

DRAFT -- To be published in DMT'10 Proceedings
(see <http://ngmdb.usgs.gov/Info/dmt/>)

"Recommended Citations for Unpublished GIS Files" -- Summary of a Discussion Session

By David R. Soller

U.S. Geological Survey
926-A National Center
Reston, VA 20192
Telephone: (703) 648-6907
Fax: (703) 648-6977
email: drsoller@usgs.gov

INTRODUCTION

The DMT'10 meeting provided several opportunities for group discussion of technical and map publication-related issues. These sessions mostly focused on cartographic techniques and methods of data preservation, and were highly informal in nature – they were quite useful, but did not lend themselves to a succinct and meaningful summary for these Proceedings. However, the narrow focus of this particular session, the spirited if inconclusive discussion, and the specific recommendations offered warrant the following summary. It is hoped that the notes presented below will be helpful to agencies and authors as they struggle to address the issue.

The Discussion Session introduction was this: “Increasingly, unpublished GIS files and related information are derived from pre-existing publications. Soon thereafter, or perhaps many years in the future, these files are used in new publications. How can we try to ensure that not only the unpublished GIS file, but also its source(s) of information, are informatively cited in new publications? It's critical to our science that, years from now, the original and authoritative source of all cited information can be found. This brief session will introduce the challenge and offer some suggestions.”

In previous DMT discussion sessions on this topic (Berquist, 1999; Richard, 2000; Berquist and Soller, 2001), a wide range of opinion was expressed regarding how authors and technical contributors should be attributed, both in the formal citation and in the metadata. Those who spoke in these sessions mostly used, or favored, similar approaches, but it was clear that “one size doesn't fit all”. A proscriptive approach, while useful as a guide, cannot suit all agencies and types of publications. The examples and opinions in this document extend the discussion to unpublished information, and should be considered in that same light – they are recommendations based on personal experience that, I hope, will contribute to improvements in managing and documenting unpublished map data.

THE CHALLENGE

As noted above, it's critical to our science that, years from now, the original and authoritative source of all cited information can be found. Here is a hypothetical, but plausible, example of how knowledge of that source can be lost (figure 1):

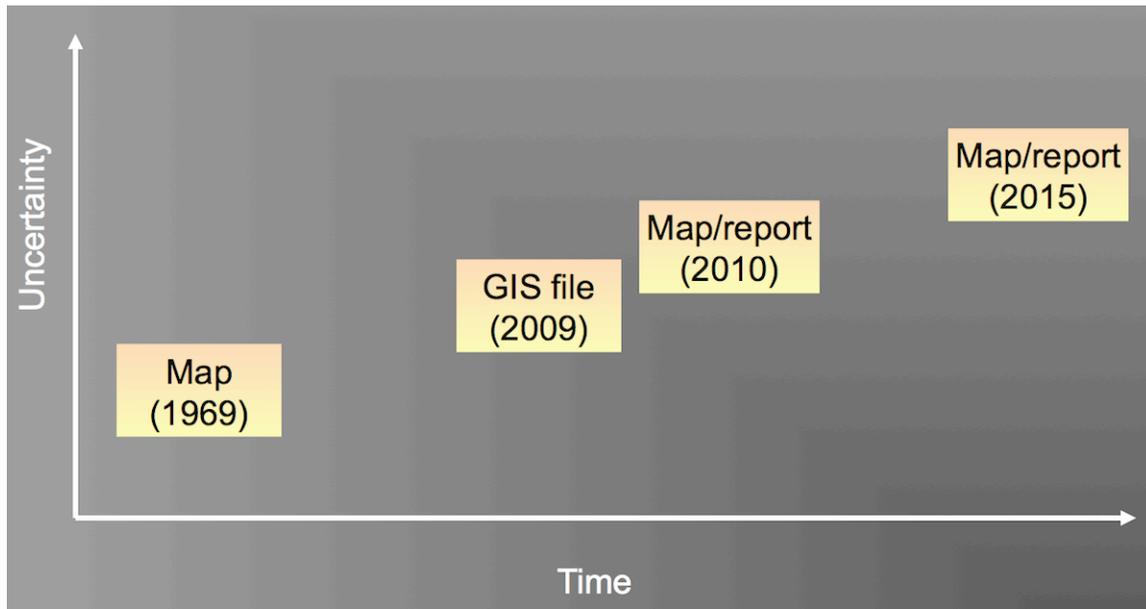


Figure 1. Four map products, described in the bulleted list below. All are derived from (at least in part) fieldwork done in preparation for publication of a geologic map in 1969. As products are subsequently produced and time elapses, the uncertainty of what constitutes the source information tends to increase.

1. In 1969, a geologist published a map. It was assigned to a geologic map series, given a series number, printed, and cataloged and archived by the Library.
2. In 2009, someone decides to digitize the map because they need it in GIS format, in order to support new mapping and research. Because they want to share it with others, they post it to their Web site, or they may simply ftp it to colleagues who request it. If the GIS file has been finalized (i.e., it's not a preliminary version), the preparer, following conventional practice, indicates in the metadata that the map published in 1969 is the source. But because the GIS file isn't published, it isn't accorded the status of a publication and is not so managed and archived (i.e., (1) the Library doesn't catalog it; (2) the agency's publications staff don't ensure that the URL and filename, and the Web page, are as informative as that expected for a publication; and (3) the information-management staff cannot be expected to archive it as part of their systematic backups).
3. In 2010, someone finds the GIS file online, and downloads it. They compile and interpret it along with other information, and formally publish a new geologic map. In that product, the 2009 GIS file is cited, since it (and not the 1969 map) was directly used. At the time, the authors of the published map decide this is

sufficient because the 2009 GIS file's metadata indicates that the 1969 map is the source.

4. In 2015, someone reads the 2010 map or paper, and wants to use it for some new research or application. Let's assume that this person tries to find the source information, in order to get a better feel for the geologic interpretations and for the quality and density of observations and other data. However, the GIS file, posted to the Web in 2009, is no longer available. There is no agency record of it because it wasn't a publication, and so this person must rely on the information provided in the 2010 publication. In terms of the geology of the region in question, the 2010 publication must now be considered the authoritative source. With additional study, the true source (the 1969 map) could be inferred, and the field notes and supporting material perhaps accessed, but this process can be greatly facilitated by more detailed citations.

SUGGESTED CITATIONS

In 2004, I helped compile a national-scale map (Soller and Reheis, 2004; revised and republished as a GIS file in Soller and others, 2009) from a wide range of sources including published maps (paper and digital) and unpublished GIS files with various levels of documentation and completeness. The suggestions offered here are mostly based on the citations devised for that map publication. They tend toward the verbose, in order to be informative but also to suggest the types of information that could be helpful in citations that you might devise. In the interest of completeness, I include citations to published GIS files. The citation types are:

- Published map
- Published GIS file that refers to the source map
- Published GIS file that doesn't refer to the source map
- Unpublished GIS file, from which a map was later printed
- Unpublished GIS file, digitized from a printed map, and later incorporated into a published GIS file
- Unpublished GIS file, digitized from a printed map, with some modification and later incorporated into a published GIS file
- Unpublished GIS file, later superseded by published GIS file
- Unpublished GIS file derived from an online database
- Descriptive notation that an unpublished GIS version of map was used
- Descriptive notation regarding modifications to published map prior to digitization and use

Suggested citation for published map:

Soller, D.R., and Reheis, M.C., 2004, Surficial Materials in the conterminous United States: U.S. Geological Survey, Open-file Report 03-275, scale 1:5,000,000, <http://pubs.usgs.gov/of/2003/of03-275/>.

DRAFT -- To be published in DMT'10 Proceedings
(see <http://ngmdb.usgs.gov/Info/dmt/>)

Suggested citation for a published GIS file that refers to the source map (from Berquist and Soller, 2001):

Smith, A.B., and Digits, C.D., 2001, Geologic map of the XYZ Quadrangle, *adapted from* Doe and Smith 1999 map: The Geological Survey, Map D-31, one Adobe Acrobat (PDF) file, scale 1:24,000, available on CDROM or <URL, if any> [*adapted from* Doe, J.K., and Smith, A.B., 1999, Geologic map of the XYZ Quadrangle: The Geological Survey, Map M-123, scale 1:24,000].

Citation for a published GIS file that doesn't refer to the source map:

Green, G.N., 1992, The digital geologic map of Colorado in ArcInfo format: U.S. Geological Survey Open-File Report 92- 507, ArcInfo file, <http://pubs.usgs.gov/of/1992/ofr-92-0507/>. [Digitized from Tweto, Ogden, 1979, Geologic map of Colorado: U.S. Geological Survey, scale 1:500,000.]

Unpublished GIS file, from which a map was later printed. The extent to which the file was further modified by the authors, before printing the map, is unknown, as somewhat implied by the bracketed note:

Fullerton, D.S., Bush, C.A., and Pennell, J.N., unpublished, Surficial deposits and materials in the eastern and central United States (east of long 102° W.), *derived from* Quaternary Geologic Atlas of the United States: U.S. Geological Survey Geologic Investigations Series I-1420, one ArcInfo file, scale 1:2,500,000. [Printed map derived from this database is available as Fullerton, D.S., Bush, C.A., and Pennell, J.N., 2003, Map of surficial deposits and materials in the eastern and central United States (east of long 102° W.): U.S. Geological Survey Geologic Investigations Series I-2789, scale 1:2,500,000, <http://pubs.usgs.gov/imap/i-2789/>.]

Unpublished GIS file, digitized from a printed map, and later incorporated into a published GIS file. The bracketed note refers the user to the published GIS, with which it has some unspecified level of commonality:

Bedford, D.R., unpublished, Digital file showing geology of California, *digitized from* Jennings, C.W., Strand, R.G, and Rogers, T.H. comps., 1977, Geologic map of California: California Div. Mines and Geology Map GDM 2, scale 1:750,000. [Some information in this file is found in Bedford, D.R., Ludington, Steve, Nutt, C.M., Stone, P.A., Miller, D.M., Miller, R.J., Wagner, D.L., and Saucedo, G.J., 2003, Geologic database for digital geology of California, Nevada, and Utah—An application of the North American Data Model: U.S. Geological Survey Open-File Report 03-135, 35 p., <http://geopubs.wr.usgs.gov/open-file/of03-135/>.]

Unpublished GIS file, digitized from a printed map, with some modification (“ adapted from”) and later incorporated into a published GIS file. The bracketed note refers the user to the published GIS, with which it has some unspecified level of commonality:

DRAFT -- To be published in DMT'10 Proceedings
(see <http://ngmdb.usgs.gov/Info/dmt/>)

Bedford, D.R., unpublished, Digital file showing geology of Utah, *adapted from* Hintze, L.F., Willis, G.C., Laes, D.Y.M., Sprinkel, D.A., and Brown, K.D., 2000, Digital geologic map of Utah: Utah Geological Survey Map 179DM, CDROM. [Information about modifications of the published map are found in Bedford, D.R., Ludington, Steve, Nutt, C.M., Stone, P.A., Miller, D.M., Miller, R.J., Wagner, D.L., and Saucedo, G.J., 2003, Geologic database for digital geology of California, Nevada, and Utah—An application of the North American Data Model: USGS Open-File Report 03-135, 35 p., <http://geopubs.wr.usgs.gov/open-file/of03-135/>.]

Unpublished GIS file, later superseded by published GIS file (but not in time for published file to be used, so uncertain whether they are identical or different):

U.S. Geological Survey, unpublished, digital file showing geology of Oregon, digitized from Walker, G.W., and MacLeod, N.S., 1991, Geologic map of Oregon: U.S. Geological Survey, scale 1:500,000, <http://geology.wr.usgs.gov/docs/geologic/or/oregon.html>. [A digital version of this map was more recently published as Walker, G.W., MacLeod, N.S., Miller, R.J., Raines, G.L., and Connors, K.A., 2003, Spatial database for the geologic map of Oregon: U.S. Geological Survey Open-file Report 2003-67, scale 1:500,000, <http://pubs.usgs.gov/of/2003/of03-067/>.]

South Dakota South Dakota Geological Survey, unpublished, digital file showing geology of South Dakota. [Contains map information later published in Martin, J.E., Sawyer, J.F., Fahrenbach, M.D., Tomhave, D.W., and Schulz, L.D., 2004, Geologic map of South Dakota: South Dakota Geological Survey Map 10, 1:500,000.]

Unpublished GIS file derived from an online database. Citation briefly notes the nature of the database query:

Belohlavy, Francis, unpublished, Digital map of soil parent materials (interpreted as bedrock types, alluvium, etc.) assembled by querying STATSGO data (from U.S. Department of Agriculture Nat. Res. Conserv. Serv.): Conservation and Survey Division, Univ. of Nebraska - Lincoln, <http://www.dnr.ne.gov/databank/statsgo1.html>.

Descriptive notation that an unpublished GIS version of map was used. Extent to which the GIS file faithfully copies the printed map is unknown:

Clayton, Lee, 1980, Geologic map of North Dakota: U.S. Geological Survey Special Map prepared in cooperation with North Dakota Geological Survey, scale 1:500,000. [Used an unpublished, digital version of this map.]

DRAFT -- To be published in DMT'10 Proceedings
(see <http://ngmdb.usgs.gov/Info/dmt/>)

Descriptive notation regarding modifications to published map prior to digitization and use:

Barnes, V.E., ed., 1992, Geologic map of Texas: Texas Bureau of Economic Geology, scale 1:500,000. [Quaternary units on the map were generalized and then digitized.]

REFERENCES

Berquist, C.R., Jr., 1999, Digital Map Production and Publication by Geological Survey Organizations: A Proposal for Authorship and Citation Guidelines, in Soller, D.R. (ed.), Digital Mapping Techniques '99 -- Workshop Proceedings: U.S. Geological Survey Open-File Report 99-386, p.

Berquist, C.R., Jr., and Soller, D.R., 2001, Map Authorship and Citation Guidelines: Summary of a Discussion Session, in Soller, D.R. (ed.), Digital Mapping Techniques '01 -- Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p.

Richard, S.M., 2000, Proposal for Authorship and Citation Guidelines for Geologic Data Sets and Map Images in the Era of Digital Publication, in Soller, D.R. (ed.), Digital Mapping Techniques '00 -- Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, p.

Soller, D.R., and Reheis, M.C., compilers, 2004, Surficial materials in the conterminous United States: U.S. Geological Survey Open-File Report 03-275, scale 1:5,000,000.

Soller, D.R., Reheis, M.C., Garrity, C.P., and Van Sistine, D.R., 2009, Map database for surficial materials in the conterminous United States: U.S. Geological Survey Data Series 425, scale 1:5,000,000, <http://pubs.usgs.gov/ds/425/>.