The Quest For Metadata At The City Of Orem, Utah

By Steven R. Johnson, Engineering Specialist

Metadata: What is it?

In the GIS community there has been a great deal of interest in metadata, but just what is metadata. Metadata describes the content, quality, condition and other characteristics of data. It is used to describe how the geospatial data and attribute data was collected and processed into its final form. Put simply, metadata is data about the data.

Some examples are beneficial in understanding what metadata is. Let's take a family portrait for example. One of the participants in the photo shoot realizes that in sharing the photo with others, it would be useful to write some information on the back of the photo. They choose to include the names of those who appear in the picture, the date and location where it was taken, the name and the phone number of the photographer and if they are really detailed, the type of camera used and the film type. All of these bits of information could prove useful at a future date.

Some other examples of metadata would include nutrition labels and library card catalogs. Nutrition labels are very similar to Federal Geographic Data Committee (FGDC) standards for geospatial data. These labels are presented in a standard format. A consumer knows exactly what to expect on a nutrition label and he knows where to find it. The same should be true with metadata as well.

Why did the City want Metadata?

Metadata was seen as a way to identify existing data. This saves the City of Orem time and effort in many ways. First it allows GIS users to locate existing geospatial data. Knowing what data is available already minimizes data duplication. It also eases the sharing of information between people and departments in the City Of Orem as well as with the public. Accurate metadata is also useful to determine whether or not the existing data is suitable for the task at hand. It also makes employee turnover much less disruptive.

Metadata is helpful with all these items because complete metadata answers many common questions. Why were these data collected? Who collected these data? Who maintains and distributes these data? What is the accuracy of the data collected? Where are the data located? What do the data describe? Which data need to be updated?

The Beginning of the Metadata Collection Effort

Metadata was first mentioned to the City Of Orem in 2000. I was invited to participate in a FGDC Metadata Training Workshop – Don't Duck Metadata on April 12, 2000 at Southern Utah University in Cedar City, Utah. During the one-day course, I learned of the FGDC Standard and metadata creation tools used by the Utah Automated Geographic Reference Center, the National Park Service, the US Forest Service, and the Bureau of Land Management. I took this information and began working on metadata for the shapefiles that I was in charge of at the City Of Orem.

Though it was mentioned in GIS Committee meetings, there was no real progress made toward establishing metadata citywide. In September 2001, a major presentation was given on the benefits of metadata in a GIS Committee meeting. In the meeting, all the participants were shown how to use ArcCatalog's metadata capabilities. At that point, three of the City's major GIS contributors including myself authored essentially complete metadata for shapefiles related to buildings, cemetery and storm sewer infrastructure. This was only a small portion of the City's spatial data holdings. The complete set of City holdings consisted of aerial photography, over 200 shapefiles and nearly 700 CAD drawings. The creation of metadata once again stalled for nearly a year.

Why did metadata creation stagnate? I believe it can be attributed to three issues. First, spatial data creators felt that they didn't have any time. Metadata creation was looked at as something that could be useful for the future, but for now, it would just be an additional burden. Second, there was no plan. No one had decided what the priority was for the creation of each metadata record. And third, there was no metadata specialist to act as a resource to answer questions from metadata creators. This metadata specialist also needed to be the person who kept a vision of the utility of the metadata and to act as the driving force for keeping the metadata creation plan moving forward.

Becoming a Metadata Specialist

The mission of a metadata specialist can be shown through three roles: a visionary, a cheerleader and a resource. As a visionary a metadata specialist keeps the value of metadata in perspective. The metadata specialist also should be aware of the plan and the priority of each metadata record. As a cheerleader, the metadata specialist should encourage. And as a resource, the metadata specialist should always be available to answer questions. The specialist should be well informed on the FGDC standard and should have solutions to common problems. In addition he/she should know the City Of Orem's policies regarding metadata.

The first step to becoming a metadata specialist was to educate myself. First I reviewed the materials from the training I attended in April 2000. I also decided to take the online course

offered through ESRI Virtual Campus: Protecting Your Investment in Data With Metadata. I found this course to be extremely helpful. It was easy to understand and referred to a great number of issues that arise as you try to comprehend the FGDC standard as well as the process involved in producing metadata. The course covers the following topics:

Metadata: What's The Big Deal?

Standards for Documenting Geospatial Data

Managing Metadata with ArcCatalog

Documenting Your Data

Data Properties Automatically Captures by ArcCatalog

How to Make Metadata Work For Your Organization

I found several of the topics discussed especially useful. Some of these topics included: how to use the FGDC Content Standard for Digital Geospatial Metadata (CSDGM); an introduction to ArcCatalog's metadata functionality; style sheets and editors and how they make your life easier; FGDC CSDGM field definitions; data properties automatically captured by ArcCatalog; planning for the documentation of legacy and future data products; who should write metadata; and good metadata practices.

After getting some good training, the next step to becoming a metadata specialist was digging in and writing some. I reviewed the metadata I had previously created. With the additional training I received, I was able to fill in a few fields I didn't previously understand. I also completed metadata for several more shapefiles that I was responsible for.

Prioritizing Metadata Creation

With a good understanding of how metadata works, it was time to set up a plan for creating metadata for existing data in the City Of Orem. Going under the assumption that metadata has the most value when associated with the data used by the most departments as well as the data most requested by the public.

A survey was circulated among GIS users and creators to determine those datasets that needed metadata first. The survey asked, "Which categories of spatial data do you use most often in your work and to satisfy requests from the public?" Respondents then were asked to give each category one of the following ratings: 0 = Never; 1 = Almost Never; 2 = Sometimes; 3 = Often; 4 = Nearly All The Time. Here's a sample of what the survey looked like:

Buildings   0   1   2   3   4

Cabling	0	1	2	3	4	
Canals		0	1	2	3	4
Cemetery		0	1	2	3	4
Census		0	1	2	3	4
Communities		0	1	2	3	4
Contours		0	1	2	3	4

The ten highest rated datasets were: Aerial Images; Parcels; Buildings; Major Street Names; Streets (Asphalt Extent); Water; Contours; Geocoded Road Centerlines; Land Use; Zoning. Metadata

Creating and Editing Metadata

Who Should Collect Metadata?

Those who are closest to the data should create metadata. The creator or compiler of the data of the will be asked to create the metadata. As a metadata specialist, I would not be able to fill all the metadata fields without an extensive interview. It would take just as much time for the data creator if I fill out the information as if they were to type it themselves.

A common concern preventing people from trying to create metadata was the complaint that there wasn't enough time. In an attempt to minimize the time required for creating metadata, I reviewed several tools designed to help. These tools included ArcCatalog, the FGDC Editor, and three ArcScripts; the Three Tab Editor, the Batch Metadata Tool and Meaux Data. All of these tools are helpful in metadata creation. Each simplifies or expedites some portion of the metadata creation process.

ArcCatalog as a Tool for Creating and Editing Metadata

ArcCatalog creates, stores and updates many attributes of metadata without any user effort. Here are some of the attributes that ArcCatalog stores: spatial reference information (including point and vector object information and raster object information); horizontal and vertical coordinate systems; and entity types and attributes

FGDC Editor included with ArcCatalog

The FGDC Editor is the default metadata creation and editing tool in ArcCatalog. It is complete

and provides access to all elements of the FGDC metadata standard. It also can be cumbersome and repetitive to use. For example, it is common for the creator, the data maintainer and the metadata reference person to be the same person. With the FGDC Editor, the only way to populate these fields is to enter the same information three times or import metadata from another file. Importing from another file can be problematic because when you import the metadata, you import all the metadata. For example, if you wanted to import the creator from another file, you get the abstract field as well. Navigation through the help files is also difficult. But in the end, it is the most complete tool available and allows access to fields that other tools don't.

Three Tab Editor

The Three Tab Editor, available on the ESRI website under "ArcScripts", condenses data entry into three forms. The fields included deal with data description (abstract, purpose, keywords, constraints, etc.), source information and contact information.

It simplifies data entry for the most commonly repeated items. For example, you only have to type your contact information once. It is also very easy to install. To start you simply set it as your default metadata editor.

It appears to be limited in utility. For example it assumes the same person will be listed as the creator, the contact person and the metadata preparer. Also, I could only enter one keyword per data set. It comes with default values that are meaningful to the script's author. Setting up your own personal defaults is not particularly intuitive. These things aside, the Three Tab Editor is a useful tool for "quick and dirty" data entry.

Batch Metadata Tool

The batch metadata tool is also available on the ESRI website under "ArcScripts". It reads a metadata template and then writes that template to several selected data sets. It replaces the built in import function in ArcCatalog. It is very effective if you have one large data set that is split into several smaller sets for distribution or if you want to write the same metadata information to several files.

I had difficulty installing the tool, though the author was helpful in resolving the problems I had. It also appears to only work smoothly if you have Visual Basic 6 installed on your machine. It did not work on a machine that had only VBA, neither did it work for a machine that had VB.NET installed until I designed my own form and used some of the code I got from ArcScripts. It is limited by requiring you to export a metadata record as a text file for a template. It also does not read XML. It is however very good at writing the same metadata for several datasets at the same time.

Meaux Data

Of the three ArcScripts I tested, Meaux Data is the easiest and most useful. It installed easily and was easy to start. The author was extremely helpful and prompt in helping to fix a problem I had in installation.

Meaux Data manages default values for your metadata. The fields you can store defaults for include: keywords, abstract, purpose, constraints, metadata creator, department, address, and data status. It includes a keyword manager tab. Meaux Data stores defaults and keywords in an environment separate from the environment for editing individual metadata records. It allows you to select which values you want to store as defaults.

The metadata record-editing interface is simplified much like the Three Tab Editor. It splits its data entry forms into four categories: data description, data status, constraints and metadata details.

Meaux Data also has a utility for importing and exporting metadata. It is the most flexible of the ArcScript tools. It still cannot replace the FGDC Editor for editing items such domains and ranges in attribute table field definitions.

Finding Metadata

Once metadata has been created, there needs to be an easy way to find the metadata. ArcCatalog is a great tool for searching for specific items in metadata. With it you can search the contents of any metadata field. For example, I was able to quickly define a search for all the spatial data on the City's server that contain the word "storm" in the title field of the metadata.
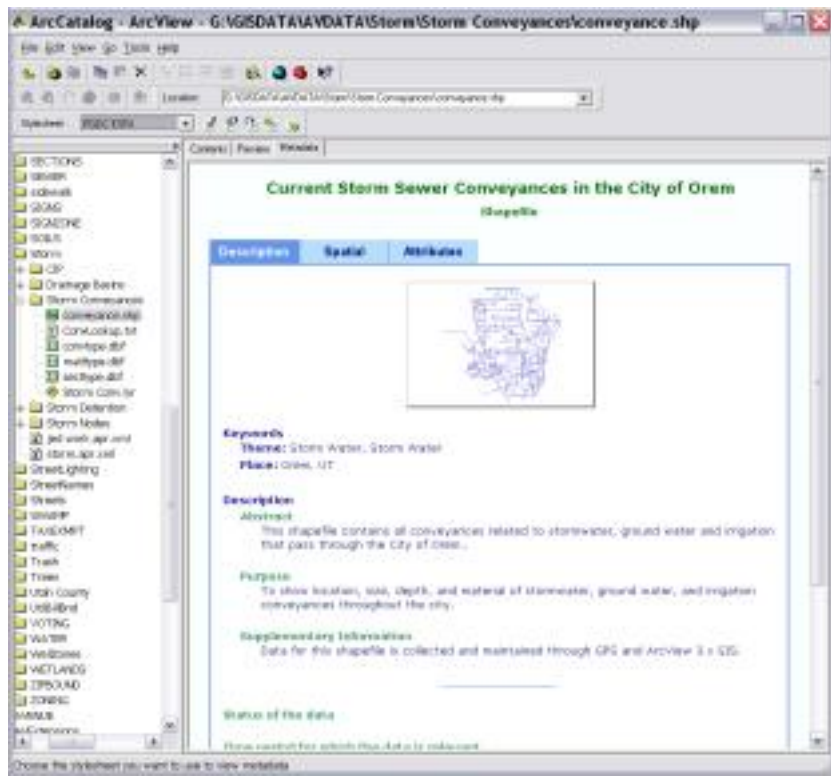
For someone who already knows how the data is stored on the server, ArcCatalog is quick and easy. Just navigate to the folder you want and click on the desired dataset. However, if you don't know the file structure on the server, ArcCatalog can be a slow utility for finding the right file. Depending on the size of you data holdings, a search can take several minutes to complete. There is no way to see all of your metadata holdings at one time.

To help find all metadata records faster, I created a webpage. I exported each metadata record into HTML. Then I created a homepage with links to all metadata records. This has become a useful alternative to searching for all metadata records in ArcCatalog.
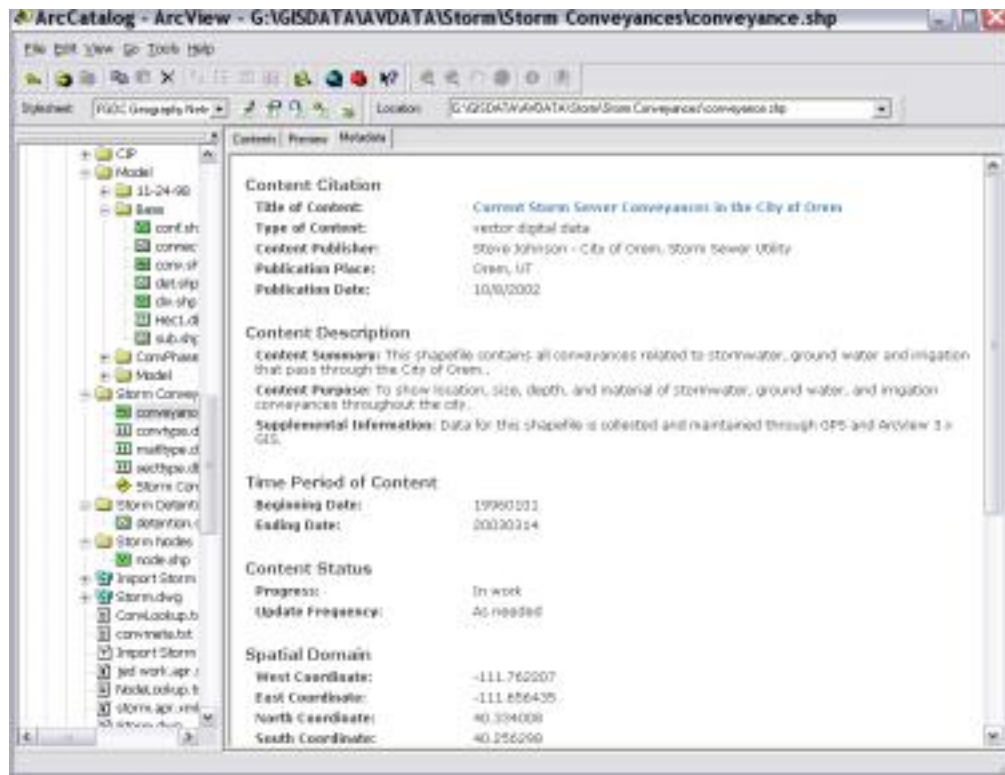
Viewing Metadata (Style Sheets)

ArcCatalog is the most useful tool for viewing individual metadata records. It makes use of items called style sheets. The style sheet you choose depends on what is important for you to see about each metadata record.

The most flexible style sheet is the *FGDC ESRI* style sheet. It is the default view for metadata in ArcCatalog. It is useful because it contains three tabs that let you focus on the data description, the spatial characteristics of the data and a listing of the fields for the attribute information. It includes a thumbnail image to quickly get a feel for what your data looks like. Clickable links allow you to expand information about fields of interest to you.
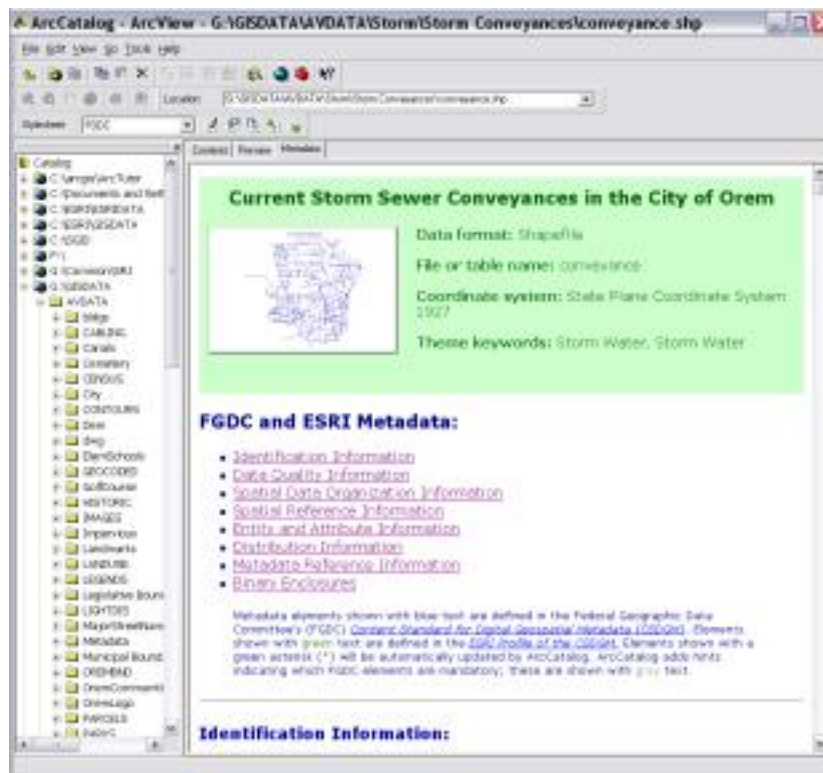
The *FGDC Geography Network* style sheet provides a quick view of the data description, spatial characteristics and constraint information. There is no information about attributes and no thumbnail is included.

The *FGDC Classic* style sheet it very complete, but is not very readable. It consists of a field name in italics followed by a field definition. A thorough knowledge of the FGDC standard is needed to be able to locate all the pertinent information.
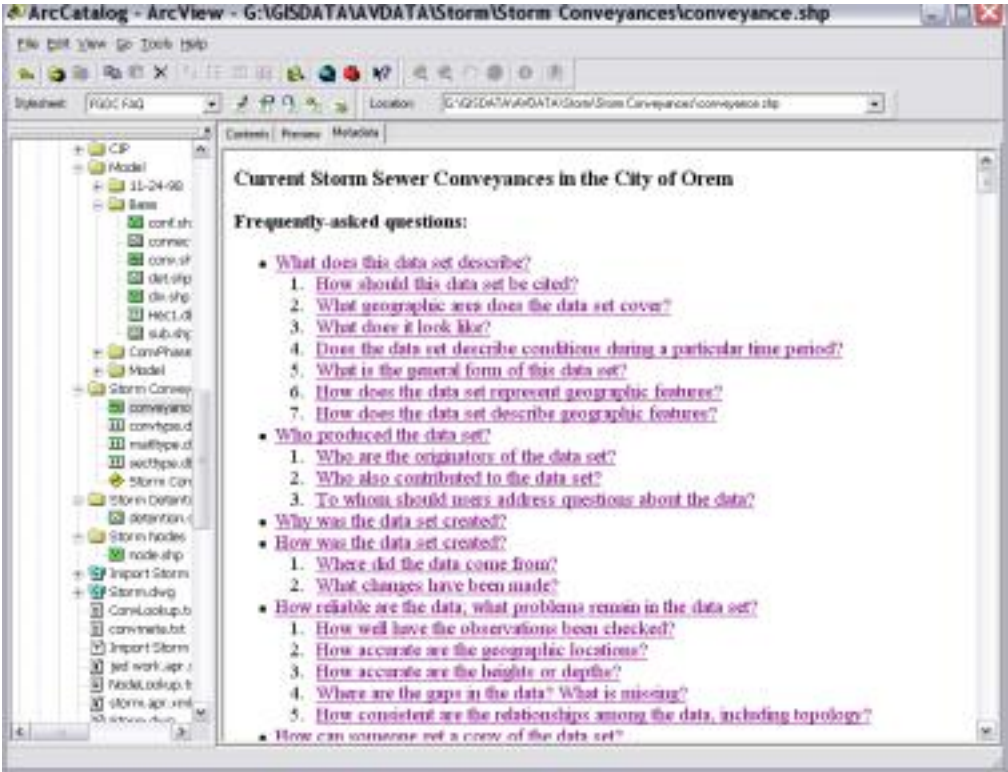
The *FGDC* style sheet is an updated version of the *FGDC Classic* style sheet. It includes a thumbnail and a quick summary of the metadata in a box at the beginning of the record. Below this box, it is very similar to the *FGDC Classic* style sheet with a thorough knowledge of the FGDC standard to navigate easily.

The FGDC FAQ style sheet is an interesting alternative view. It turns the metadata into a series of answers to "frequently asked questions". Some of the questions include: What does this data set describe? Why was the data set created? How can someone get a copy of the data set? In

addition, the FGDC FAQ style sheet also includes the most useful way of summarizing domains of attribute table fields.

Status of Metadata now

So how are we doing? In the past year there has been much progress toward creating metadata. We have gone from just a few files with metadata to 43 files with metadata. Of these, 28 files are complete. This progress is encouraging and we look forward to continuing to create and maintain metadata at the City Of Orem.