Tutorials on Data Management

Lesson 8: How to Write Quality Metadata

DataONE

Lesson Topics

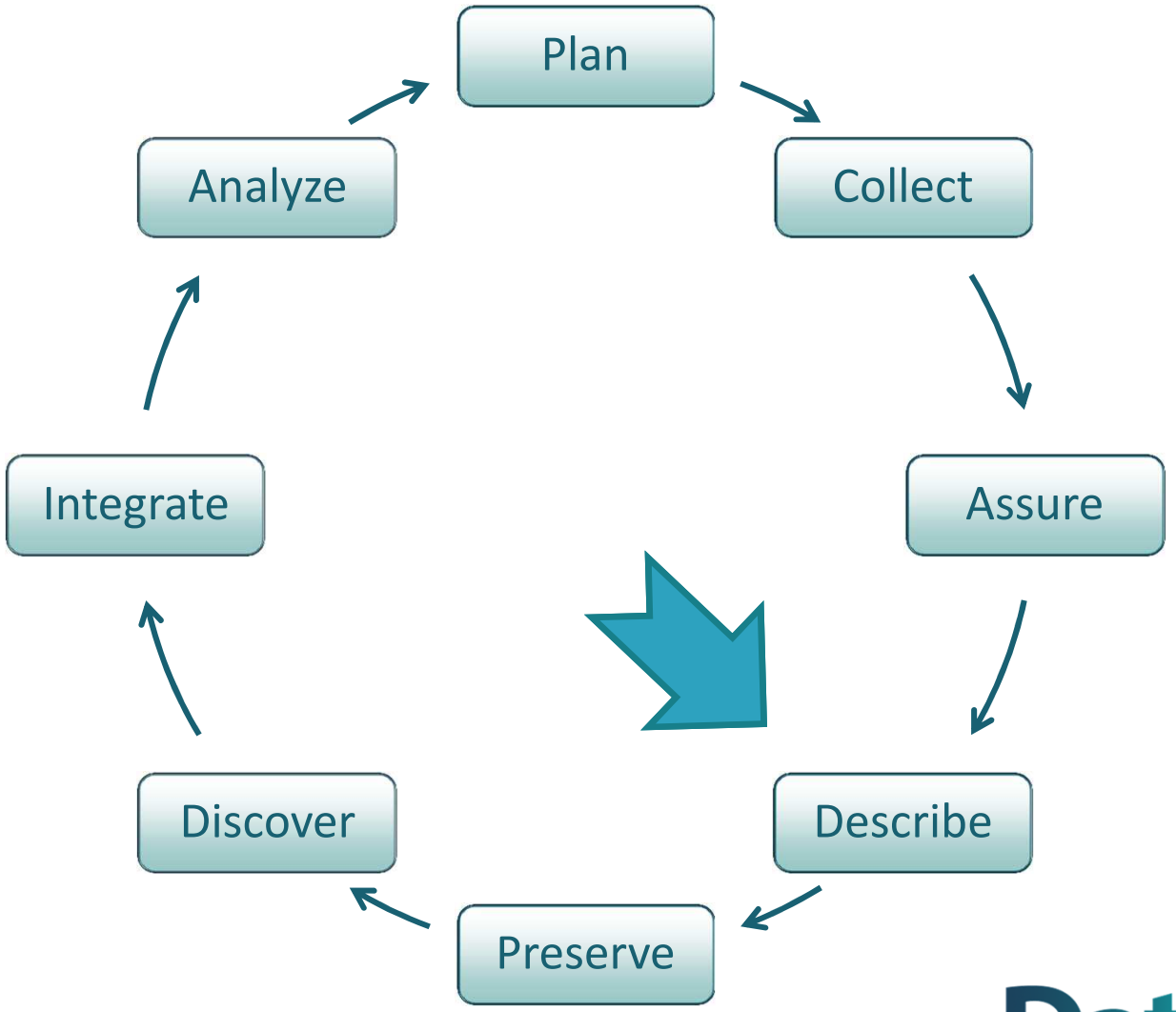
- Preparing to write metadata
- Tips for writing a quality metadata record

Learning Objectives

- After completing this lesson, the participant will be able to:
  - List steps to prepare to write metadata
  - Explain how to write good metadata

DataONE

The Data Life Cycle



DataONE

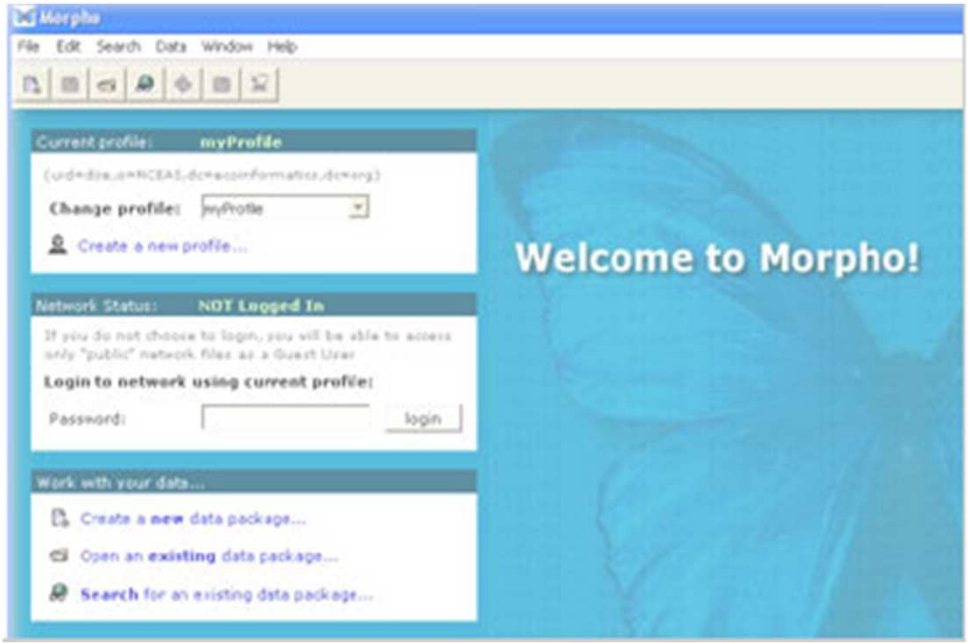
Steps to Create Quality Metadata

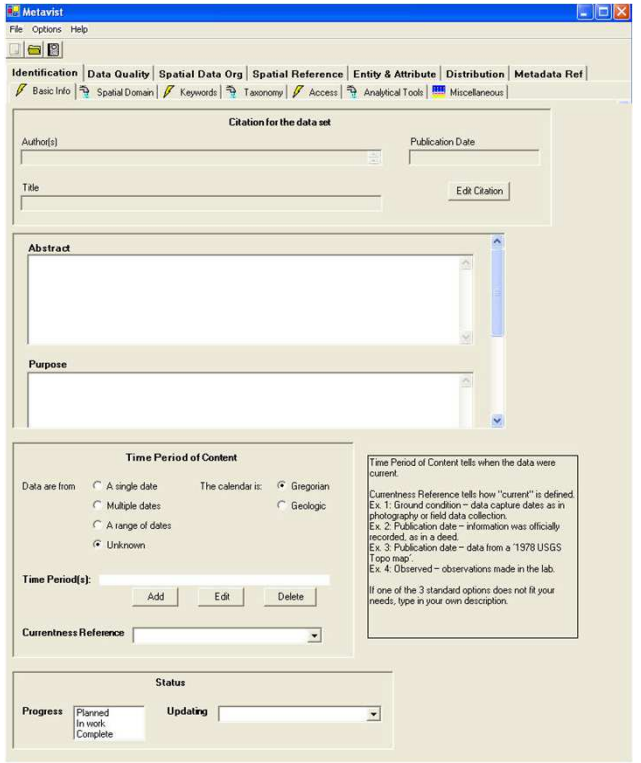
- Organize your information
  - Did you write a project abstract to obtain funding for your proposal? Re-use it in your metadata!
  - Did you use a lab notebook or other notes during the data development process that define measurements and other parameters?
  - Do you have the contact information for colleagues you worked with?
  - What about citations for other data sources you used in your project?
- Write your metadata using a metadata tool
- Review for accuracy and completeness
- Have someone else read your record
- Revise the record, based on comments from your reviewer
- Review once more before you publish

DataONE

Steps to Create Quality Metadata, con't

- Write your metadata using a metadata tool
- Review for accuracy and completeness
- Have someone else read your record
- Revise the record, based on comments
- Review once more before you publish





DataONE

Steps to Create Quality Metadata, con't

- Review for accuracy and completeness
- Have someone else read your record
- Revise the record, based on comments from your reviewer
- Review once more before you publish

CC image by Shelly Munkberg on Flickr

CC image by mujalifah on Flickr

Southwest Non-native Invasive Plant Database (SWEMP07)

Metadata:

- Identification_Information
- Data_Quality_Information
- Spatial_Data_Organization_Information
- Spatial_Reference_Information
- Entity_and_Attribute_Information
- Distribution_Information
- Metadata_Reference_Information

Identification_Information:
Citation:
Citation_Information:
Originator:
Kathryn Thomas and Patricia Guertin. U.S. Geological Survey, Southwest Biological Science Center (USGS-SBSC)
Publication_Date: 20070508
Title: Southwest Non-native Invasive Plant Database (SWEMP07)
Geospatial_Data_Presentation_Form: tabular digital data
Online_Linkage:
http://sbsc.wr.usgs.gov/research/projects/swepic/swemp/swempA5.asp
Description:
Abstract:
The Southwest Exotic Plant Mapping Program (SWEMP) is a collaborative effort between the United States Geological Survey and federal, tribal, state, county and NGO partners in the southwest. This project is an ongoing effort to compile and distribute regional data on the occurrence of non-native invasive plants in the southwestern United States. The database represents the known sites (represented by a point location, i.e. site) of non-native invasive plant infestations within Arizona and New Mexico, and adjacent portions of California, Colorado, Nevada and Utah. These data, collected from 1911 to 2006, represent the field observations of various state, federal, tribal and county agencies, along with some specimen data from Herbaria. The SWEMP database is published at least once a year and consists of a compilation of all data submitted up to the date of publication.
Purpose:
This dataset was created to provide a regional perspective on non-native invasive plant distributions. It can be used to assist land managers, as well as the public, to review the locations and extent of reported infestations. These data can ultimately help guide management strategies and policies for the control of non-native invasive plant species. All plant species in the database are non-native as defined by the USDA PLANTS database 2007; the extent to which they are invasive has not been determined.

DataONE

Tips for Writing Quality Metadata

- Do not use jargon
- Define technical terms and acronyms:
  - CA, LA, GPS, GIS : what do these mean?
- Clearly state data limitations
  - E.g., data set omissions, completeness of data
  - Express considerations for appropriate re-use of the data
- Use “none” or “unknown” meaningfully
  - None usually means that you knew about data and nothing existed (e.g., a “0” cubic feet per second discharge value)
  - Unknown means that you don’t know whether that data existed or not (e.g., a null value)

DataONE

Tips for Writing Quality Metadata

Titles, Titles, Titles…

- Titles are critical in helping readers find your data
  - While individuals are searching for the most appropriate data sets, they are most likely going to use the title as the first criteria to determine if a dataset meets their needs.
  - Treat the title as the opportunity to sell your dataset.
- A complete title includes: What, Where, When, Who, and Scale
- An informative title includes: topic, timeliness of the data, specific information about place and geography

DataONE

Tips for Writing Quality Metadata

- A Clear Choice: Which title is better?

CC image by dolfi on Flickr

- *Rivers*

OR

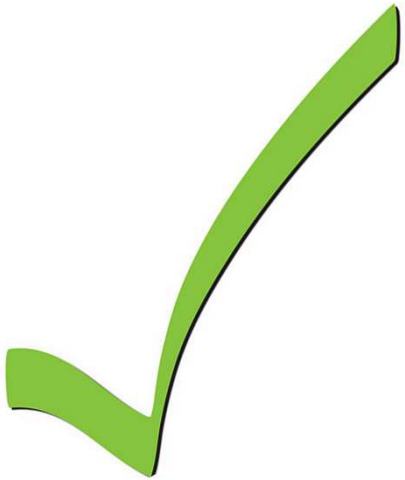
- *Greater Yellowstone Rivers from 1:126,700 U.S. Forest Service Visitor Maps (1961-1983)*

Greater Yellowstone (where) Rivers (what) from 1:126,700 (scale) U.S. Forest Service (who) Visitor Maps (1961-1983) (when)

DataONE

Tips for Writing Quality Metadata

- Be specific and quantify when you can! The goal of a metadata record is to give the user enough information to know if they can use the data without contacting the dataset owner.



Vague: We checked our work and it looks complete.

Specific: We checked our work using a random sample of 5 monitoring sites reviewed by 2 different people. We determined our work to be 95% complete based on these visual inspections.

DataONE

Tips for Writing Quality Metadata

- Select keywords wisely
- Use descriptive and clear writing
- Fully qualify geographic locations
- Use thesauri for keywords whenever possible
- Example: USGS Biocomplexity Thesaurus (over 9,500 terms)

DataONE

Tips for Writing Quality Metadata

- Remember: a computer will read your metadata
- Do not use symbols that could be misinterpreted: Examples: ! @ # % { } | / \ < > ~
- Don’t use tabs, indents, or line feeds/carriage returns
- When copying and pasting from other sources, use a text editor (e.g., Notepad) to eliminate hidden characters

DataONE

Summary

- Review your final product:

Does the documentation present all the information needed to use or reuse the data?

- Remember: a well-written title and good keywords are critical in data discovery

The full slide deck may be downloaded from:

http://www.dataone.org/education-modules

Suggested citation:

DataONE Education Module: How to Write Good Quality Metadata. DataONE. Retrieved Nov12, 2012. From http://www.dataone.org/sites/all/documents/L08_WriteQualityMetadata.pptx