Top ten metadata mistakes

by Peter Schweitzer

I would enlarge the problem beyond the metadata itself to the information processing in general and the process of data management. Here's my list, Letterman-style:

12. (for Arc/Info users) **Taking time to document things that are consequences of the GIS**, like making detailed descriptions of AREA, PERIMETER, LPOLY#, RPOLY#, FNODE#, TNODE#, cover#, cover-ID, and the like.

11. (for Arc/Info users) **Simply dumping the results of ITEMS into an Entity_and_Attribute_Overview and calling that enough.** People need to know the units of measured variables, and "percent" is **not** a unit of measure.

10. **Putting too much faith in mp.** Human review is the thing that really matters. mp can help, but isn't the arbiter of what is and what is not good metadata. Prioritize errors like this, from most serious (fix) to least serious (understand and let go):
   1. Indentation problems
   2. Unrecognized elements
   3. Misplaced elements
   4. Too many of some element
   5. Missing elements
   6. Empty elements
   7. Improper element values
   8. Warnings and upgrades

9. **Making too many metadata records.** People who try to document every GIS coverage or data table can wear themselves out. Some aggregation is good for both the producer and the user. Ancillary coverages can be described as Source_Information.

8. **Not making enough metadata records.** Trying to cram all of the information about an entire research program into a single metadata record will drive you and your potential users crazy. Split when sources, processing, or spatial reference varies.

7. **Agonizing over the most difficult elements.** These include, but probably aren't limited to

   - Latitude_Resolution
   - Longitude_Resolution
   - Abscissa_Resolution
   - Ordinate_Resolution
   - Entity_Type_Definition_Source
   - Attribute_Definition_Source
and, to a lesser extent,

Attribute_Accuracy_Report
Logical_Consistency_Report

6. Misunderstanding Enumerated_Domain as numerical values. Attributes measure (or count),
categorize, or characterize real things. Those functions are expressed as Range_Domain (for
measures or counts) and Enumerated_Domain (for categories). The hardest part is how to describe
attributes whose values are simply text describing something, or are names like place names. This is
a deficiency in the FGDC standard; there should be another type of Attribute_Domain_Values for
descriptions; a better alternative might be to make Attribute_Domain_Values mandatory if applicable
and not applicable if the values are sufficiently descriptive.

5. Substituting statements about precision for statements about accuracy. I do this often,
because what I know is how variable the values are, and I don't know the true values that they
estimate.

4. Larding the metadata with uninformative values. People can honestly disagree about this, but
I find it aggravating to see more than a few N/a, unknown, not applicable, implied, or see
above (or below). Reasonable defaults should be assumed. For example, if no Process_Contact
is indicated, people should assume that the Point_of_Contact either did the processing or knows
who did, and that the people who did are either Originators or are listed in Data_Set_Credit.
Likewise, any of the elements

Entity_Type_Definition_Source
Attribute_Definition_Source
Enumerated_Domain_Value_Definition_Source

if missing, should be assumed to have the value "this report" or some similar self-reference.

3. Choosing a tool because it's free or because it's commercial. Making, maintaining, reviewing,
and reading metadata cost so much more time and energy than the tools do that price per se
shouldn't direct the choice of tools.

2. Not recognizing that "the metadata problem" involves not only tools and training, but also
work-flow strategy and even the philosophy governing how your organization interacts with
the users of its data.

1. Not asking for help from the community. Beyond all the hype and promises (empty and
fulfilled), beyond all the tools, training, and technology, what NSDI has done is bring a common
language and common purpose to a highly diverse group of people, and we have found in each
other consolation, challenge, and care.