USGS
science for a changing world



South Florida Information Access (SOFIA) Metadata for the U.S. Geological Survey Greater Everglades Priority Ecosystems Science

Beginning in 1995 the U.S. Geological Survey (USGS) funded scientific research to support the restoration of the Greater Everglades area and to supply decision makers and resource mangers with sound data on which to base their actions. However, none of the research and resulting data is useful if it can't be discovered, can't be assessed for utility in an application, can't be accessed, or is in an undetermined format. The decision was made early in the USGS Place-Based Studies (PBS) program to create a "one-stop" entry for information and data about USGS research results. To facilitate the discovery process some mechanism was needed to allow standardized queries about data. The FGDC metadata standard has been used to document the South Florida PBS data from the beginning.

A survey of project chiefs in 1996 showed that most felt metadata are important to their work and in assessing other data for use. However, the great majority of them replied that they had no money or personnel allocated to creating the metadata. Thus the metadata portion of SOFIA was created and funded to aid the scientists in fulfilling the requirement for FGDC-compliant metadata.

When a new project is authorized, a metadata record is created. As the project progresses and data becomes available, the project record is updated and metadata is created for data posted on the Data Exchange pages. The software currently used to create most of the metadata is Spatial Metadata Management System* (SMMS) from Intergraph (figure 1). The version in use has an Access Database to store common elements such as contact, and citation information, the metadata standard identification information, and other information used in all of the records. The newly created or revised metadata record is sent to the project chief or data producer for review and comments. Changes requested by the reviewer are made and the record is finalized. It is then exported from SMMS in ASCII text format for input to the USGS metadata parser program (mp). The record is checked for errors using mp. When all errors are remedied, the record is again run through mp and output in several different formats.

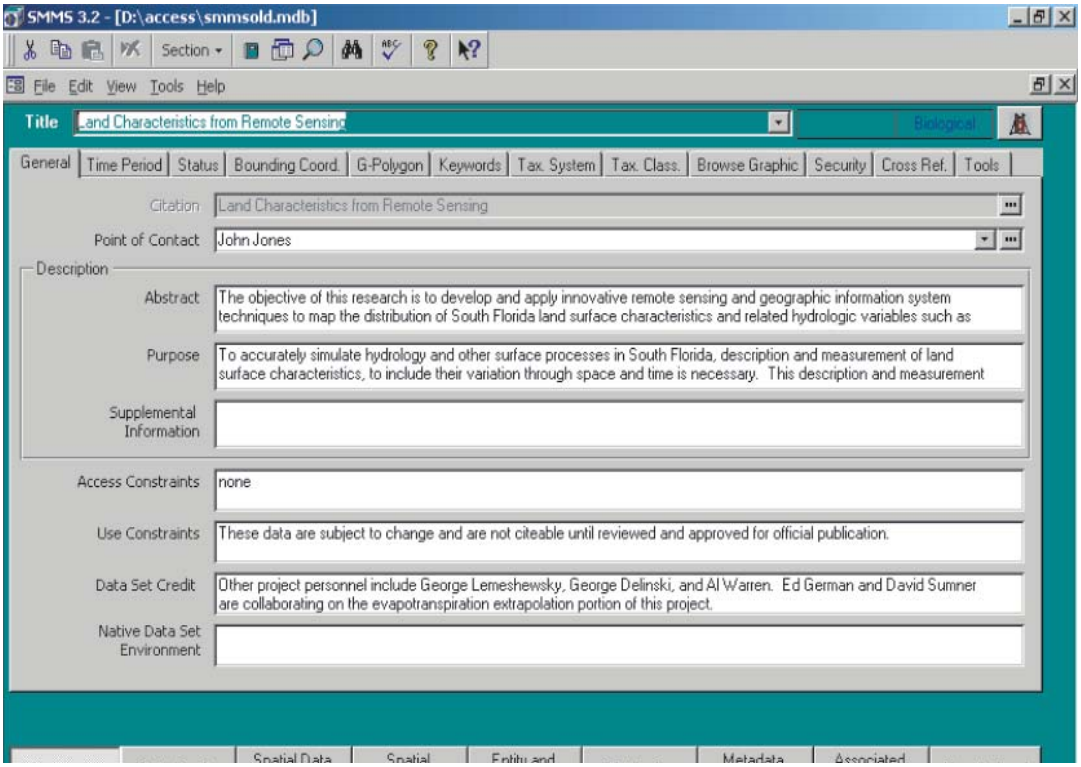


Figure 1: SMMS input screen

A second method of creating metadata records uses Extensible Markup Language or XML (figure 2). The USGS has developed an application, XMLInput, that was initially for use by contractors, cooperators, and partners. The application uses XML-based templates that can be filled in. An advantage is that the project chief or data

provider need only fill in certain information instead of having to create a complete FGDC metadata record. The information in the template can be merged with existing information to complete the full metadata record.

A third method of metadata creation is through ArcCatalog from ESRI and is used for data sets (figure 3). ArcCatalog extracts certain metadata elements from the information stored with e-data file(s) and places it in the appropriate FGDC metadata elements.

At present all FY 2004 funded projects have current metadata and over 100 data sets have been documented.

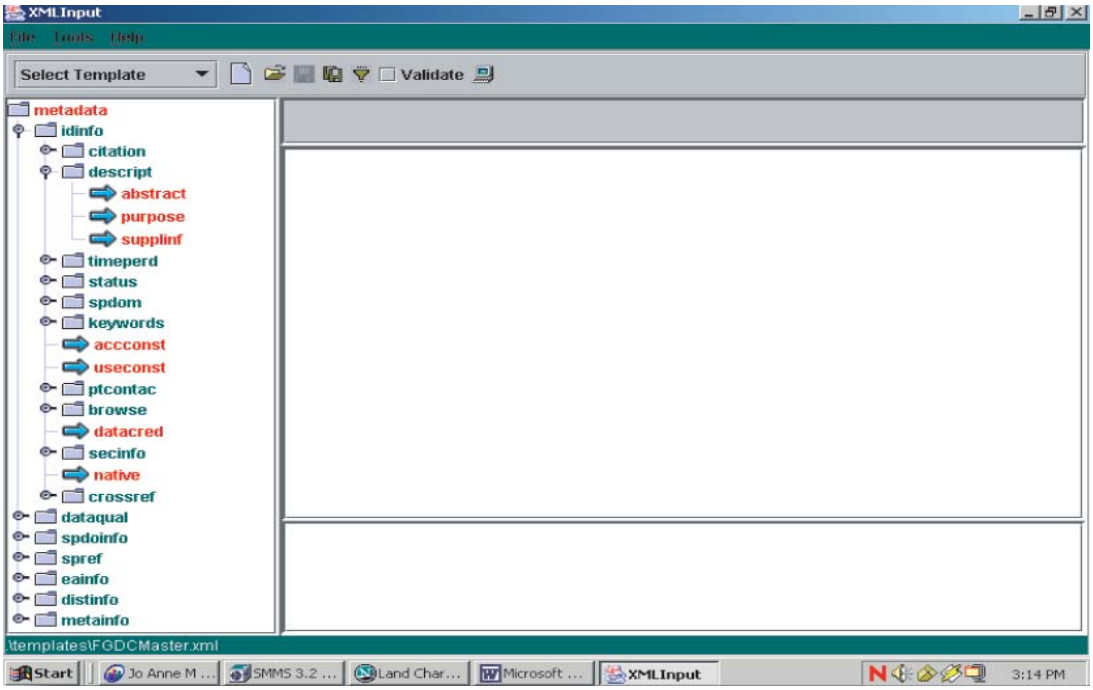


Figure 2: XML input template screen

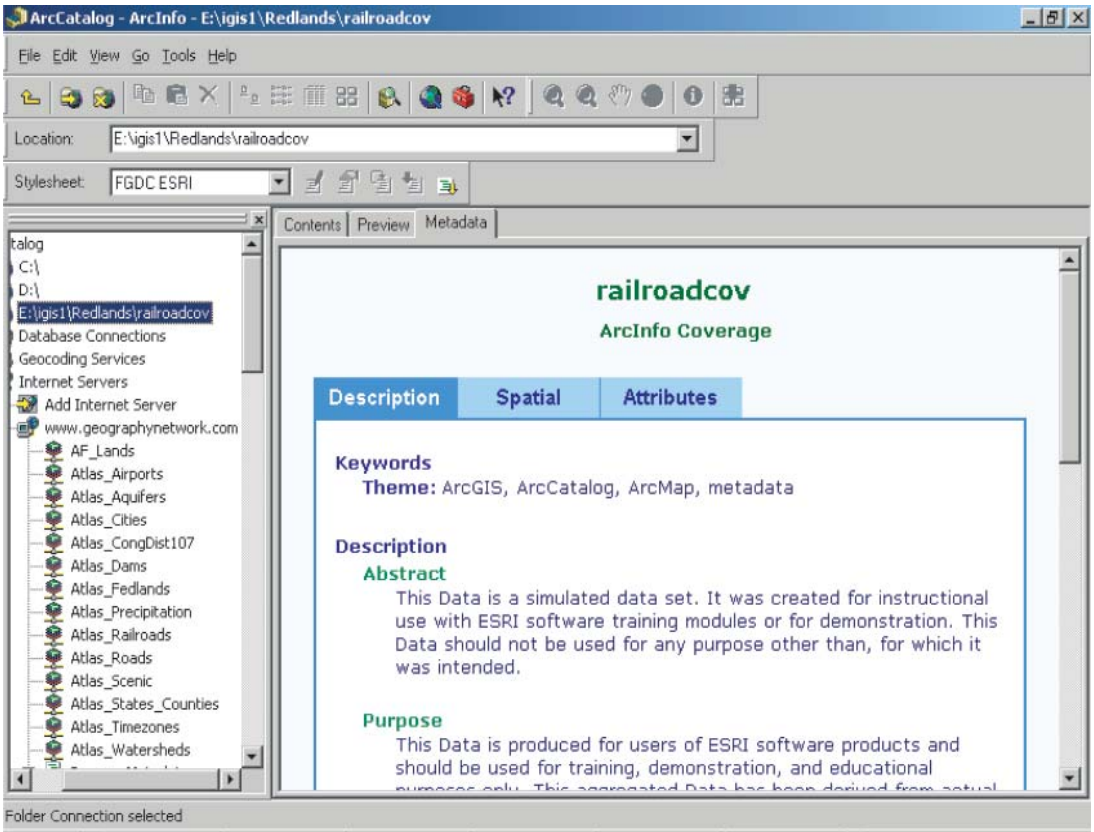


Figure 3: ArcCatalog screen

The formats used for SOFIA metadata records are html, text, and xml. A variation on the html format is called Question and Answer format. This format extracts information from the FGDC metadata format and answers a series of questions. This format is also referred to as "metadata in plain language".

Examples of the various formats are shown in figures 4 through 7.

MP, metadata parser program, has the ability through the use of a configuration file to link the various formats so that a user can go from one to another using the links at the top of the record.

Some points to remember to make the metadata more useful.

1. There are only two required sections in the FGDC minimal metadata record; however information in the other five sections make the data more useful.

2. Much of the information used to populate the metadata record exists in the form of yearly work plans, progress summaries, papers, posters, and other publications.

3. Metadata is not really difficult and there are SOFIA project personnel available to create the metadata and to make any necessary changes.

4. Good metadata makes your data useful to others even if you move on to other projects or organizations.

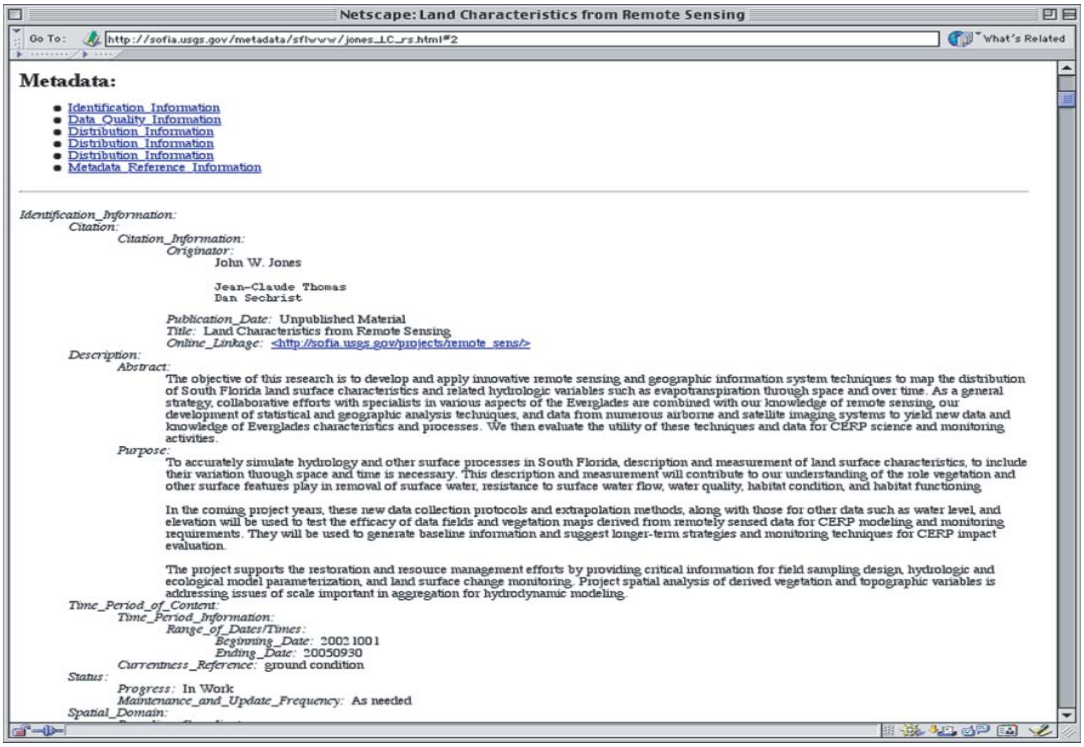


Figure 4: FGDC html format

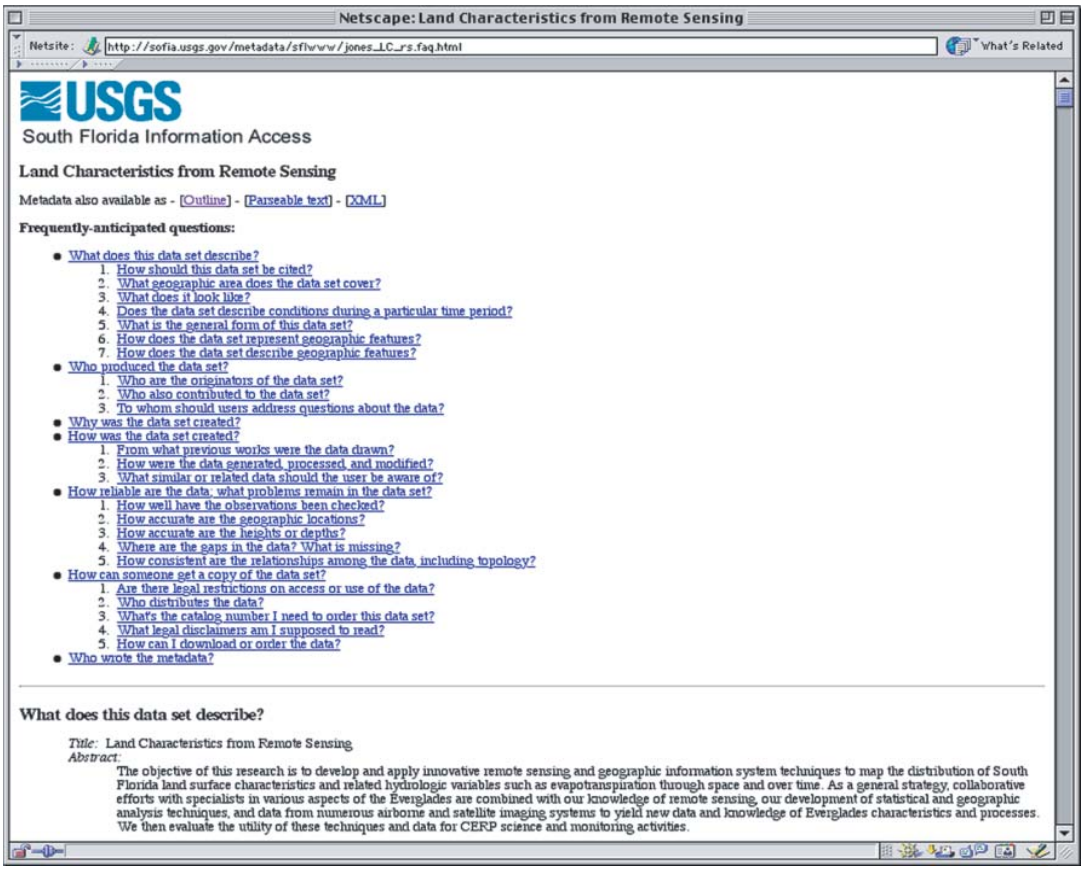


Figure 5: Question and Answer format

Information

Information on XTME, TKME, cns, and mp can be found at geology.usgs.gov/tools/metadata/.

XMLInput software is available for download at ftp://ftpext.usgs.gov/pub/cr/mo/rolla/release/xmlinput/

An explanation of Metadata in plain language can be found at geology.usgs.gov/tools/metadata/tools/doc/ctc/.

The FGDC metadata standard is available at www.fgdc.gov/metadata/contstan.html.

Information on SMMS is available at www.intergraph.com/gis/smms/.

SOFIA wants the metadata for your data sofia.usgs.gov/metadata/

For further information about metadata, contact:

Jo Anne Stapleton
jastapleton@usgs.gov
703 648-4592 voice
703 648-4614 fax

Roy Sonenshein
sunshine@usgs.gov
305 717 5824 voice
305 717-5801 fax

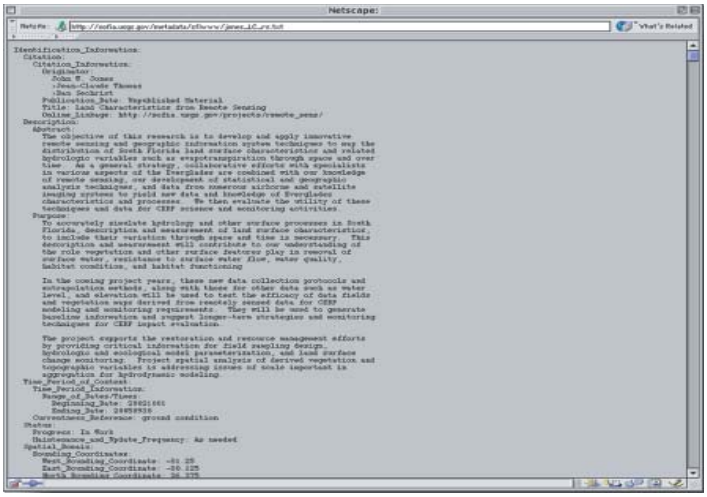


Figure 6: ASCII text format

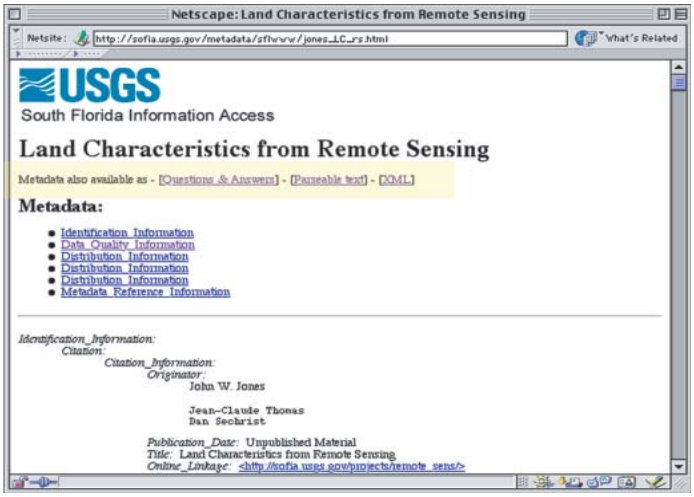


Figure 8: Example of the top of a metadata record showing the links (highlighted in yellow)

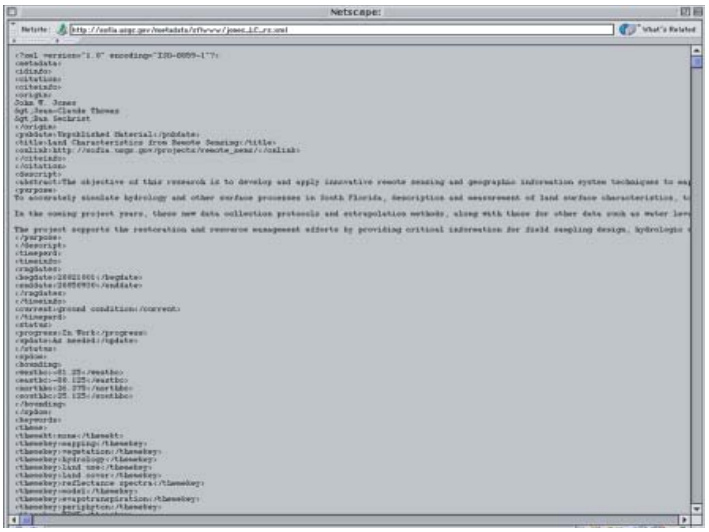


Figure 7: XML format

For information on other USGS products and services, call 1-888-ASK-USGS, or visit the USGS Geography Products Web site at erg.usgs.gov/isb/pubs/pubslists/.

For additional information, visit the ask.usgs.gov Web site or the USGS home page at www.usgs.gov.

** Any use of trade, product, or firm name is for descriptive purposes only and does not imply endorsement by the U.S. Government.*

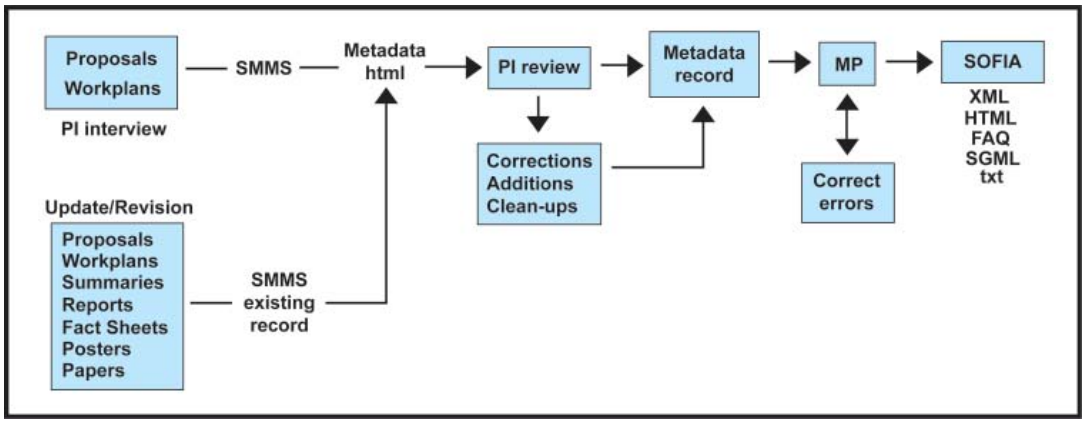


Figure 8: Diagram of metadata record creation and revision.
(The process is the same for revising metadata starting with the existing metadata record.)