



Best Practices & Standards: *Metadata*

Date of Adoption: February 19, 2009

Introduction:

The Montana Association of Geographic Information Professionals (MAGIP) is a non-profit professional association that works to foster technical cooperation and promote the development of sound policy and practices that will support the efficient and effective use of geographic information systems in Montana. As such, the Technical Committee recognizes that a set of formalized best practices will not only encourage the use of standards but will make it easier to share data and save GIS practitioners work and time.

A best practice can be defined as the most efficient and effective way of achieving a goal or objective, improving business results through repeatable procedures that have proven themselves over time for large numbers of people. For a best practice to be effective, it must allow for the easy transfer of ideas, knowledge, and standards, while at the same time remaining flexible and adaptable to change as business needs and technology change.

It is important to note that the adoption of a best practice by MAGIP is not a mandate for its use, nor is a best practice applicable to every situation. Only in the instance where a best practice is recognized as a standard (Metadata in the Montana GIS Portal, for example) is it required and enforced. By committing to use a best practice, however, an individual acknowledges the benefits to one's organization and to the Montana GIS community as a whole.

Overview:

At its core, metadata is data about data, providing users with a formalized and practical description of a dataset's content, quality, condition and other characteristics. It represents the 'Who', 'What', 'When', 'Where', 'Why' and 'How' of the resource. In the context of GIS, however, metadata is also a critical component to sharing tools, data and maps. Much like a card in the library's card catalog describes a book to give readers information about its content, metadata provides information to GIS users to help determine if a geographic dataset is appropriate for their needs.

In Montana, the central location for sharing GIS data is the [GIS Portal](#). In that context, a metadata document is an XML file containing descriptive information about a GIS data set that also conforms to a set metadata standard. That standard allows the portal software to load it into a searchable index, provide basic descriptive information about the data set, and provide information that allows users to retrieve the data set.

With that in mind, MAGIP has adopted the [Metadata Standard for the Montana GIS Portal](#) as a best practice to help facilitate data sharing in the state of Montana.

Description:

The Metadata Standard for the Montana GIS Portal divides its requirements into different levels of compliance according to whether they are 1) mandatory technical requirements necessary for the correct operation of the portal software, 2) mandatory informational content to ensure minimal necessary information for the evaluation, retrieval, and use of the data set by portal users, and 3) recommended content that allows users to fully understand the data.

The following is a discussion of these requirements and the data elements that comprise each requirement type:

1) **Mandatory Technical Requirements**

- **FGDC Compliance**

Metadata documents submitted to the portal must be valid XML documents that conform to the structure set out in the Federal Geographic Data Committee's (FGDC) [Content Standard for Digital Geospatial Metadata \(CSDGM\), version 2](#).

- **GIS Portal Functional Requirements**

The following metadata elements are displayed by the portal when users find metadata records or are used by the portal to classify and provide links to the data. Documents without proper entries in these fields will not be accepted by the portal. In this and subsequent lists, the metadata element name is followed by the XML path to the element in the metadata file and a description of the required content.

- **Title** (Overview/Resource Identification/Citation/Title). Metadata must contain a title for the data set. The title should describe the region covered by the data set (such as "Montana", "Bozeman", or "Smith Watershed") and the subject (such as "Highways" or "Water Wells"). If the data set, or the particular copy of the data set referenced by the metadata, is not being actively maintained to be up-to-date, the title should include a date, for example "2006 Color Orthophoto of Helena, Montana" or "Yellowstone County School Districts, 1980".
- **Summary** (Overview/Item Description/Summary) is a concise summary of important information about the data set. In about three paragraphs or less, say everything you would want someone to know about the data, assuming this is all they were going to read.
- **Time Period of Content** (Overview/Citation/Dates) is a date, dates, or range of dates when the data set was valid. Dates must be entered in the metadata in YYYYMMDD format, such as 20080131. If you do not know the exact date, you may just fill out a year or year and month, as in 1995 or 199708. The FGDC standard allows you to fill this section with a publication date, but this is strongly discouraged, unless you have no information about when the data was really collected.
- **Contact Organization** (Resource/Points of Contact/Contacts) is used by the portal to fill out the Publisher seen by users in the metadata search results. The portal's advanced search page allows users to search for values from this field.
- **Point of Contact** (Resource/Points of Contact/Contacts) should be a person or office that can answer questions about the data set. The address, city, state, and ZIP code must be filled out.
- **Theme Keywords** (Overview/Topics&Keywords/Topic Categories) and (Overview/Topics&Keywords/New Theme Keywords). There must be a theme keyword section that contains a themekt element (Theme Keyword Thesaurus) whose value is "ISO 19115 Topic Category" and one or more themekey elements (Theme Keywords) whose values are chosen from the list in [Appendix A](#). The portal's advanced search page allows users to restrict their searches to records containing one of these keywords.
- **Place Keywords** (Overview/Topics&Keywords/New Place Keywords). The portal will show these in the Coverage Area section of the metadata results.
- **Online Linkage** (Resource/Distribution/Digital Transfer Options) is a link to a downloadable data file, a web site about the data, or a specification that allows the portal to add a web service to its map viewer. See [Appendix B](#) for the rules for this element.
- **Bounding Coordinates** (Resource/Extents/Geographic Extent/ Bounding Box) are the latitude and longitude coordinates of a rectangle that encloses the region covered by the data set. The portal's advanced search page allows users to search for data sets whose bounding coordinates intersect, or are within, any geographic region.
- **Resource Description** (Overview/Citation/Presentation Form) must be filled out with a choice from a pre-set list of values. Its value is shown in the Content Type section of the search results and it helps control how the portal creates a link to the data. The portal's advanced search page allows users to search for values from this field. See [Appendix B](#) for the rules for this element.
- **Global Unique ID** (Overview/Citation/Identifiers). For the portal to recognize that a revised document you publish is the same as a document you previously published, you must insert a Global Unique ID into the document. Instructions for this are in [Appendix C](#). For metadata file identifier see ((Metadata/Details/File Identifier).

2) **Mandatory Informational Content**

The portal managers will review your metadata and, if it is available, the data it describes. If they feel they do not understand it or if the following sections are not filled out properly, they may ask you to provide more information before they publish your metadata. In this and subsequent lists, the metadata element name is followed by the XML path to the element in the metadata file and a description of the required content.

- **Originator** (Overview/Citation/Contact/Role/Originator) or (Resource/Points of Contact/Role/Originator) is the agency or person primarily responsible for creating the data set. If, for example, this data was obtained from the Census Bureau and you made several corrections to the data, you would still probably list the Census Bureau as the originator. You may enter multiple originators if you feel this is necessary.
- **Publication Date** (Overview/Citation/Dates) may be the first time this data set was made available to the public or the date when this specific version of the data was released.
- **Publisher** (Resource/Points of Contact/Role/Publisher) or (Overview/Citation/Contact/Role/Publisher) is probably YOUR ORGANIZATION. Who is making this particular version of the data available to the public? If this version of the data is essentially unchanged from something you obtained from somewhere else (aside from easily accomplished format or projection changes), this should be the organization you obtained it from that is primarily responsible for making it public. If the organization does not make it public but provided it to you, with permission to distribute it, you might want to claim status as the publisher.
- **Spatial Reference Information** (Resource/Spatial Reference) is the definition of the coordinate system and datum of the data. If your metadata tool does not fill this section out automatically, refer to the instructions in the CSDGM workbook. Use the EPSG database to find the correct code for the spatial reference system you are using.
- **Entity and Attribute Information** (Resource/Fields/Entity and Attribute Information) is where you should put a list of the tables and data fields that come with your data, and explanations of what they contain. Fields that have a limited domain should have a list of the allowable values and, if applicable, their meanings. See [Appendix D](#) for some examples.
- **Distributor** (Overview/Citation/Contact/Role/Distributor) contains information on who to contact about getting a copy of the data.
- **Metadata Date** (Metadata/Citation/Resource Citation/ Dates) is the latest revision date of the metadata. Dates must be entered in the metadata in YYYYMMDD format, such as 20080131. If you do not know the exact date, you may just fill out the year or year and month, as in 1995 or 199708.
- **Metadata Contact** (Metadata/Contacts) is the person who wrote the metadata.

3) **Recommended Content**

You are strongly encouraged to fill out these fields. In some cases, the portal managers may ask you to fill some of these out if they feel an element from this list is especially important for your data set. In this and subsequent lists, the metadata element name is followed by the XML path to the element in the metadata file and a description of the required content.

- **Purpose** (idinfo/descript/purpose) describes a specific purpose the data was developed for, or something you want to see other people use it for. You don't have to imagine a purpose for the data if you don't want to put anything here.
- **Access Constraints** (Resource/Constraints/Legal Constraints). If there restrictions on who may obtain the data, or promises they have to make beforehand, put them here.
- **Use Constraints** (Resource/Constraints/General Constraints). Are there uses that the data should be restricted to or discouraged from?
- **Progress** (Resource/Details/Status) indicates whether the data set can be considered as a finished product. If the data is constantly being updated, you can still say it is "complete", as in having all the latest data you know about.
- **Update Frequency** (Resource/Maintenance/Update Frequency). How often is the data updated?
- **Browse Graphic** (Resource/Details/Browse Graphic). If you have some sort of on-line picture or map that features the data and put a link to it in this section, the portal has a place where it will show users a thumbnail image of the picture and/or a link to it.
- **Completeness Report** (Resource/Quality/Data Quality/Report). Is there some subset of the data that is missing? For example, "No data is available for Dawson County" or "Streams less than two miles long have been omitted".

- **Attribute Accuracy Report** (Resource/Quality/Data Quality/Report). If there is something you should say about the accuracy of any of the data's attribute fields, put it here. For example, "Each well ID number was independently checked against the source map by two technicians" or "Standard deviation for Calcium concentration in control samples was 8 mg/L".
- **Horizontal Positional Accuracy Report** (dataqual/posacc/horizpa/horizpar). Tell us what you know about the accuracy of the coordinates, such as "GPS coordinates were not differentially corrected, and the receiver reported accuracy values of between 10 meters and 40 meters during the survey."
- **Vertical Accuracy Report** (dataqual/vertacc/vertaccr). Describe the accuracy of any elevation information in the data set. "Elevations were estimated from a topographic map whose contour interval is 80 feet".
- **Source Information** (Resource/Lineage/Data Source). What documents or data sets did you obtain information for this data set from? See [Appendix E](#) for examples.
- **Process Steps** (Resource/Lineage/Process Step). What did you do to create the data? See [Appendix E](#) for examples.

Implementation:

1) Writing Metadata With Geospatial Metadata Tools

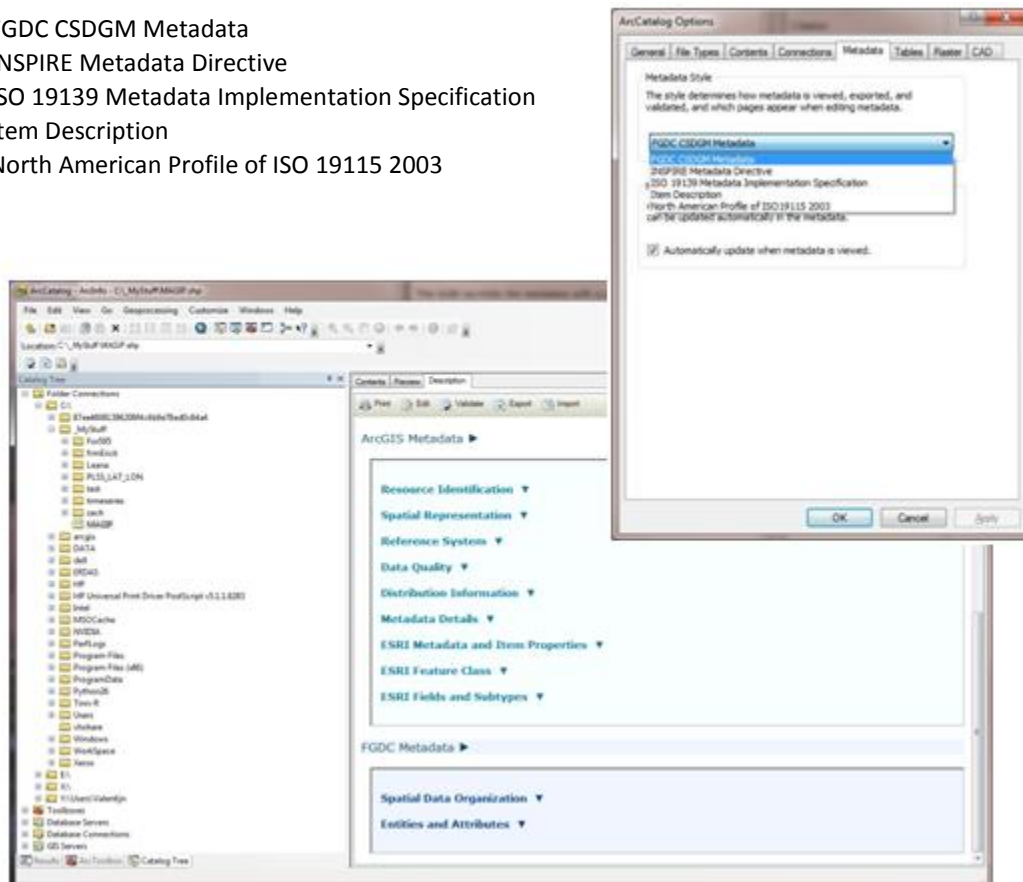
There are a variety of [free and commercial software tools](#) available to support metadata development. Each tool offers a range of features and capabilities that can be catered and specialized to your needs.

2) Writing Metadata With ESRI's ArcCatalog

The most common tool for writing metadata (and thus the focus of this document) is ESRI's ArcCatalog. Any item in ArcCatalog, including folders and file types such as Word documents, can have metadata. Once created, metadata is copied, moved and deleted along with the item when it is managed with ArcCatalog.

While in ArcCatalog, click on 'Customize' in the main menu. Then select the 'Metadata' tab and pick the desired style of Metadata documentation. The following style options are available:

- FGDC CSDGM Metadata
- INSPIRE Metadata Directive
- ISO 19139 Metadata Implementation Specification
- Item Description
- North American Profile of ISO 19115 2003

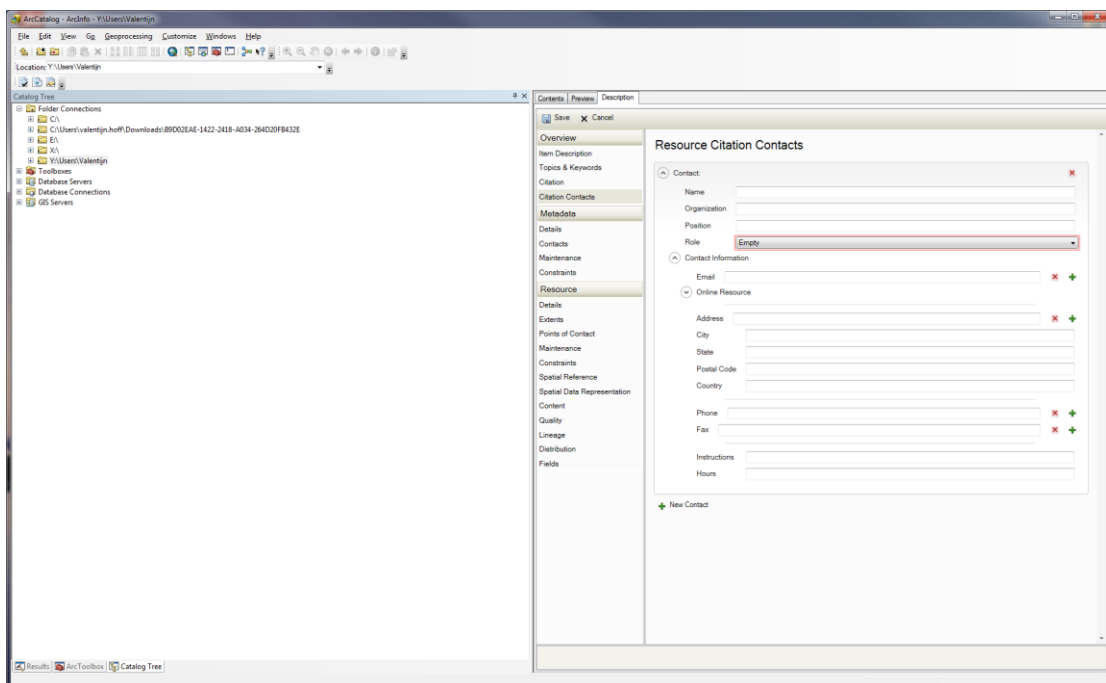


For the simplest documentation set the style to 'Item Description'. This option only stores text data such as summary, description and credits. It does not store geoprocessing history, spatial information or field properties, etcetera—so it is not a true metadata. When you choose any one of the other options, the metadata are displayed both as ArcGIS Metadata and as FGDC Metadata. Once you edit the metadata, it can be displayed in either format.

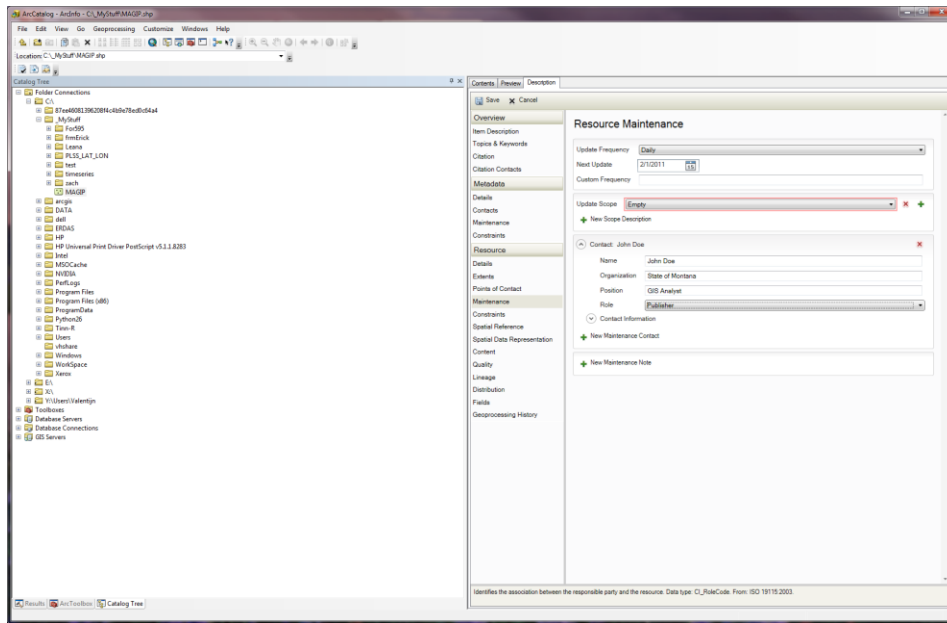
In order to edit the metadata, click on the 'Edit' button under the 'Description' tab in ArcCatalog. The left side of the Description pane has a list of items which are part of the complete metadata of a dataset, file or document. By clicking on each item you can edit it in the right side of the Description pane. Editing the metadata is a straight forward operation.

Properties, such as the extent of a feature class, are automatically derived from the item by ArcCatalog and added to the metadata. You may wish to change your settings so that ArcCatalog doesn't automatically create or update your metadata, which ultimately writes scores of notations about what ArcCatalog thinks is required in your metadata that you may then have to track down and fix.

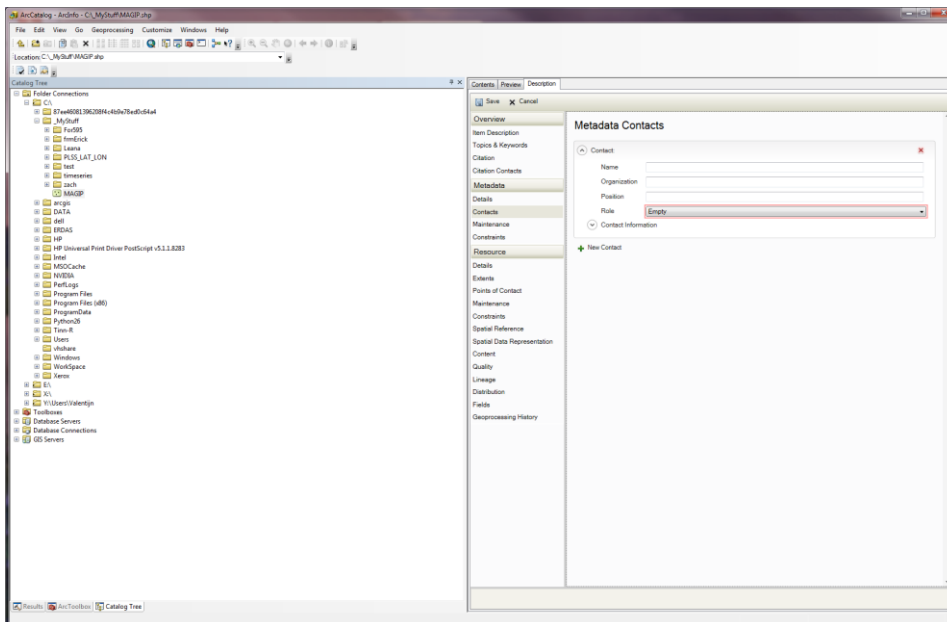
Metadata contains a lot of information about you and your organization. In many ways, it is much easier to make metadata for a layer if you can start by loading a metadata template that already has your contact information in it. To create the template, begin by selecting a layer that has no metadata and then click the Description Tab. Next, click the Edit button to open the editing pane. Information about you and your organization can be added to the metadata under Citation Contacts, Metadata contacts, Resource Points of Contact and Maintenance Contact. All have identical fields for personal information.



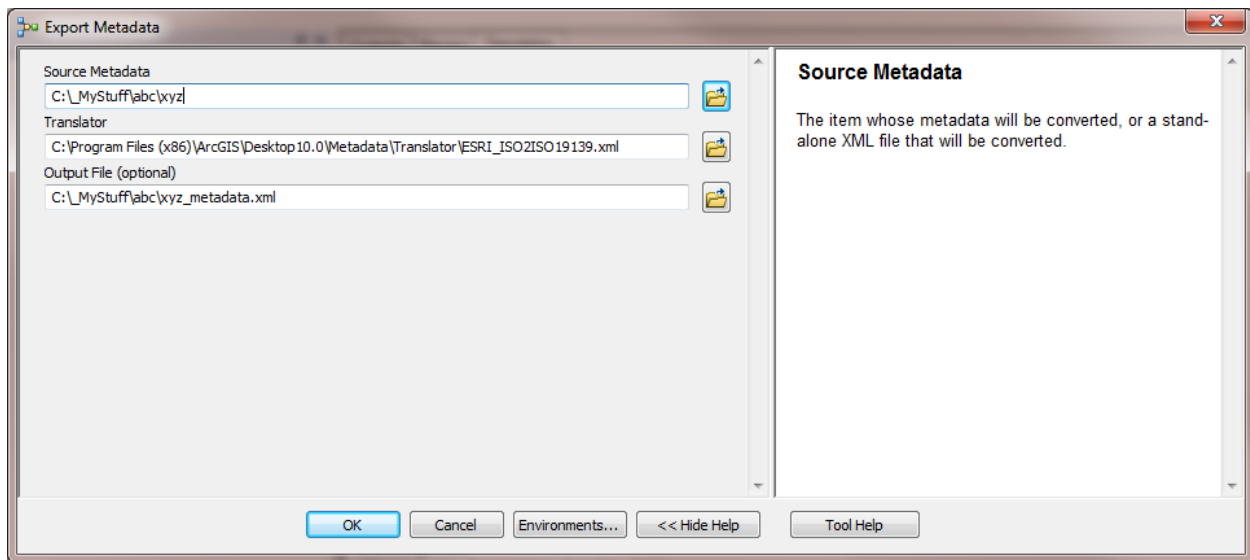
After filling out your Contact Information in these sections, the Constraint Section can then be used to insert your standard liability statement. For users who wish to publish their data to the Montana GIS Portal, it is important to fill out the Resource Description field. This field tells the Portal what type of resource is being submitted (valid types are given in [Appendix A](#)). You can put the value you will use for most of your metadata here—but if you have different types of resources available, this is an important field to think about for each document.



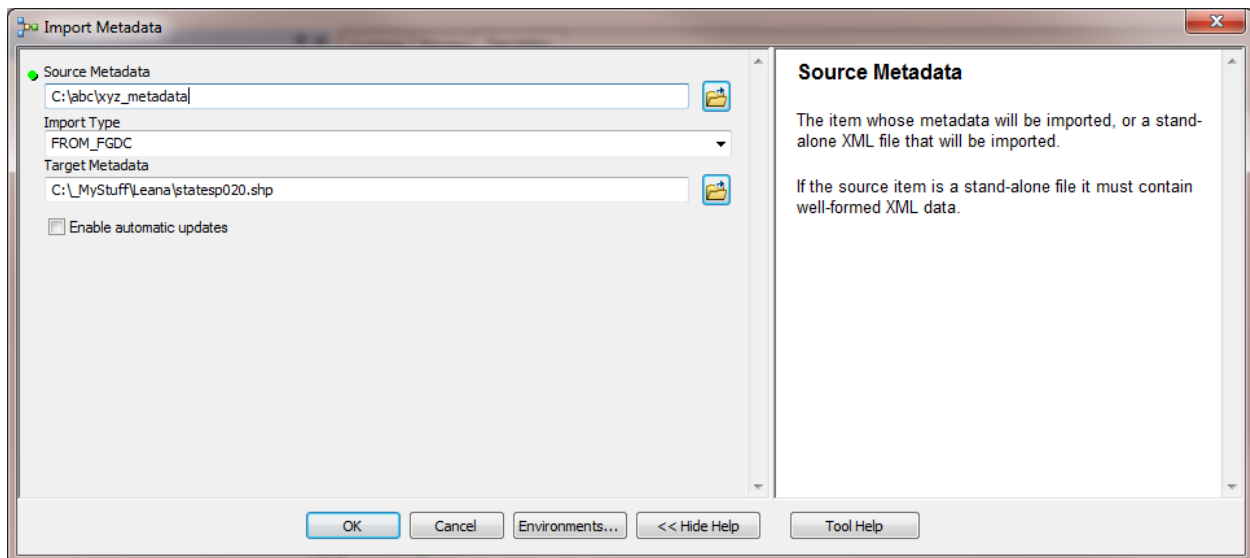
The Metadata Reference Section contains information about the individual who wrote the metadata and when it was created. Again, the Contact Section has the same format as the Point of Contact and the Distributor. Once this information has been entered, the metadata should then be saved.



Once the information has been saved, it should then be exported to FGDC XML format so it can be imported into future metadata files—simply click the Export Metadata button, browse to the location where you wish to save the file, and choose the FGDC compliant XML format.



For future projects, you are now able to import the FGDC compliant XML template you just saved into any layer for which you wish to write metadata. To import the template, select the item from the Catalog tree and click the Import Metadata button. It is probably a good idea to uncheck the box that enables automatic updates of your metadata, as that tends to cause problems with the title of your metadata document upon import.



After you have imported the template, ArcCatalog will still make entries in the metadata indicating what you are required to fill out, but there are a lot fewer to track down than if you hadn't used the template.

At this point, each of the required fields (discussed in the Description Section) must be manually populated with information about the data by using the Edit button under the Description tab in ArcCatalog. *The lists of required and recommended metadata content provided in this document all display the metadata element names in bold, followed by the XML path to each element in the metadata file (shown in parentheses).*

References:

[ESRI ArcGIS Desktop Help – About Metadata](#)

[Federal Geographic Data Committee](#)

[Federal Geographic Data Committee - Content Standard for Digital Geospatial Metadata \(Version 2.0\)](#)

[Montana GIS Portal](#)

[Montana GIS Portal Metadata Publisher's Guide](#)

[Montana GIS Portal Metadata Standard](#)

[Montana GIS Portal Metadata Writing Guide](#)

Document History:

February 19, 2009	Adopted by MAGIP Board of Directors
May 04, 2010	Documentation Standardized by MAGIP Technical Committee
April 26, 2011	ArcGIS 10 Update Added by MAGIP Technical Committee