Intermediate Metadata

A tracking document to use while developing data

This document is for you to use AS you work on your research project. You can consider this a DRAFT until you are ready to write formal metadata and contribute your metadata and data to a data repository (you will be able to cut/paste information from this document into your formal metadata).

***This will help you because it will****:*

1. Make it easier to create final standards-based metadata
2. Help you write the methods sections of your manuscripts and thesis

***Hold these questions in mind as you work on this:***

1. What details would someone need to know about my data, analysis, and results if I wasn’t around to tell them? (i.e., more details than are typically included in a manuscript!)
2. What information would someone need to know to reproduce exactly what I did? Also handy if you need to go back and repeat or troubleshoot steps in your process.
3. What details are relevant to methods sections of manuscripts I plan to write?

***What next, you have finished your metadata and are ready to upload data to the web:***

1. Visit <https://www.northwestknowledge.net/>
2. Select “Contribute Data”
3. Request an Account
4. Once you have a username and password, log-in and follow the step-by-step guide to enter metadata and upload your data files

|  |  |
| --- | --- |
|  |  |
| **Title**(the WHAT*)* | *A brief description of your data; include 'what', 'where,' and 'when'. Examples: LiDAR data for Clear Creek Watershed, Idaho (2009); Water clarity for Lake Coeur d'Alene from IDEQ source data, Idaho (2005-2014)* |
|  |
| **Summary & Purpose**(the WHAT & the WHY*)* | *A brief statement of why you are creating this dataset and a short description of the data (e.g., what it is, how it might be used, variables included, etc.). This paragraph is similar to an abstract for a poster or manuscript.* |
|  |
| **Keywords** | *Think carefully about keywords because this is often how people locate your data in a search engine. In addition to keywords for topic, consider including keywords for location, grant number, and method. Include both broad (example: biology) and specific (example: lesser goldfinch) terms.* |
|  |
| **Funding Sources** | *How was your work funded? Under the MILES grant (Idaho EPSCoR IIA-1301792)? List all sources of funding that supported your work, add the grant number if possible.* |
|  |
| **Data Production and Source**(the HOW)\* You may or may not need to fill in all these field depending on your situation | *What are the meanings of column headings in your spreadsheet and/or attribute fields in your GIS attribute table? This is especially important if you use codes or shorthand notations. Consider including units in your column headings, e.g. “Depth\_m" rather than “Depth” to note that depth values are in meters.* *\*You may include all details here or provide a separate text document that lists all the headings or attribute fields and their meaning, which you store alongside your data.* |
|  |
| *How was the data collected? Where? How often? By whom? Using what instruments, including manufacturer? How were the instruments deployed and data recorded (e.g., by hand, field notebooks; in situ, instrument recorded)? What parameters and procedures were used to calibrate the instrument? What variables were measured, including units?* |
|  |
| *Did you build upon someone else’s published dataset? List any published data you are using as input to your analysis, include the data source, version (number or data acquired), url (if available), and DOI (if available).*  |
|  |
| *What did you do to your raw data to alter it from its original form for analysis? Example: You applied a correction factor to your in situ dissolved oxygen data and smooth over small data gaps to create a ‘clean’ time series records.*Example: Y*ou clipped it to your geographic area of interest and transformed it to different units. Then you aggregated it into a new classification.**\* You may include all details on this sheet or provide a reference to another location where you track this information.*  |
|  |
| *What steps did you take to analyze your data? How did you take your data and turn it into information? What software, models, modeling platforms, and/or programming languages did you use, including what version? For each step include any parameters, settings, code, etc. that you used.* *\* You may include all details on this sheet or provide a reference to another location where you track this information.*  |
|  |
| **Temporal Extent**(the WHEN) | *What time period does your data cover? (example: July 1994 to August 2012)* |
|  |
| **Date**(the WHEN) | *When was the file last updated or modified? (example: created on 21 December 2013)* |
|  |
| **Data Update Status**(the WHAT*)* | *Is the dataset* ***complete*** *(you anticipate no further work and consider it final) or is it* ***continually updated****? Example: a data product that updates with new climate data inputs on a weekly basis is continually updated.* |
|  |
| **Collection or Single Dataset?**(the WHAT*)* | *Is the data you are describing a single* ***dataset*** *or a collection of datasets that have the same specification known as a* ***series****? Example: A climate data series includes the same variables but over 1 year at one-day intervals; this series would contain 365 datasets.* |
|  |
| **Data Update Frequency** | *If you expect there to be updates, how often do you anticipate this will happen? Example: continual, daily, weekly, fortnightly, monthly, quarterly, biannually, annually, as needed, irregular, not planned, unknown.* |
|  |
| **Location or Spatial Extent**(the WHERE) | *What spatial area does your data cover? You will need to either select an area on a map or enter bounding coordinates when you create the final metadata. For now, at a minimum, describe the location of your data. Example: the state of Idaho, W Lon -117.531786, E Lon -110.655421, S Lat 41.946097, N Lat 49.039542.*  |
|  |
| **Reference System**(the WHERE) | *If you are working with geographic or GIS data, include the coordinate system (example: UTM) and datum (example: NAD83) used for your data. This will allow people to project your GIS files correctly.* |
|  |
| **Authors**(the WHO) | *Who are the author(s) of this dataset? Consider who might be the author of a manuscript describing the data (likely PI, post-doc, or graduate student). You will add contact detail for each when you enter metadata. For now, list names and institutions.* |
|  |
| **Data Format**(the WHAT*)* | *In what format will your final dataset be published? For example, csv, raster, NetCDF, vector, text. You might have multiple to list.* |
|  |
| **Use Restrictions?** | *What do you know about how the data should and shouldn’t be used that you want to convey to someone who wants to use your data but doesn’t have the privilege of working with you directly?* |
|  |
| **Publishing & Archiving** | *Where do you plan to publish and store your data when it is complete and ready to share, e.g., NKN’s data portal, BSU’s data portal, other public data repositories? Is there anywhere else online where others can access your data, e.g. GIS web services? Do you plan to request a digital object identifier (DOI) for your final original data so that it can be cited by others and listed as a data publication on your CV? [Note: NKN is can issue DOIs after data is uploaded. Contact Carrie Roever for more details.]**\*\* You can publish your metadata and data and ‘embargo’ your data for a certain time period (1-2 yr) to ensure that you have time to publish your results in manuscripts before the data is publically available.*  |
|  |
| **Other Details, Notes to Yourself** | *List any other details relevant to the questions listed at the top of this document.* |
|  |