



NWT CIMP Data Management Requirements

All NWT CIMP funded projects are required to submit a Data Management Plan (DMP) and Project Metadata. Dataset Metadata is also required if data will be submitted directly to NWT CIMP. This guideline will be reviewed every 3 years and either amended or updated as required.

Data Management Requirements	Reference	Due Dates
Data Management Plan	Appendices A & B	To accompany the interim report (Year 1), with annual updates as needed.
Project Metadata	Appendix C	To accompany the interim report (Year 1), with annual updates as needed.
Dataset Metadata	Appendix D	To accompany the submission of datasets as project deliverables. Exact timelines depend on deliverable schedule.

Data Management Plan

The DMP describes how the Project Lead will manage and share information generated by the project. It *must* be submitted to NWT CIMP with the interim report in the first year of funding and updated annually as needed. The DMP must be appropriate for the type of data being collected and reflect the best practices and standards in the area of research or monitoring being proposed.

NWT CIMP has provided a template (Appendix A) to guide development of your DMP, and an example (Appendix B) of a completed DMP. Other DMP templates can be used, provided they capture the same information.

Project Metadata

Project Metadata is the information that enables data/information to be discovered, evaluated, and reused for other purposes. As with the DMP, it *must* be submitted to NWT CIMP with the interim report in the first year of funding and updated annually if needed.

Each project *must* complete the Project Metadata template provided in Appendix C. Instructions to guide the development of your Project Metadata are also provided in Appendix C.

Dataset Metadata

Dataset Metadata is the information that documents a dataset's structure. Each dataset submitted to NWT CIMP must be accompanied by Dataset Metadata.



NWT CIMP recognizes that the datasets produced by each project vary widely. If there is an existing metadata standard that is applicable to your project, NWT CIMP encourages you to use it. Examples include Darwin Core, Global Biodiversity Information Facility, and Ecological Metadata Language. However, some projects may need to develop their own Dataset Metadata format. If a dataset metadata format is required for aquatic data, the [NWT CIMP Metadata Standards for Reporting Water Quality Information in the NWT](#) template can be used.

When creating your Dataset Metadata, each data field name (e.g., column name) must be listed and accompanied with a description of the information contained within each data field. Key information, such as the format of the date (e.g., YYYY-MM-DD), Geographic Coordinate System (e.g., NAD27, NAD83, UTM), and units, should be included so that the data are understood. If applicable, detection limits or measurement error should be included as well. For categorical data, all categories should be listed and defined.

An option for clear, concise data management is to generate multiple datasets that can be linked together through common data fields. Common data fields (e.g., site ID, survey respondent ID) allow different datasets to be linked for analysis, without duplicating the same information in many places which may introduce errors. For example, if your project collects multiple types of data at the same sites across many years, one dataset could contain all site level information (e.g., site coordinates, habitat characteristics), with separate datasets for each type of data collected (e.g., fish abundance, zooplankton diversity, phytoplankton diversity, and water chemistry). The site information and the ecological data could be linked by common site ID column in all datasets. Each of these datasets should be accompanied by dataset metadata; common data fields should be identified in the metadata and named consistently across all datasets.

If the data are published elsewhere online (e.g., another online repository such as Dryad or OpenScience), provide a link to the dataset metadata and data and submit nothing further. NWT CIMP has provided an example (Appendix D) to guide development of your Dataset Metadata.

Data File Formats

Datasets should be saved in a durable file format to enable access by anyone in the future. This means using readily available programs and that don't require proprietary software. Common examples include:

Text documents	Spreadsheets/ tabular data	Web pages	Images	Audio files	Video files
.TXT	.CSV	.HTML	.PNG	.MP3	.MP4
.DOCX	.XLSX	.XML	.JPG	.WAV	.AVI
		.XSLT	.TIFF		.MPG



Appendix A: NWT CIMP Data Management Plan Template

Project Title and Number (CIMP###):

Project Lead (name):

Lead Organization:

Project Objectives: *(insert from your proposal)*

This Data Management Plan describes how the Project Lead will manage and share information generated by the project. Please complete all sections below.

- 1) Describe the data, samples, software, presentations or curriculum materials, etc. that will be produced during the project.

- 2) Identify who is responsible for creating Project Metadata (mandatory) and Dataset Metadata (mandatory if data is to be shared with NWT CIMP). Describe/provide the details on methods and procedures etc. When applicable, NWT CIMP encourages the use of existing community standards for Dataset Metadata.

- 3) Describe the plans for access and sharing knowledge, including appropriate protection of privacy, confidentiality, security, licensing, intellectual property, or other rights and requirements. Address the following:
 - a. Will a data sharing agreement be required?



b. Who will data be made available to and when?

4) Describe the plans for archiving physical samples. Address the following:

- a. Where and how will physical samples be stored?
- b. How long will they be stored and who is responsible for storage?
- c. May they be used in future research projects, and if so, what permissions must be obtained?

5) Describe the plans for archiving data. Address the following:

- a. Where and how will electronic data be stored?
- b. How will the data be stored and backed up during the research?
- c. How will the data be stored at the end of the research?
- d. Who holds the authoritative copy?
- e. May the data be used in future research projects, and if so, what permissions must be obtained?



Appendix B: Data Management Plan Example

Project Title and Number (CIMP###): CIMP643

Project Lead (name): Jane Doe

Lead Organization: GNWT

Project Objectives: *(insert from your proposal)*

This project aims to develop the basis for a cumulative impact monitoring program for aquatic effects of an area of intense oil and gas exploration in the Sahtu region, in particular, for watersheds draining the eastern foothills of the Mackenzie Mountains.

The main objectives are to:

- 1- Build a partnership between government, communities, regulators and researchers to develop a watershed-based cumulative impact study in areas of oil and gas exploration in the Sahtu region.
- 2- Develop a framework to determine watershed characteristics, disturbance level and human use within key watersheds of development and community interest.
- 3- To develop baseline water quality conditions and to assess the health of stream ecosystems as a baseline against which to examine the cumulative impacts of development.
- 4- Contribute to capacity building and decision making in the region.

This Data Management Plan describes how the Project Lead will manage and share information generated by the project. Please complete all sections below.

- 1) Describe the data, samples, software, presentations or curriculum materials, etc. that will be produced during the project.

Data collected/to be collected: surface water grabs from 10 streams in the Central Mackenzie Valley – processed by Taiga Labs (Yellowknife), benthic invertebrate samples using ECCC’s CABIN protocol, physical site characteristics collected using the CABIN protocol.

Deliverables (over the 3 years of funding): water quality data to be uploaded to the Mackenzie DataStream, benthic invertebrate data to be uploaded to ECCC’s CABIN database, presentations at NWT Geoscience, participation in the Norman Wells and Tulita Experiential Science Program, community meetings in Norman Wells and Tulita, publication of data collected in scientific literature.



- 2) Identify who is responsible for creating Project Metadata (mandatory) and Dataset Metadata (mandatory if data is to be shared with NWT CIMP). Describe/provide the details on methods and procedures etc. When applicable, NWT CIMP encourages the use of existing community standards for Dataset Metadata.

Jane Doe will be responsible for the Project Metadata, Dataset Metadata and providing detailed methods. ECCC's CABIN protocol (<https://www.canada.ca/en/environment-climate-change/services/canadian-aquatic-biomonitoring-network.html>) will be used to collect and archive benthic invertebrate data.

- 3) Describe the plans for access and sharing knowledge, including appropriate protection of privacy, confidentiality, security, licensing, intellectual property, or other rights and requirements. Address the following:
 - a. Will a data sharing agreement be required?
 - b. Who will data be made available to and when?

Data will be made available to the communities (Norman Wells and Tulita) immediately after data quality control is complete. Water quality data will be uploaded to the Mackenzie DataStream (<https://mackenziedatastream.ca>) annually and available to the general public. Benthic invertebrate data will be uploaded annually to ECCC CABIN and made available by request. No data sharing agreements are required.

- 4) Describe the plans for archiving physical samples. Address the following:
 - a. Where and how will physical samples be stored?
 - b. How long will they be stored and who is responsible for storage?
 - c. May they be used in future research projects, and if so, what permissions must be obtained?

Water samples will be disposed of 2 months after analysis by Taiga Labs. Unless another researcher asks for physical access to our invertebrate samples, they will be disposed of by the analytical lab that was contracted for identification.

- 5) Describe the plans for archiving data. Address the following:
 - a. Where and how will electronic data be stored?
 - b. How will the data be stored and backed up during the research?
 - c. How will the data be stored at the end of the research?
 - d. Who holds the authoritative copy?
 - e. May the data be used in future research projects, and if so, what permissions must be obtained?



By the end of the research project, all of the water quality data will be archived on the Mackenzie DataStream (MDS) website. Benthic invertebrate data will be archived on ECCC's CABIN database. Water quality and benthic invertebrate data will also be archived on the GNWT's server.

Over the course of the project, data will be saved on both the GNWT server and an external drive. In the field, data collected will be transcribed and backed up daily. MDS will hold the authoritative copy of the water quality data, and ECCC CABIN will hold the authoritative copy of the benthic invertebrate data. All data collected may be shared with the public. No permissions are required to obtain water quality data as it is available on the MDS website. CABIN data also can be obtained, but the interested party must first contact the administrators at ECCC-CABIN to access the data.



Appendix C: NWT CIMP Project Metadata Template

Project Title and Number (CIMP###):

Title(s) of All Included Datasets

- Title each dataset included in the NWT CIMP project in a way that describes the data collected, geographic context, research site, and time frame (what, where, when; e.g., Long-term vegetation data from Qikitarug-Herschel Island, Yukon, 2005-2015).
- There may be multiple separate datasets (e.g., water chemistry, phytoplankton, zooplankton, fish datasets for a series of sites).

Data Citation

- Format for data citation if used in other research (e.g., as unpublished data, journal reference, raw data citation).

Study Site

- Identify where the data were collected. Please verify the spelling of your study site using the [Canadian Geographical Names Data Base](#) (e.g. Nunavik, Québec, Canada).
- We encourage the use of traditional Indigenous names for geographical features and populated places.

Project Objectives

- A summary of the study objectives and the broader application of the research taken from the Project Proposal.



Abstract (plain language)

- A plain language description of the project and information/data collected (e.g., interviews, physical and chemical variables, imagery, recordings, maps).
- In single/few species monitoring, state the focal species.
- Include basic methods and site selection.

Responsible Parties

- Provide the names and contact information of the people associated with the data and identify their roles. The definitions of the roles can be found below.

Role	Name and Affiliation	Contact Information <i>(email and phone number)</i>
Project Lead		
Originator <i>Primary responsible party (first author) for the dataset.</i>		
Point of Contact <i>A persistent point of contact for acquiring knowledge about the dataset, such as a professor or permanent employee.</i>		
Collaborator <i>Other parties who contributed to the dataset</i>		

Links to External Data Repository

- If the data are published elsewhere online, please provide the link to the data and dataset metadata.



Links to Publications

- If applicable, provide a link to any publications using this data.

Data Status

- Please choose one of the following: Complete, In Progress, Planned

Maintenance and Updating Frequency of Project Metadata

- Identify the planned maintenance and updating frequency (e.g., Annually, Daily, Monthly, Weekly, None Planned, Unknown)

Data Security

- Identify the appropriate protection level for privacy, confidentiality, security, intellectual property, or other rights and requirements. Protection levels may vary across datasets, or apply to the entire NWT CIMP project. Options are:
 - Public
 - Limited: Data involve human subjects
 - Limited: Data involve intellectual property issues related to local or Traditional Knowledge
 - Limited: Release of data may cause harm to the environment or the public
 - Limited: Pre-existing data have been used and are subject to access restrictions
 - Limited: Currently under analysis for publication



Study Site Coordinates

- Coordinates encompass your study area and are usually a box spanning from the most northern to the most southern latitudes along with the most western and eastern longitudes. Bounding coordinates must be entered in decimal degrees. If you have a small study site, please enter the same latitude in both the North and South and the same longitude in both West and East.
- Coordinates MUST be between -90 and 90 for latitude and between -180 and 180 for longitude.

North (Latitude N): South (latitude N): West (longitude E): East (longitude E):
--

Time Period

- Specify the time period during which the data were collected including the start and end dates.

--



Appendix D: Dataset Metadata Example

Dataset Title: Sahtu river site characteristics 2005-2015.csv

Column Name	Description	Units (if applicable)
StreamID	Unique stream identification number	
Latitude	Latitude	NAD83/decimal degrees
Longitude	Longitude	NAD83/decimal degrees
WatershedID	Unique watershed identification number	
Area	Watershed Area	Hectares/km ²
YearBurned	The most recent year of a forest fire within this stream's watershed prior to data collection in 2015 based on the National Fire Database	

Dataset Title: Benthic Invertebrate community assemblage data from Sahtu rivers 2015.csv

Column Name	Description	Units (if applicable)
StreamID	Unique stream identification number	
CollectionDate	YYYY-MM-DD	
128689*	Abundance of genus <i>Eukiefferiella</i>	Individuals/L
128707*	Abundance of genus <i>Euryhapsis</i>	Individuals/L
128737*	Abundance of genus <i>Heterotrissocladius</i>	Individuals/L
128771*	Abundance of genus <i>Krenosmittia</i>	Individuals/L
128874*	Abundance of genus <i>Orthocladius</i>	Individuals/L

***Note:** Numeric column headings refer to unique and stable Taxonomic Serial Number in case of reclassification in the future