1. <i>No unique and persistent identifiers have been explicitly assigned for each dataset.</i></p> <p>F2. <i>Rich metadata, no ideal machine-readable metadata format has been provided.</i></p> <p>F3. <i>Metadata does not include the identifier.</i></p> <p>F4. Indexed in the Google search engine’s database.</p>
NLIS	<p>F1. <i>No unique and persistent identifiers have been explicitly assigned.</i></p> <p>F2. <i>Rich metadata, no consistent machine-readable metadata format has been provided.</i></p> <p>F3. <i>Metadata does not include the identifier.</i></p> <p>F4. Indexed in the Google search engine’s database.</p>

Table 2 Summary of FAIR-F evaluation; italic text displays FAIR non-compliance.

ACCESSIBILITY

A1: The datasets in both the NC and NLIS data repositories are accessible through Hypertext Transfer Protocols (HTTPs).

HTTP(s) is a standard, open, free, and universally implementable protocol (A1.1) that also supports authentication/authorization when necessary (A1.2), depending on the access policies established by the data developers.

In our most recent access to the NC and NLIS websites, neither site required authentication/authorization, ensuring that the data are openly accessible to all users. However, according to the FAIR-A1.2 sub-principle, FAIR data do not necessarily refer to open-access data. Instead, data publishers can specify the conditions under which data are accessible based on user rights; in other words, highly protected data can nonetheless meet FAIR standards (Go-fair.org).

A2: In both NC and NLIS, there is no public evidence indicating the metadata preservation.

According to the FAIR-A2 sub-principle, metadata should be permanently accessible even when the data are no longer available. This requirement implies that the metadata should be persistent, accessible, and stored separately from the data (Go-fair.org). If the primary data are missing, lost, or inactive, a separate and persistent digital metadata resource allows users to understand that a specific dataset with a set of features was previously available for research purposes. Providing metadata in appropriate formats (F2), with persistent identifiers (F3) registered in searchable resources (F4), is essential for meeting the FAIR requirement of preserved metadata accessibility (see Table 3).

DATA REPOSITORY	ACCESSIBILITY
NC	<p>A1. Retrievable in an open, free access protocol, and no authentication/authorization is currently required.</p> <p>A2. <i>No guarantee for metadata preservation.</i></p>
NLIS	<p>A1. Retrievable in an open, free access protocol, and no authentication/authorization is currently required.</p> <p>A2. <i>No guarantee for metadata preservation.</i></p>

Table 3 Summary of FAIR-A evaluation; italic text displays FAIR non-compliance.

INTEROPERABILITY

I1: Both NC and NLIS partially use formal, accessible, shared, and broadly applicable language for knowledge representation.

Although both NC and NLIS use standard and widely accepted scientific languages in water science to describe the (meta)data, the NC metadata are released in ‘.pdf’ and ‘.xlsx’ formats, and the data are published in the ‘.xlsx’ format following a consistent (meta)data structure. In the NLIS, (meta)data are provided in plain text-based descriptors across all selected dashboards, where data are downloadable in the ‘.csv’ format in one single dashboard (i.e., ‘Precipitation Projections and Winter Trends on the US Native Lands’), following an inconsistent (meta)data structure.

To enhance interoperability, data should be released in a machine-friendly format, including but not limited to JSON-DL, XML, or RDF, as previously discussed in F2. For geospatial data, formats such as GeoJSON, Shapefile, and NetCDF are also recommended because of their effectiveness in map generation, geospatial analysis platforms, terrestrial modeling tools, and geospatial data archives (Butler *et al.*, 2016; Hassell *et al.*, 2017; Jhummarwala, Prajapati and Potdar, 2017). Additionally, the metadata can be structured using standards such as the Dublin Core Schema to ensure consistency and interoperability.

I2: Both NC and NLIS partially use FAIR vocabularies.

The absence of adequate documentation, clear definitions, and direct links for some variables, tribes, and references within NC and NLIS metadata and datasets hinders an efficient understanding, interpretations, and use of the data. Using formal FAIR languages in (meta)data facilitates the effective discovery, interpretation, validation, and processing information by both humans and machines. This sub-principle ensures that each concept, term, and vocabulary used in the (meta)data is effectively documented, findable through unique and persistent identifiers (F1), accessible by standard communication protocols (A1), and interoperable by employing a widely applicable language for knowledge communication (I1) (Go-fair.org; Wilkinson *et al.*, 2018).

I3: Both the NC and NLIS provide links to the relevant digital resources.

In both data repositories, resolvable links to the original data products on which they are built, reference articles describing the data development methodology, and relevant digital resources, portals, and data partners are properly cited and acknowledged.

Enriching the (meta)data by referencing the resources and articles through which the data is developed is essential for understanding the data development process, understanding data origins, and ensuring data validity (dos Santos, Peroni and Mucheroni, 2022) (see Table 4).

DATA REPOSITORY	INTEROPERABILITY
NC	<p>I1. <i>(Meta)data are provided in a partially formal, accessible, shared, and broadly applicable language formats.</i></p> <p>I2. <i>(Meta)data partially use FAIR vocabularies.</i></p> <p>I3. The metadata provides the key references.</p>
NLIS	<p>I1. <i>(Meta)data are provided in a partially formal, accessible, shared, and broadly applicable language formats.</i></p> <p>I2. <i>(Meta)data partially use FAIR vocabularies.</i></p> <p>I3. The metadata provides the key references.</p>

Table 4 Summary of FAIR-I evaluation; italic text displays FAIR non-compliance.

REUSABILITY

R1: In both NC and NLIS, metadata do not contain the data usage license, provide details on data provenance, and partially align with community standards.

NC provides important data descriptions, including data provenance, developers/sponsors, development methods, spatiotemporal coverage, variable definitions and units, and contact information for any inquiries. Similarly, the NLIS offers a necessary data description, including data provenance, developers/sponsors, development methods, spatiotemporal coverage, data importance and limitations, and a contact email address for any inquiries.

However, both NC and NLIS have considerable potential to improve (meta)data reusability for research applications. These improvements include specifying a clear data usage license for which the data can be used (R1.1), providing data provenance in both human- and machine-readable formats (R1.2), and ensuring that the (meta)data aligns with the scientific community standards (R1.3). For the latter task, any specific standards established by Indigenous communities also should be considered. A consistent (meta)data structure in format, file naming, variable definitions, and data applications and limitations is also highly useful in enhancing data reusability (see Table 5).

DATA REPOSITORY	REUSABILITY
NC	<i>R1. No clear data usage license is provided, details on data provenance is provided, and partially follows the community standards.</i>
NLIS	<i>R1. No clear data usage license is provided, details on data provenance is provided, and partially follows the community standards.</i>

Table 5 Summary of FAIR-R evaluation; italic text displays FAIR non-compliance.

Although both the NC and NLIS repositories offer valuable Indigenous WCE datasets, a specific set of recommendations has been developed as a part of FAIRification to address the identified gaps and non-compliances with FAIR principles. These recommendations provide a first-level set of actions for data managers/developers to implement:

(a) NC

- Assigning unique and persistent identifiers for each dataset and including them in the metadata.
- Ensuring that the preserved metadata remain as comprehensive as possible, including but not limited to data accuracy, data limitation, spatiotemporal coverage updates, resolution, and identifiers in both human- and machine-readable formats.
- Providing the datasets in appropriate human- and machine-readable formats.
- Registering the (meta)data in various searchable resources for broader discoverability.
- Specifying a clear data usage license.
- Maintaining (meta)data with both WCE data and Indigenous community standards.

(b) NLIS

- Assigning unique and persistent identifiers for each dataset and including them in the metadata.
- Ensuring that the preserved metadata remain as comprehensive as possible, including but not limited to data accuracy, variable definitions and units, spatiotemporal coverage updates, resolution, and identifiers in a dedicated persistent metadata digital resource in both human- and machine-readable formats in all dashboards.
- Providing the datasets in consistent, both human- and machine-readable formats in all dashboards.
- Registering the (meta)data in various searchable resources for broader discoverability.
- Specifying a clear data usage license.
- Maintaining (meta)data with both WCE data and Indigenous community standards, with a consistent (meta)data structure.

The FAIRification of Indigenous data repositories is the first step that Carroll *et al.* (2021) recommended for operationalizing FAIR-CARE principles in the entire data lifecycle for already existing and newly created data in both exclusively Indigenous and a mix of Indigenous and non-Indigenous datasets. Thus, the development of a data management plan should be encouraged through the proposal phase of Indigenous database development projects. In turn, funding agencies should consider and provide the financial support to guarantee the implementation of FAIR principles. Funding sources typically require a data management plan in research proposals but may not weigh evaluation scoring rubrics to recognize the FAIRification of the data products. Additionally, proposal budget justifications can include resources to support professional development opportunities and training for projects and community members to manage data. If the management plans implement these principles in the WCE data, evaluators should recognize this practice in the evaluation scoring rubric. The implementation of sound data management plans will contribute to the overall research project quality. This implementation plan should be considered as an ongoing effort in each stage of the data lifecycle, including design, infrastructure, collection, storage, and reuse for sustained FAIR compliance (Jennings *et al.*, 2025). Both NC and NLIS are acknowledged for providing a wide range of WCE datasets across reservations in the US, and after implementing the FAIRification processes recommended above, subsequent datasets will provide a step forward in enhancing visibility and reusability for data-driven water resource research on reservations.

Notably, FAIR data principles are not designed to examine data quality or reliability; instead, the evaluation of these aspects is the responsibility of the data user and is based on a specific data application (Go-fair.org). In this research, we aimed to highlight the importance of implementing FAIRification for Indigenous WCE data publishers and users through emphasizing the need to enhance data discoverability and usability; such an approach would facilitate the broader use of data across a wide range of data-driven topics in water resource research, including, but not limited to, surface and groundwater modeling and management, water balance analysis, flood control and risk assessment, water quality modeling, supply and demand management and operation, ecological modeling and assessment, and quantifying the effects of climate change on water resources.

Despite there being a remarkable potential for technical studies and data-driven research to address the water-related problems of Native communities in several cases, data scarcity poses barriers for scholars, water managers, policymakers, and stakeholders (Young, Colby and Thompson, 2019). Only a limited number of data-driven case studies have been conducted; those that do exist highlight the critical role of valid data generation and application in research, with a focus on enhancing the well-being of Indigenous peoples. Examples include analyzing the geospatial distribution of groundwater contaminants across the Navajo Nation (Hoover *et al.*, 2017), health risk assessment and water treatment performance evaluation in the Crow Reservation (Eggers *et al.*, 2018), surface water quality sampling across the Yukon River Basin conducted in collaboration with the Alaska Native Tribes and Canadian First Nations (Wilson *et al.*, 2018), groundwater quality assessments in the Omaha and Santee Sioux Reservations (McGinnis and Davis, 2001), and groundwater sampling in Alaska (Rowles III *et al.*, 2020). These data-driven studies can contribute to informed decision-making regarding water and natural resource management in tribal communities.

The ‘availability’ of data falls out of the FAIR data principles’ scope, and in the case of Indigenous data, decisions regarding data publication and accessibility are in the scope of the sovereign right of the tribes, who determine whether and how their data should be shared. The core of CARE data principles is founded on data protection practices to preserve Indigenous data and ensure that data usage aligns with the interests of Native communities. However, publishing, sharing, and preserving non-sensitive Indigenous WCE data that follows FAIR principles in research data repositories, in tandem with tribal agreements, can remarkably advance data-driven water resource research focused on Native communities.

This study focused on Indigenous-led WCE data repositories. Data repositories managed by federal and state organizations, which record, document, and publish Indigenous data as part of larger databases, were not included in this analysis.

CONCLUSIONS

The lack of FAIR WCE data is a notable barrier in conducting data-driven water resource research relevant to tribal communities. The FAIRification of Indigenous data is an initial step toward operationalizing the integrated FAIR-CARE principles that support the development, sharing, and protection of these data. Our assessment of publicly available Indigenous WCE databases such as NC and NLIS outlines key areas for improving the data's findability, accessibility, interoperability, and reusability. The NC and NLIS data repositories provide valuable spatiotemporal hydrological, climatic, and environmental datasets across US reservations in various data structures and formats. However, these data repositories do not fully align with the FAIR guiding principles, as outlined below:

- The absence of unique and persistent identifiers hinders data discovery and tracing.
- Inconsistent and nonmachine-readable (meta)data formats make working with these datasets challenging and time-consuming.
- Metadata could be considerably more explanatory regarding data accuracy and limitations, variable definitions, updated information on spatiotemporal and resolution coverage, and data usage terms and conditions.
- Registering datasets published by the NC and NLIS on multiple data-search platforms is strongly recommended to increase their visibility to researchers.
- Specifying a clear data usage license and permissions can substantially enhance research efficiency, facilitate data reuse, promote transparency, and protect the rights of both data developers and users.

The FAIR Indigenous WCE data can considerably enhance data discoverability and usability for both Indigenous and non-Indigenous scholars, thereby broadening the scope and impact of data-driven water resource research relevant to tribal communities. Because this research was limited to the web-based, Indigenous-led WCE data repositories, an evaluation of FAIR principles in Indigenous WCE datasets documented and released by federal and state institutions, as well as non-web-based data, is recommended for future work. Additionally, FAIRness requirements may be different in different communities due to the community-specific standards. Therefore, evaluating the FAIR data metrics in the context of the Indigenous community could be a valuable topic for further research.

ADDITIONAL FILES

The additional files for this article can be found as follows:

- **Appendix 1.** FAIR Assessment Questionnaire. DOI: <https://doi.org/10.5334/dsj-2025-031.s1>
- **Appendix 2.** Abbreviations. DOI: <https://doi.org/10.5334/dsj-2025-031.s2>

ACKNOWLEDGEMENTS

We thank the Native Climate and Native Land Information System for providing the preliminary data used in this study. We also thank Cathy Middlecamp for her insightful comments on improving our manuscript. AI has been used in some cases for grammar checks.

FUNDING INFORMATION

We acknowledge support from the Indigenous Environment & Science Working Group, Office of Vice Chancellor for Research and Graduate Education & Nelson Institute, University of Wisconsin-Madison, and the Center for Water Policy, School of Freshwater Science, University of Wisconsin-Milwaukee.

COMPETING INTERESTS

The authors have no competing interests to declare.

P.S. conceptualized the study, carried out the implementation, and prepared the manuscript, and G.B. conceptualized the study, supervised the research, and reviewed and prepared the manuscript.

AUTHOR INFORMATION

P.S. is a Research Associate at the Nelson Institute for Environmental Studies at the University of Wisconsin-Madison. Her main research focuses on water resource management, database management, and data science.

G.B. is a member of the Crow Tribe and an Assistant Professor at the Nelson Institute for Environmental Studies at the University of Wisconsin-Madison. Her main research focuses on developing resource sovereignty frameworks for Indigenous communities.

AUTHOR AFFILIATIONS

Parisa Sarzaeim  orcid.org/0000-0003-3661-2423

Nelson Institute for Environmental Studies, University of Wisconsin-Madison, Madison, WI 53706, US

Grace Bulltail  orcid.org/0000-0003-1306-3302

Nelson Institute for Environmental Studies, University of Wisconsin-Madison, Madison, WI 53706, US

REFERENCES

- Ali, B.** and **Dahkhaus, P.** (2022) 'The role of FAIR data towards sustainable agricultural performance: A systematic literature review', *Agriculture*, 12(2), p. 309. Available at: <https://doi.org/10.3390/agriculture12020309>
- Ananthakrishnan, R., Chard, K., D'Arcy, M., Foster, I., Kesselman, C., McCollam, B., Pruyn, J., Rocca-Serra, P., Schuler, R. and Wagner, R.** (2020) 'An open ecosystem for pervasive use of persistent identifiers', *Practice and Experience in Advanced Research Computing 2020: Catch the Wave*, PEARC '20, 2020. Portland, OR, USA, 26–30 July. New York, NY: Association for Computing Machinery, pp. 99–105. Available at: <https://doi.org/10.1145/3311790.3396660>
- Awad, A., Elgohary, R., Moawad, I. and Roushdy, M.** (2020) 'Metadata extraction for low-quality semi-structured spreadsheets', in A.E. Hassanien, A. Azar, T. Gaber, D. Oliva and F. Tolba (eds.) *International conference on artificial intelligence and computer*. Cham: Springer, pp. 448–457. Available at: https://doi.org/10.1007/978-3-030-44289-7_42
- Bark, R.H., Garrick, D.E., Robinson, C.J. and Jackson, S.** (2012) 'Adaptive basin governance and the prospects for meeting Indigenous water claims', *Environmental Science & Policy*, 19–20, pp. 169–177. Available at: <https://doi.org/10.1016/j.envsci.2012.03.005>
- Bayer, J.M., Scully, R.A., Diabola, E.K., Courtwright, J.L., Hirsch, C.L., Hockman-Wert, D., Miller, S.W., Roper, B.B., Saunders, W.C. and Snyder, M.N.** (2023) 'Sharing FAIR monitoring program data improves discoverability and reuse', *Environmental Monitoring and Assessment*, 195, p. 1141. Available at: <https://doi.org/10.1007/s10661-023-11788-4>
- Berber, F. and Yahyapour, R.** (2017) 'A high-performance persistent identifier management protocol', *International Conference on Networking, Architecture, and Storage (NAS), 2017*. Shenzhen, China, 7–9 August. Red Hook, NY: Curran, pp. 1–10. IEEE. Available at: <https://doi.org/10.1109/NAS.2017.8026839>
- Bulltail, G.** (2024) 'Water resource management and land tenure on allotted and tribal lands', *Environmental Research Letters*, 19. Available at: <https://doi.org/10.1088/1748-9326/ad5574>
- Bulltail, G. and Walter, M.T.** (2020) 'Impacts of coal resource development on surface water quality in a multi-jurisdictional watershed in the Western United States', *Journal of Contemporary Water Research & Education*, 169(1), pp. 79–91. Available at: <https://doi.org/10.1111/j.1936-704X.2020.03333.x>
- Butler, H., Daly, M., Doyle, A., Gillies, S., Hagen, S. and Schaub, T.** (2016) *The GeoJSON format*. RFC Editor. Available at: <https://doi.org/10.17487/RFC7946>
- Carroll, S.R., Herczog, E., Hudson, M., Russell, K. and Stall, S.** (2021) 'Operationalizing the CARE and FAIR Principles for Indigenous data futures', *Scientific Data*, 8. Available at: <https://doi.org/10.1038/s41597-021-00892-0>
- Chief, K., Meadow, A. and Whyte, K.** (2016) 'Engaging southwestern tribes in sustainable water resources topics and management', *Water*, 8(8). Available at: <https://doi.org/10.3390/w8080350>
- Cohn, T.C., Berry, K., Whyte, K.P. and Norman, E.** (2019) 'Spatio-temporality and tribal water quality governance in the United States', *Water*, 11(1). Available at: <https://doi.org/10.3390/w11010099>

- Cozzetto, K., Chief, K., Dittmer, K., Brubaker, K., Gough, R., Souza, K., Ettawageshik, F., Wotkyns, S., Opitz-Stapleton, S., Duren, S. and Chavan, P.** (2013) 'Climate change impacts on the water resources of American Indians and Alaska Natives in the U.S.', in J.K. Maldonado, B. Colombi and R. Pandya (eds.) *Climate change and Indigenous peoples in the United States*. Cham: Springer, pp. 61–76. Available at: https://doi.org/10.1007/978-3-319-05266-3_6
- Craig, R.K.** (2022) 'Tribal water rights and tribal health: The Klamath Tribes and the Navajo Nation during the COVID-19 pandemic', *St. Louis University Journal of Health Law & Policy*, 16(1). Available at: <https://scholarship.law.slu.edu/jhlp/vol16/iss1/5/>
- David-Chavez, D.M. and Gavin, M.C.** (2018) 'A global assessment of Indigenous community engagement in climate research', *Environmental Research Letters*, 13. Available at: <https://doi.org/10.1088/1748-9326/aaf300>
- dos Santos, E.A., Peroni, S. and Mucheroni, M.L.** (2022) 'The way we cite: Common metadata used across disciplines for defining bibliographic references', *26th International Conference on Theory and Practice of Digital Libraries, TPDL, 2022*. Padua, Italy, 20–23 September. Cham: Springer, pp. 120–132. Available at: https://doi.org/10.1007/978-3-031-16802-4_10
- Doyle, J.T., Redsteer, M.H. and Eggers, M.J.** (2013) 'Exploring effects of climate change on Northern Plains American Indian health', *Climatic Changes*, 120, pp. 135–147. Available at: https://doi.org/10.1007/978-3-319-05266-3_11
- Eggers, M.J., Doyle, J.T., Lefthand, M.J., Young, S.L., Moore-Nall, A.L., Kindness, L., Other Medicine, R., Ford, T.E., Dietrich, E., Parker, A.E., et al.** (2018) 'Community Engaged Cumulative Risk Assessment of Exposure to Inorganic Well Water Contaminants, Crow Reservation, Montana', *International Journal of Environmental Research and Public Health*, 15(1), pp. 76. <https://doi.org/10.3390/ijerph15010076>
- Fillmore, H.M. and Singletary, L.** (2019) 'Climate data and information needs of indigenous communities on reservation lands: insights from stakeholders in the Southwestern United States', *Climatic Change*, 169, p. 37. Available at: <https://doi.org/10.1007/s10584-021-03285-9>
- Gamble, J.L., Balbus, J., Berger, M., Bouye, K., Campbell, V., Chief, K., Conlon, K., Crimmins, A., Flanagan, B., Gonzalez-Maddux, C., Hallisey, E., Hutchins, S., Jantarasami, L., Khoury, S., Kiefer, M., Kolling, J., Lynn, K., Manangan, A., McDonald, M., Morello-Frosch, R., Redsteer, M.H., Sheffield, P., Thigpen Tart, K., Watson, J., Whyte, K.P. and Wolkin, A.F.** (2016) 'Populations of concern', in A. Crimmins, J. Balbus, J.L. Gamble, C.B. Beard, J.E. Bell, D. Dodgen, R.J. Eisen, N. Fann, M.D. Hawkins, S.C. Herring, L. Jantarasami, D.M. Mills, S. Saha, M.C. Sarofim, J. Trtanj and L. Ziska (eds.) *The impacts of climate change on human health in the United States*. New York, NY: Skyhorse, pp. 247–286. Available at: <https://doi.org/10.7930/J0Q81B0T>
- Gautam, M.R., Chief, K. and Smith, W.J., Jr.** (2013) 'Climate change in arid lands and Native American socioeconomic vulnerability: The case of the Pyramid Lake Paiute Tribe', in J.K. Maldonado, B. Colombi and R. Pandya (eds.) *Climate change and Indigenous peoples in the United States*. Cham: Springer, pp. 77–91. Available at: https://doi.org/10.1007/978-3-319-05266-3_7
- Ge, Y., Jin, Y., Stein, A., Chen, Y., Wang, J., Wang, J., Cheng, Q., Bai, H., Liu, M. and Atkinson, P.M.** (2019) 'Principles and methods of scaling geospatial Earth science data', *Earth-Science Reviews*, 197. Available at: <https://doi.org/10.1016/j.earscirev.2019.102897>
- GO FAIR.** (n.d.) *FAIR principles*. Available at: <https://www.go-fair.org/fair-principles/> (Accessed: 14 October 2024).
- Hassell, D., Gregory, J., Blower, J., Lawrence, B.N. and Taylor, K.E.** (2017) 'A data model of the Climate and Forecast metadata conventions (CF-1.6) with a software implementation (cf-python v2. 1)', *Geoscientific Model Development*, 10(12), pp. 4619–4646. Available at: <https://doi.org/10.5194/gmd-10-4619-2017>
- Hoagland, S.J.** (2016) 'Integrating traditional ecological knowledge with western science for optimal natural resource management', *IK: Other Ways of Knowing*, 3(1), pp. 1–15. Available at: <https://doi.org/10.18113/P8ik359744>
- Hoover, J., Gonzales, M., Shuey, C., Barney, Y. and Lewis, J.** (2017) 'Elevated arsenic and uranium concentrations in unregulated water sources on the Navajo Nation, USA', *Exposure and Health*, 9(2), pp. 113–124. Available at: <https://doi.org/10.1007/s12403-016-0226-6>
- Jeong, S., Hou, R., Lynch, J.P., Sohn, H. and Law, K.H.** (2019) 'A scalable cloud-based cyberinfrastructure platform for bridge monitoring', *Structure and Infrastructure Engineering*, 15(1), pp. 82–102. Available at: <https://doi.org/10.1080/15732479.2018.1500617>
- Jennings, L., Anderson, T., Martinez, A., Sterling, R., David-Chavez, D., Garba, I., Hudson, M., Garrison, N.A. and Russo Carroll, S.** (2023) 'Applying the "CARE Principles for Indigenous Data Governance" to ecology and biodiversity research', *Nature Ecology & Evolution*, 7, pp. 1547–1551. Available at: <https://doi.org/10.1038/s41559-023-02161-2>
- Jennings, L., Jones, K., Taitingfong, R., Martinez, A., David-Chavez, D., Anolani Alegado, R., Tofighi-Niaki, A., Maldonado, J., Thomas, B., Dye, D., Weber, J., Spellman, K.V., Ketchum, S., Duerr, R., Johnson, N., Balch, J. and Russo Carroll, S.** (2025) 'Governance of Indigenous data in open earth systems science', *Nature Communications*, 16. Available at: <https://doi.org/10.1038/s41467-024-53480-2>

- Jhummarwala, A., Prajapati, A. and Potdar, M.B.** (2017) 'GeoDigViz: A spatio-temporal model for massive analysis of shapefiles.' *2017 International Conference on Computing, Communication and Automation, ICCCA, 2017*. Greater Noida, India, 5–6 May. Piscataway, NJ: The Institute of Electrical and Electronics Engineers, Inc., pp. 1240–1245. Available at: <https://doi.org/10.1109/CCAA.2017.8229989>
- Kinkade, D. and Shepherd, A.** (2021) 'Geoscience data publication: Practices and perspectives on enabling the FAIR guiding principles', *Geoscience Data Journal*, 9(1), pp. 177–186. Available at: <https://doi.org/10.1002/gdj3.120>
- Lannom, L., Koureas, D. and Hardisty, A.R.** (2020) 'FAIR data and services in biodiversity science and geoscience', *Data Intelligence*, 2(1–2), pp. 122–130. Available at: https://doi.org/10.1162/dint_a_00034
- Lertzman, D.A. and Vredenburg, H.** (2005) 'Indigenous peoples, resource extraction and sustainable development: An ethical approach', *Journal of Business Ethics*, 56, pp. 239–254. Available at: <https://doi.org/10.1007/s10551-004-3528-8>
- McGinnis, S. and Davis, R.** (2001) 'Domestic well water quality within tribal lands of eastern Nebraska', *Environmental Geology*, 41, pp. 321–329. Available at: <https://doi.org/10.1007/s002540100389>
- McNeeley, S.M.** (2017) 'Sustainable climate change adaptation in Indian country', *Weather, Climate, and Society*, 9(3), pp. 393–404. Available at: <https://doi.org/10.1175/WCAS-D-16-0121.1>
- Murphy, M., Lewis, L., Sabogal, R.I. and Bell, C.** (2009) 'Survey of unregulated drinking water sources on the Navajo Nation', *137th Annual Meeting of the American Public Health Association, APHA, 2009*. Philadelphia, PA, USA, 7–11 November. Available at: <http://apha.confex.com/apha/137am/webprogram/Paper208881.html>.
- Nair, S.S. and Jeevan, V.K.J.** (2004) 'A brief overview of metadata formats', *DESIDOC Bulletin of Library & Information Technology*, 24(4), pp 3–11. Available at: <https://doi.org/10.14429/dbit.24.4.3629>
- Native Climate.** *Climate data*. Available at: <https://native-climate.com/about/> (Accessed: Feb 2024).
- Native Land Information System.** *Dashboards*. Available at: <https://nativeland.info/> (Accessed: Feb 2024).
- Reid, A.J., McGregor, D.A., Menzies, A.K., Eckert, L.E., Febria, C.M. and Popp, J.N.** (2024) 'Ecological research "in a good way" means ethical and equitable relationships with Indigenous peoples and lands', *Nature Ecology & Evolution*, 8, pp. 595–598. Available at: <https://doi.org/10.1038/s41559-023-02309-0>
- Roche, D.G., Kruuk, L.E.B., Lanfear, R. and Sandra, A.B.** (2015) 'Public data archiving in ecology and evolution: how well are we doing?', *PLoS Biology*, 13(11). Available at: <https://doi.org/10.1371/journal.pbio.1002295>
- Rowles, III, L.S., Hossain, A.I., Aggarwal, S., Kirisits, M.J. and Saleh, N.B.** (2020) 'Water quality and associated microbial ecology in selected Alaska Native communities: Challenges in off-the-grid water supplies', *Science of the Total Environment*, 711. Available at: <https://doi.org/10.1016/j.scitotenv.2019.134450>
- Royster, J.V.** (2011) 'Conjunctive management of reservation water resources: Legal issues facing Indian tribes', *Idaho Law Review*, 47(2), pp. 255–272.
- Sarzaeim, P. and Bulltail, G.** (2023) *FAIR and CARE data principles assessment for sustainable water resources management on reservation lands*. *CUAHSI Hydroinformatics Blog*. (Accessed: 14 October 2024).
- Sarzaeim, P., Mahmoud, Q.H., Azim, A., Bauer, G. and Bowles, I.** (2023b) 'A systematic review of using machine learning and natural language processing in smart policing', *Computers*, 12(12). Available at: <https://doi.org/10.3390/computers12120255>
- Sarzaeim, P., Munoz-Arriola, F., Jarquin, D., Aslam, H. and De Leon Gatti, N.** (2023a) 'CLIM4OMICS: A geospatially comprehensive climate and multi-OMICS database for maize phenotype predictability in the United States and Canada', *Earth System Science Data*, 15(9), pp. 3963–3990. Available at: <https://doi.org/10.5194/essd-15-3963-2023>
- Sun, Z., Di, L., Fang, H., Guo, L., Yu, E., Tang, J., Zhao, H., Gaigalas, J., Zhang, C., Lin, L., Yu, Z., Zhong, S., Wang, X., Tan, X., Jiang, L., Chen, Z., Xu, Z. and Sun, J.** (2019) 'Advanced cyberinfrastructure for agricultural drought monitoring', *8th International Conference on Agro-Geoinformatics, 2019*. Istanbul, Turkey, 16–19 July. Available at: <https://doi.org/10.1109/Agro-Geoinformatics.2019.8820694>
- VanDerslice, J.** (2011) 'Drinking water infrastructure and environmental disparities: Evidence and methodological considerations', *American Journal of Public Health*, 101, pp. 109–114. Available at: <https://doi.org/10.2105/AJPH.2011.300189>
- Wilkinson, M., Dumontier, M., Aalbersberg, I., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. and Mons, B.** (2016) 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data*, 3. Available at: <https://doi.org/10.1038/sdata.2016.18>

- Wilkinson, M., Sansone, S., Schultes, E., Doorn, P., Olavo, L., da Silva Santos, B. and Dumontier, M.** (2018) 'A design framework and exemplar metrics for FAIRness', *Scientific Data*, 5. Available at: <https://doi.org/10.1038/sdata.2018.118>
- Williamson, B., Provost, S. and Price, C.** (2023) 'Operationalising Indigenous data sovereignty in environmental research and governance', *Philosophy, Theory, Models, Methods, and Practice*, 2(1–2), pp. 281–304. Available at: <https://doi.org/10.1177/26349825221125496>
- Wilson, N.J., Montoya, T., Arseneault, R. and Curley, A.** (2021) 'Governing water insecurity: navigating indigenous water rights and regulatory politics in settler colonial states', *Water International*, 46, pp. 783–801. Available at: <https://doi.org/10.1080/02508060.2021.1928972>
- Wilson, N.J., Mutter, E., Inkster, J. and Satterfield, T.** (2018) 'Community-based monitoring as the practice of Indigenous governance: A case study of Indigenous-led water quality monitoring in the Yukon River Basin', *Journal of Environmental Management*, 210, pp. 290–298. Available at: <https://doi.org/10.1016/j.jenvman.2018.01.020>
- Yang, C., Raskin, R., Goodchild, M. and Gahegan, M.** (2010) 'Geospatial cyberinfrastructure: Past, present and future', *Computers, Environment and Urban Systems*, 34(4), pp. 264–277. Available at: <https://doi.org/10.1016/j.compenvurbsys.2010.04.001>
- Young, R., Colby, B. and Thompson, G.** (2019) 'Tribal water rights, community economies, and adaptive water institutions in the western United States', *Journal of Natural Resources Policy Research*, 9(1), pp. 74–102. Available at: <https://doi.org/10.5325/naturesopolirese.9.1.0074>
- Yue, P., Gong, J., Di, L. and He, L.** (2012) 'Automatic geospatial metadata generation for earth science virtual data products', *GeoInformatica*, 16, pp. 1–29. Available at: <https://doi.org/10.1007/s10707-011-0123-x>
- Zhu, M. and Cole, J.M.** (2022) 'PDFDataExtractor: A tool for reading scientific text and interpreting metadata from the typeset literature in the portable document format', *Journal of Chemical Information and Modeling*, 62(7), pp. 1633–1643. Available at: <https://doi.org/10.1021/acs.jcim.1c01198>

TO CITE THIS ARTICLE:

Sarzaeim, P. and Bulltail, G. (2025). Evaluation of FAIR Principles in Indigenous Water-Climate-Environment (WCE) Data Repositories. *Data Science Journal*, 24: 31, pp. 1–16. DOI: <https://doi.org/10.5334/dsj-2025-031>

Submitted: 13 November 2024

Accepted: 08 September 2025

Published: 06 November 2025

COPYRIGHT:

© 2025 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Data Science Journal is a peer-reviewed open access journal published by Ubiquity Press.