

Defintion: Metadata

Metadata is documentation about the data. It describes the content, quality, condition, and other characteristics of a dataset. Metadata allows data to be discovered, accessed, and reused.



Data Life Cycle

Formatting tips

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



CC-BY-NC-SA coloneljohnbritt

Steps to creating quality metadata

- Organize your information
- Use a metadata tool
- Review for accuracy and completeness
- Have someone else read your metadata
- Revise the record based on reviewer feedback
- Review once more before you publish

Be specific and quantify when you can

One option: We checked our work and it looks complete.

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

Titles

Titles are critical in helping readers find your data

- » They will use it as the first criteria for determining whether your data meets their needs

Complete title includes:

- » what | where | when | who | scale

Might also include:

- » topic | timeliness of data | place and geography

Use titles as an opportunity to sell your dataset!

Which title is better?

Choice A: Rivers

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

The final step

Review your final product to make sure that the documentation presents all the information needed to use or reuse the data.

Tips for writing quality metadata

- Do not use jargon
- Define technical terms and acronyms
- Clearly state data limitations, such as data omissions and completeness, and express considerations for appropriate reuse of the data
- Select keywords carefully
- Use descriptive and clear writing
- Fully qualify geographic locations
- Use thesauri for keywords whenever possible

Use “none” vs. “unknown” meaningfully

None: usually means that you knew about the data and that none existed (value of “0”)

Unknown: usually means that you don’t know whether data existed or not (value of “null”)

Local contact information

