­­­­NPS Metadata Template (2024)­­­

Purpose and Introduction

Metadata is key component of a data package, as it helps others to understand your data and assess whe­ther it’s a good fit for a particular purpose. This template can help organize all the bits of information that must come together to create metadata inside of your data package. Overall guidance on metadata and data packages can be found on the Data Publication Best Practices SharePoint.

Data Package Title

(Include **what, where,** and **when**. E.g. “Monthly Water Quality Data from Horsetooth Reservoir, Colorado: 2010-2019”)

|  |
| --- |
|  |

Metadata Filename

(Similar to Data Package Title, should be informative. Be sure it ends in **\_metadata** to comply with data package specifications. This will become the file name of your .xml. Example: RMNP\_Mammals\_2020\_metadata)

|  |
| --- |
|  |

Data Files, Names, and Descriptions

(List your data files, give them an informative name and description. Descriptions should be unique and about 10 words long)

|  |  |  |
| --- | --- | --- |
| **Data File**  **(e.g. SEUG\_crustClassData.csv)** | **Informative Name**  **(e.g. SEUG LTVM Biocrust Data)** | **Description**  **(e.g. Biological soil crust development class data)** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Taxonomic Information

(List the data file(s) with your taxonomic information, including the scientific name field within that data file. We suggest using [DarwinCore](https://dwc.tdwg.org/terms) for column names, such as “scientificName”. If your data package does not have taxonomic data, skip this step.)

|  |  |
| --- | --- |
| **Data File**  **(e.g. qry\_Export\_AA\_VegetationDetails.csv)** | **Scientific Name Column**  **(e.g. scientificName)** |
|  |  |

Geographic Information

(List the data file(s) that contain geographic information. Please ensure your geographic information is in **decimal degrees**. If your coordinates are in UTMs, the the [convert\_utm\_to\_ll()](https://nationalparkservice.github.io/QCkit/reference/convert_utm_to_ll.html) function in [QCkit](https://nationalparkservice.github.io/QCkit/) can help. If your data package does not have geographic information, you can skip this step.)

|  |  |  |  |
| --- | --- | --- | --- |
| **Data File**  **(e.g. qry\_Export\_AA\_points.csv)** | **Decimal Latitude Column**  **(e.g. decimalLatitude)** | **Decimal Longitude Column**  **(e.g. decimalLongitude)** | **Site Name Column** **(e.g. Point\_ID)** |
|  |  |  |  |

Content Units

(These are the park units where data were collected. If the data package includes data from more than one park, they can all be listed. For instance, if data were collected in all network park units, each unit should be listed separately rather than by the network code.)

|  |
| --- |
|  |

Producing Units

(This is the unit(s) responsible for generating the data package. It may be a single park (ROMO) or a network (ROMN). It may be identical to the units listed in the previous step, overlapping, or entirely different.)

|  |
| --- |
|  |

Data Collection Status

Ongoing  Complete

Timeframe

|  |  |
| --- | --- |
| **Begin Date** | **End Date (leave empty for ongoing)** |
|  |  |

Abstract

(Include what, why, where, when, and how.)

|  |
| --- |
|  |

Methods

(Describes the data creation methods. Includes enough detail for future users to correctly use the data. Be specific about the study design and field and lab methods for collecting and processing the data. Protocol can be cited. It may also be appropriate to cite the datasets that were ingested to generate the data package, software (e.g. R), packages (e.g. dplyr, ggplot2) or custom scripts.)

|  |
| --- |
|  |

Creators

**(These are the people who will show up as authors in the dataset citation.** These are the individuals who have provided intellectual or other significant contributions to the creation of this dataset, much like the authors of a research paper. Valid EML requires at least one person with a **creator** role.)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **First Name** | **Middle Initial** | **Last Name** | **Organization** | **Email address** | **ORCID ID (optional)** | **Role in project** |
|  |  |  |  |  |  | Creator |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

Other personnel names and roles

(Who should a data user contact with questions about these data? You **must** enter a person or organization name to serve as the **contact** for this dataset. If this is the same person as the creator, list that person twice. You may also list other personnel who participated in the project (such as field crew, lab tech, data entry etc.) Persons serving more than one role are listed on separate lines. Other roles (e.g. Field Technician) will be listed as associated parties to the data. Their specific role (e.g. “Field Tech” will also be listed in metadata))

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **First Name** | **Middle Initial** | **Last Name** | **Organization** | **e-mail address** | **ORCID ID (optional)** | **Role in project** |
|  |  |  |  |  |  | Contact |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

Dissemination Level

(Select a Dissemination Level based on the CUI status of your dataset. This is a required step. You can choose from one of five dissemination codes. Watch out for the spaces!)

PUBLIC *(Does NOT contain CUI.)*

FED ONLY *(Contains CUI. Only federal employees should have access.)*

FED CON *(Contains CUI. Only federal employees and federal contractors should have access.)*

NOCON *(Contains CUI. Federal, state, local, or tribal employees may have access, but contractors cannot.)*

DL ONLY *(Contains CUI. Should only be available to a named list of individuals.)*

|  |
| --- |
| **DL ONLY Names** |
|  |
|  |
|  |
|  |
|  |
|  |

More information about these codes can be found at: <https://www.archives.gov/cui/registry/limited-dissemination>).

Intellectual Rights Statement

(The EMLeditor tool can assist in modifying the intellectual rights statement that is embedded in metadata. Government works shared with the public are usually public, i.e. public domain, or CC0, and anything with CUI should be set to ‘restricted’.)

Public *(Does not contain CUI. The intellectual rights will read: “This work is in the public domain. There is no copyright or license.”)*

CC0 *(Does not contain CUI. The intellectual rights will read: “The person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law. You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.”)*

Restricted (*Contains CUI.*)

Keywords

(List keywords below and separate with commas. Using keywords from a controlled vocabulary (CV) will improve the future discovery and reuse of your data. The LTER CV is a good source for keywords. Access the LTER CV [here](http://vocab.lternet.edu/vocab/vocab/index.php). Also, please determine one or two keywords that best describe your park, station, and/or project (e.g., Trout Lake Station, NTL LTER).)

|  |
| --- |
|  |

Data Table(s)

(Provide a Table Name, Table Description, and description of each column in your data table):

* **Column Name**: This name must be exactly as it appears in the dataset. Please avoid special characters (like & or \), dashes and spaces. Underscores are permissible. Do not begin a column name with a number.
* **Description**: Please give a specific definition of the column name. This can be lengthy.
* **Class:** Column class. Valid options are **numeric**, **categorical**, **character**, and **date**.
* **Unit:** Identify units for all numeric variables. Please avoid special characters and describe units in this pattern: e.g., microSiemenPerCentimeter, microgramPerLiter, absorptionPerMolePerCentimeter
* **Date Time Format**: Please tell us exactly how the date and time is formatted: e.g. mm/dd/yyyy hh:mm:ss plus the time zone and whether or not daylight savings was observed. ISO 8601 date format of YYYY-MM-DD or YYYY-MM-DD hh:mm:ss is preferred.
* **Missing Value Code**: If a code for ‘no data’ is used, please specify: e.g., -99999
* **Missing Value Code Explanation**: Definition of missing value code.

**Table name:** (A short name for this table)

**Table description:** (Add brief description of table contents)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Column name** | **Description** | **Class** | **Unit** | **Date Time Format** | **Missing Value Code** | **Missing Value Code Explanation** |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

(Copy this table to document more than one data table.)

Categorical Variables (Catvars)

(Describes categorical variables of a data table (if any columns are classified as categorical in table attributes).)

* **Attribute Name**: Column name
* **Code**: Categorical variable
* **Definition :** Definition of categorical variable

|  |  |  |
| --- | --- | --- |
| **Attribute** | **Code** | **Definition** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

(Copy this table to document more than one data table.)

Data Release Report (DRR)

(Indicate whether there is an associated DRR with your data package.)

No  Yes, it already exists  Yes, plan to generate one with the Data Strike Team

(If you have a DRR and there is an existing reference for it on DataStore, fill out the table below. Otherwise, you can skip this step.)

|  |  |
| --- | --- |
| **DRR Title** | **DRR Reference Number (from DataStore)** |
|  |  |

Additional notes and comments

|  |
| --- |
|  |