Top Ten Ways to Create Metadata That's Just Ducky

10. Create metadata *during* the development process

Write down what you know while it's fresh in mind. Metadata should be part of the process of creating data rather than a separate activity put off until the details are forgotten.

9. Use starter templates

Much of the metadata will be the same for all of your data sets, for example, DISTRIBUTOR CONTACT information. Why type it again and again? Instead, enter this information once, save it as a 'starter template,' and then concentrate on those elements that do change, such as ABSTRACT and LINEAGE.

8. Share the task

Delegate responsibility for particular elements to the people who are most familiar with them. For example, data developers are the people most familiar with the data quality and data organization elements; managers can tackle distribution policy and use constraints. Look at other people's metadata for good examples and adopt their language if it fits your data set.

7. Avoid defining your data too finely or too broadly

It's easy to become overwhelmed trying to individually document every data table and layer. On the other hand, trying to cover all of your data resources with a single metadata record will drive both you and your data users crazy. A good rule of thumb is to consider how the data resource is used – as a component of a broader data set or as a stand-alone product that may be mixed and matched with a range of other data resources.

6. Spend some time on LINEAGE

This element describes the data sources used and how a data set was created. For each processing step, provide where possible) the name, date and scale of the source data, a description of the processing steps performed, scanning or digitizing specs, software used, tolerances set, etc. This information is hard to remember later, so it's best written down as the data is created.

5. Describe entities and attributes

Spell out what those cryptic codes and abbreviations in your data set stand for. How else would anyone know what "land use code = 11" means? Include a description of each category's definition e.g., how was 'mixed forest' defined . If you have a detailed data dictionary, refer the user to it.

4. Understand data quality

Add any information that could help a user decide whether the data is accurate enough for their project. Data sets that are not useful for surveying purposes may be perfectly fine for county-level planning. Even if statistical tests are not done, any information about known problems is helpful.

COMPLETENESS: Identify data omitted from the data set that might normally be expected, as well as the reason for the exclusion. This may include geographic exclusions, 'data was not available for Smith County'; categorical exclusions, 'municipalities with populations under 2500 were not included in the study'; and definitions used, '*floating marsh* was mapped as *land*'.

ATTRIBUTE ACCURACY: Assess how "true" the attribute values may be. Add any information you have from field checks or other quality control checks.

POSITIONAL ACCURACY: How accurately is your data located? Whatever you know can be helpful ('some areas in the north appear to be off by at least 100 meters'). For more information on positional accuracy, see the Minnesota Governor's Council's *Positional Accuracy Handbook* at: https://mn.gov/mngeo/assets/nssda_o_tcm1235-758190.pdf

3. Rely on your own judgment, not just on metadata tools

Tools help with tasks that can be automated, but tools can never replace human judgment about what's important. Only people can write concise, informative abstracts; distribution policies; lineage. Think of metadata as a way for people to describe to other people important information that they should know about the data. What information would you want to receive from someone else?

2. Prioritize the backlog

If you're like many people, you have a pile of undocumented data sets. Prioritize them using whatever criteria make sense to you. Which data sets do you use a lot? Which ones are you most tired of explaining verbally? Which ones are most current? Start on the highest priority items and leave the CDs labeled "Miscellaneous Stuff" for last.

and finally…

---

1. Don't Duck Metadata!

If you think the cost of creating metadata is too high, you haven't compiled the costs of **not** creating metadata: loss of information with staff changes, data redundancy, data conflicts, liability, misapplications, and decisions based upon poorly documented data. The hardest part is getting started; once you start, you'll quickly see the benefits.

For more information about documenting data using the Minnesota Geographic Metadata Guidelines, see: https://mn.gov/mngeo/gis-data-and-maps/metadata-resources/

or contact the Minnesota Geospatial Information Office (MnGeo) at:

600 N Robert Street
Suite 2800
St. Paul, MN 55146
(651) 201-2499
gisinfo.mngeo@state.mn.us

This handout was created by MnGeo (formerly the Minnesota Land Management Information Center) as part of the Federal Geographic Data Committee's "Don't Duck Metadata" initiative. It was freely adapted from "Top Ten Metadata Errors" at: www.fgdc.gov/metadata/documents/top10metadataerrors.pdf

Updated September 2026