



Metadata Wizard: An Easy-to-Use Tool for Creating FGDC– CSDGM Metadata for Geospatial Datasets in ESRI ArcGIS Desktop

By Drew A. Ignizio, Michael S. O'Donnell, and Colin B. Talbert

Open-File Report 2014–1132

**U.S. Department of the Interior
U.S. Geological Survey**

U.S. Department of the Interior
SALLY JEWELL, Secretary

U.S. Geological Survey
Suzette M. Kimball, Acting Director

U.S. Geological Survey, Reston, Virginia: 2014

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit <http://www.usgs.gov> or call 1-888-ASK-USGS

For an overview of USGS information products, including maps, imagery, and publications, visit <http://www.usgs.gov/pubprod>

To order this and other USGS information products, visit <http://store.usgs.gov>

Suggested citation:

Ignizio, D.A., O'Donnell, M.S., and Talbert, C.B., 2014, Metadata wizard—An easy-to-use tool for creating FGDC–CSDGM metadata for geospatial datasets in ESRI ArcDesktop: U.S. Geological Survey Open-File Report, 2014–1132, 14 p., <http://dx.doi.org/10.3133/ofr20141132>.

ISSN 2331-1258 (online)

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

Although this information product, for the most part, is in the public domain, it also may contain copyrighted materials as noted in the text. Permission to reproduce copyrighted items must be secured from the copyright owner.

Contents

Abstract	1
Introduction.....	1
Downloading and Installation.....	2
Using the Metadata Wizard	6
Parameter 1: “Input Dataset or XML Record”	7
Parameter 2: “Working Directory”	7
Parameter 3: “Create Stand-Alone FGDC–CSDGM XML Metadata Record” (Optional Toggle).....	8
Parameter 4: “Run Wizard Using Custom Starter Template” (Optional Toggle).....	8
Software Workflow/Processing Steps.....	8
Metadata Wizard Data Entry Graphical User Interfaces.....	10
Entity and Attribute Builder Graphical User Interface	10
Main Metadata Editor Graphical User Interface.....	11
Tool Updates and Maintenance.....	13
Data and Software Disclaimer	14
Acknowledgments	14
Reference Cited.....	14

Figures

Figure 1. Screen capture of the USGS ScienceBase page where the Metadata Wizard can be downloaded.....	3
Figure 2. Screen capture of the Metadata Wizard toolbox in the ArcCatalog ‘Catalog Tree.....	4
Figure 3. Screen capture of the Metadata Wizard toolbox in the ArcCatalog ‘Catalog Tree,’ showing description.....	4
Figure 4. Screen capture of the ‘Add Toolbox’ prompt from the ArcCatalog ArcToolbox pane.....	5
Figure 5. Screen capture showing the Metadata Wizard in the ArcToolbox pane	6
Figure 6. Screen capture of the Metadata Wizard tool dialogue box.....	7
Figure 7. Processing sequence and operational steps performed by the Metadata Wizard.....	9
Figure 8. Screen capture of the Metadata Wizard’s Entity and Attribute Builder GUI.....	10
Figure 9. Screen capture of the Metadata Wizard’s Main Metadata Editor GUI, tab 1	11
Figure 10. Screen capture of the Metadata Wizard’s Main Metadata Editor GUI, tab 2	12
Figure 11. Screen capture of the Metadata Wizard’s preview pane	13

Metadata Wizard: An Easy-to-Use Tool for Creating FGDC–CSDGM Metadata for Geospatial Datasets in ESRI ArcDesktop

By Drew A. Ignizio¹, Michael S. O'Donnell², and Colin B. Talbert²

Abstract

Creating compliant metadata for scientific data products is mandated for all federal Geographic Information Systems professionals and is a best practice for members of the geospatial data community. However, the complexity of the The Federal Geographic Data Committee's Content Standards for Digital Geospatial Metadata (FGDC–CSDGM), the limited availability of easy-to-use tools, and recent changes in the ESRI software environment continue to make metadata creation a challenge. Staff at the U.S. Geological Survey Fort Collins Science Center have developed a Python toolbox for ESRI ArcDesktop to facilitate a semi-automated workflow to create and update metadata records in ESRI's 10.x software. The U.S. Geological Survey Metadata Wizard tool automatically populates several metadata elements: the spatial reference, spatial extent, geospatial presentation format, vector feature count or raster column/row count, native system/processing environment, and the metadata creation date. Once the software auto-populates these elements, users can easily add attribute definitions and other relevant information in a simple Graphical User Interface. The tool, which offers a simple design free of esoteric metadata language, has the potential to save many government and non-government organizations a significant amount of time and costs by facilitating the development of FGDC–CSDGM compliant metadata for ESRI software users. A working version of the tool is now available for ESRI ArcDesktop, version 10.0, 10.1, and 10.2 (downloadable at <http://www.sciencebase.gov/metadatawizard>).

Introduction

Creating metadata to document finalized scientific data products is mandated for all federal Geographic Information Systems (GIS) professionals and is also a best practice for the greater geospatial data community (Executive Order 12906). The Federal Geographic Data Committee's Content Standards for Digital Geospatial Metadata (FGDC–CSDGM) is one of several well-defined metadata formats, and currently constitutes the most widely-used metadata standard among federal employees and GIS users. However, the complexity of the FGDC–CSDGM standard and the limited availability of easy-to-use tools continue to make metadata creation a challenge for many geospatial professionals. This has resulted in many scientific datasets being poorly or incompletely documented; in some cases, in spite of best practices or formal requirements, data producers simply forego metadata creation.

¹ Cherokee Services Group—contractor to U.S. Geological Survey, Fort Collins, Colo.

² U.S. Geological Survey

To address this challenge, the U.S. Geological Survey Fort Collins Science Center (FORT) has developed a Python toolbox for ESRI ArcGIS Desktop 10.x to facilitate a semi-automated process for creating and updating metadata records. The U.S. Geological Survey (USGS) Metadata Wizard tool streamlines metadata creation by simplifying the process of creating and editing FGDC–CSDGM metadata content when it is associated with spatial datasets (for example, geodatabase feature classes, Geotiffs, and shapefiles), stand-alone tabular data (for example, Dbase, ArcInfo, and geodatabase tables), or in the form of a stand-alone extensible markup language (XML) file—a common format for metadata records.

After a series of changes were introduced in the way that metadata is handled in the ESRI 10.x software, editing the FGDC–CSDGM metadata content associated with a dataset became more complicated, as did the process of exporting metadata content to well-formatted XML files compliant with the FGDC–CSDGM standard. By overcoming the technical difficulties associated with working with metadata in ESRI’s software, and developing a user interface designed to capture critical information while avoiding the technical language of the FGDC–CSDGM standard, the Metadata Wizard tool represents a valuable resource for metadata development and review.

The tool will automatically populate several metadata elements: the spatial reference, spatial extent, geospatial presentation format, vector feature count or raster column/row count, native system/processing environment, and the metadata creation date. Users are then provided with two Graphical User Interfaces (GUIs) to enter additional metadata information. The first GUI (the Entity and Attribute Builder) captures information for the entity/attribute section (that is, a description of the tabular data associated with the dataset, if applicable), while the second GUI (the Main Metadata Editor) allows a user to enter other general metadata information. With its simple interface, free of esoteric metadata language, the Metadata Wizard has the potential to save many government and non-government organizations a significant amount of time and costs, and can help facilitate the development of FGDC–CSDGM compliant metadata for all ESRI ArcGIS Desktop users.

The tool provides valuable functionality to expedite metadata creation by automatically retrieving information inherent to geospatial datasets, providing a semi-automated form to describe tabular field values, using default values, and providing for easy template importation. The Metadata Wizard also contains a built-in viewer to preview a cleanly formatted XML metadata record. The application can be easily installed on any standard version of ESRI ArcGIS Desktop 10.x and dovetails with other metadata efforts and existing workflows by allowing the easy export of XML metadata records.

Downloading and Installation

The complete Metadata Wizard file bundle, which contains all the necessary materials to install and use the tool on a standard version of ESRI ArcGIS Desktop, can be acquired as a bundled “zip” package from the following Uniform Resource Locator (URL): <http://www.sciencebase.gov/metadatawizard> . At this URL, site visitors can download the tool (fig. 1), read a description of how to install the toolbox on their local machine, view a general description of the tool’s operational workflow, and review a list of updates that have been made to the most recent version of the software. Three versions of the tool are available; each respective to a version of ESRI’s ArcGIS Desktop (10.0, 10.1, and 10.2.x).

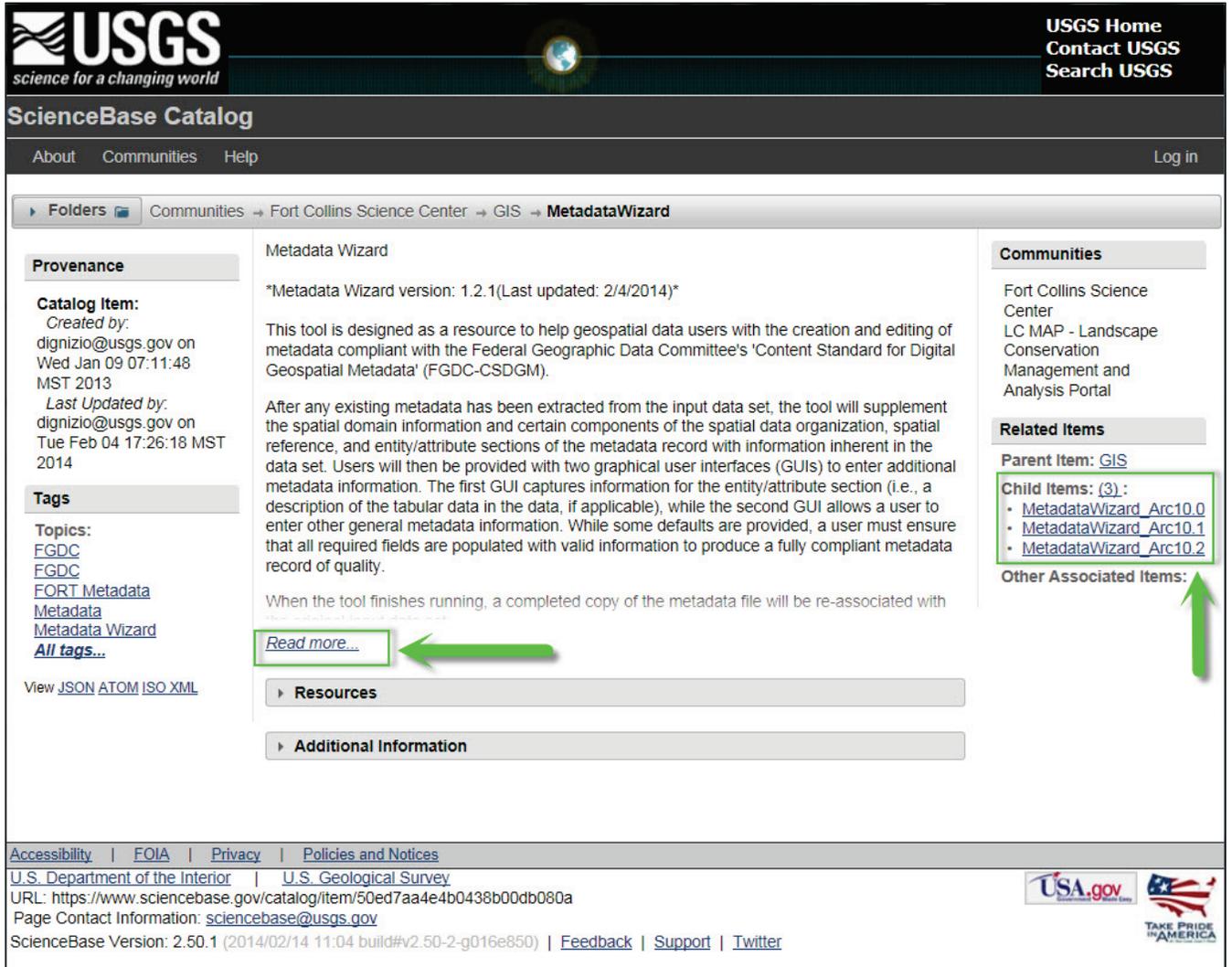


Figure 1. Screen capture of the USGS ScienceBase page where the Metadata Wizard can be downloaded.

To install the tool, users should download the appropriate versioned package for their ESRI installation and extract (“unzip”) the contents locally to their machine or to a shared network location. From within the ESRI ArcCatalog application, users have two options to access and use the tool. The first option is to use the ‘Catalog Tree’ (by default, this is the left-most pane within ArcCatalog) and navigate to the folder where the toolbox (USGS_MetadataTools.tbx) was extracted. The toolbox will be recognized by ArcCatalog and the contents of the downloaded, unzipped folder will be visible (fig. 2).

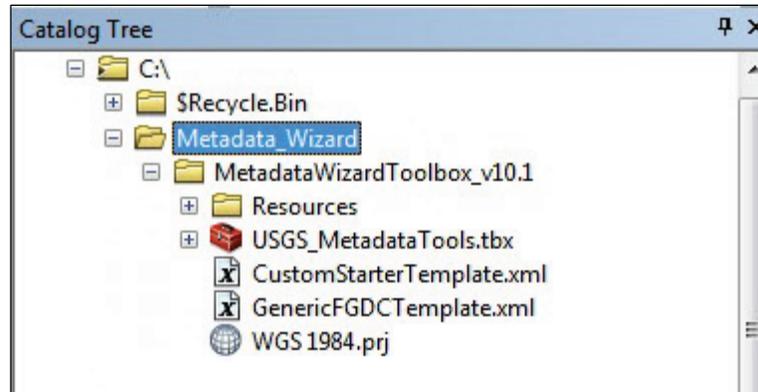


Figure 2. Screen capture of the Metadata Wizard toolbox in the ArcCatalog 'Catalog Tree.'

Accessing the tool from the Catalog Tree in this manner will also allow a user to view the ArcCatalog item description for the 'MetadataWizard' tool contained in the USGS_MetadataTools.tbx toolbox (fig. 3).

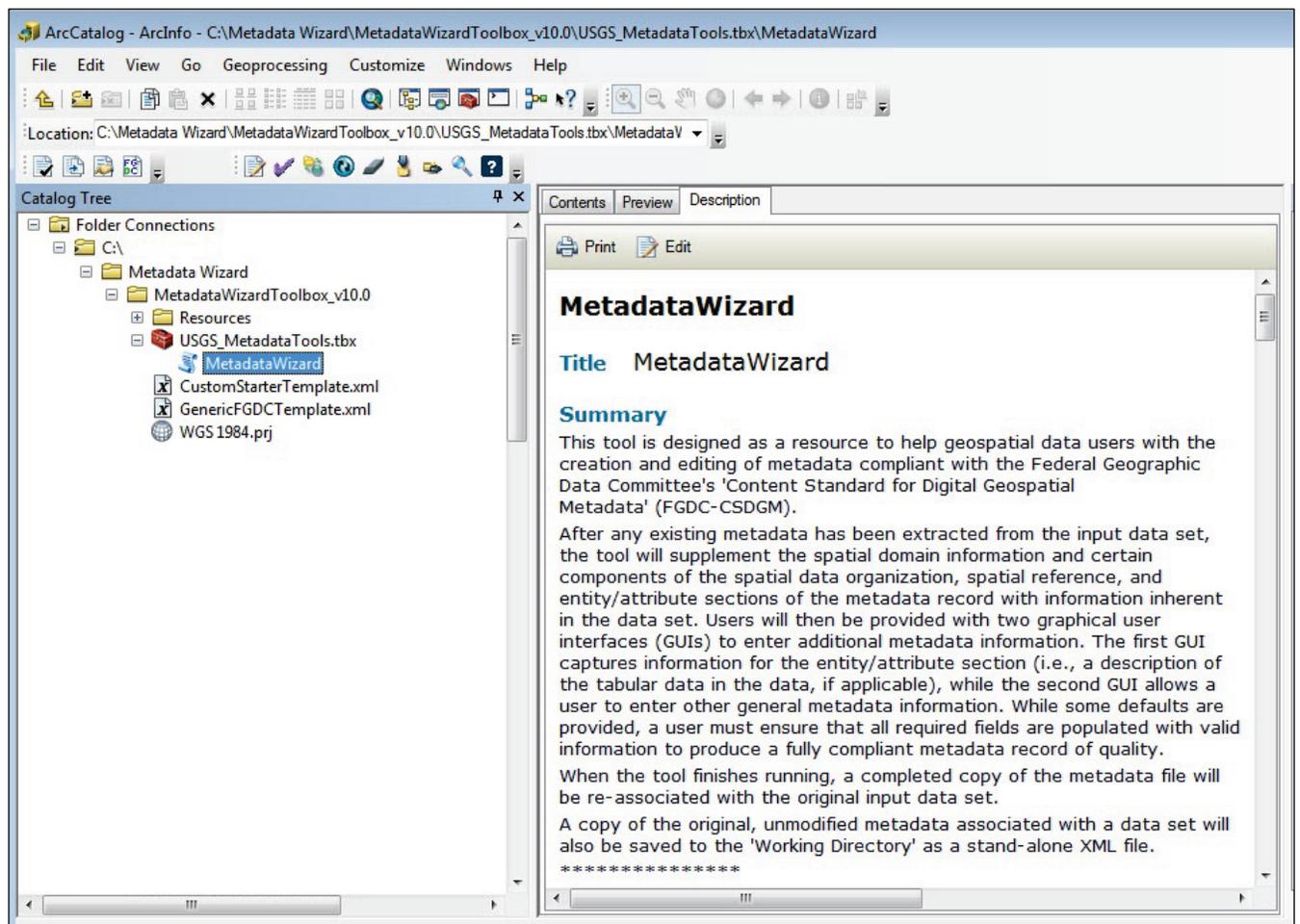


Figure 3. Screen capture of the Metadata Wizard toolbox in the ArcCatalog 'Catalog Tree,' showing description.

The description provided includes information about the general workflow of the software, the software version, additional usage notes, and an infographic that illustrates how the tool will run for each of several different input types (see fig. 7).

A second method for accessing and using the tool in ArcCatalog is to simply ‘add’ the toolbox in the ArcToolbox pane (fig. 4). By right-clicking in the white space of the ArcToolbox pane, users are prompted to navigate to the folder containing the toolbox.

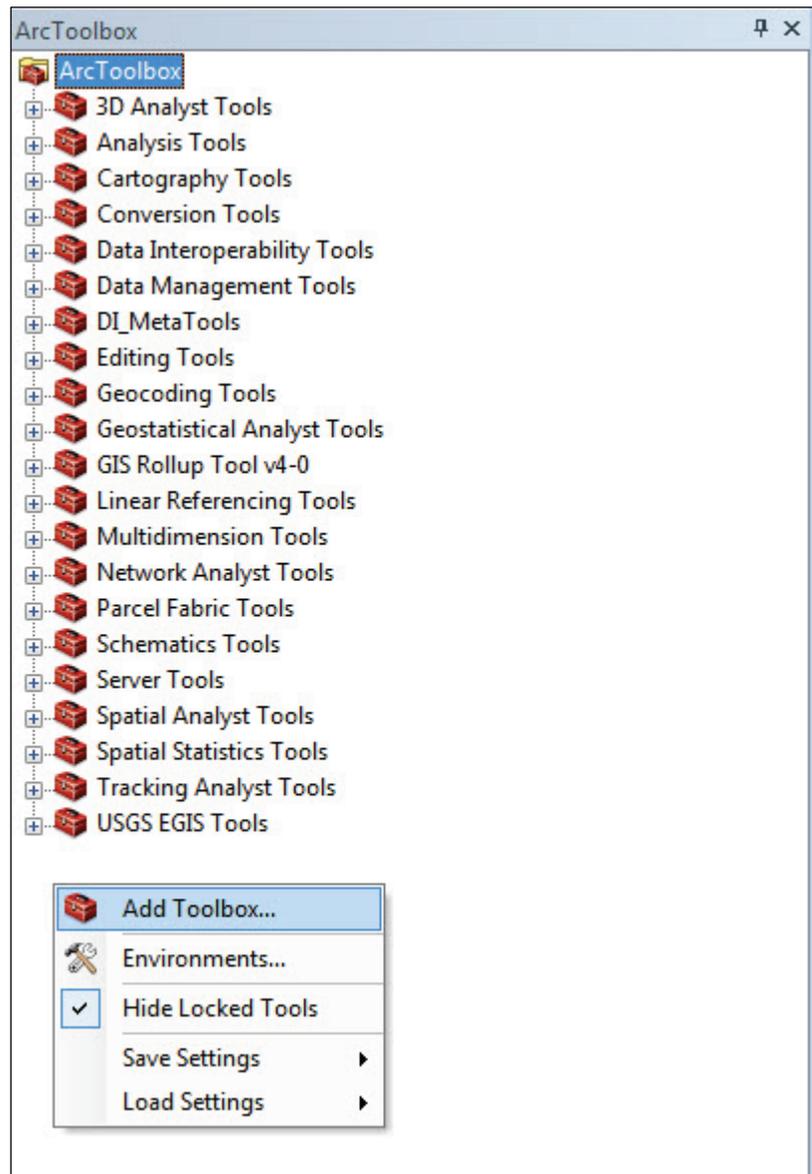


Figure 4. Screen capture of the ‘Add Toolbox’ prompt from the ArcCatalog ArcToolbox pane.

After choosing the ‘Add Toolbox’ option, navigating to the folder location where the Metadata Wizard ‘.tbx’ file was extracted, and selecting the file, the toolbox will appear in the user’s ArcCatalog ArcToolbox (fig. 5).

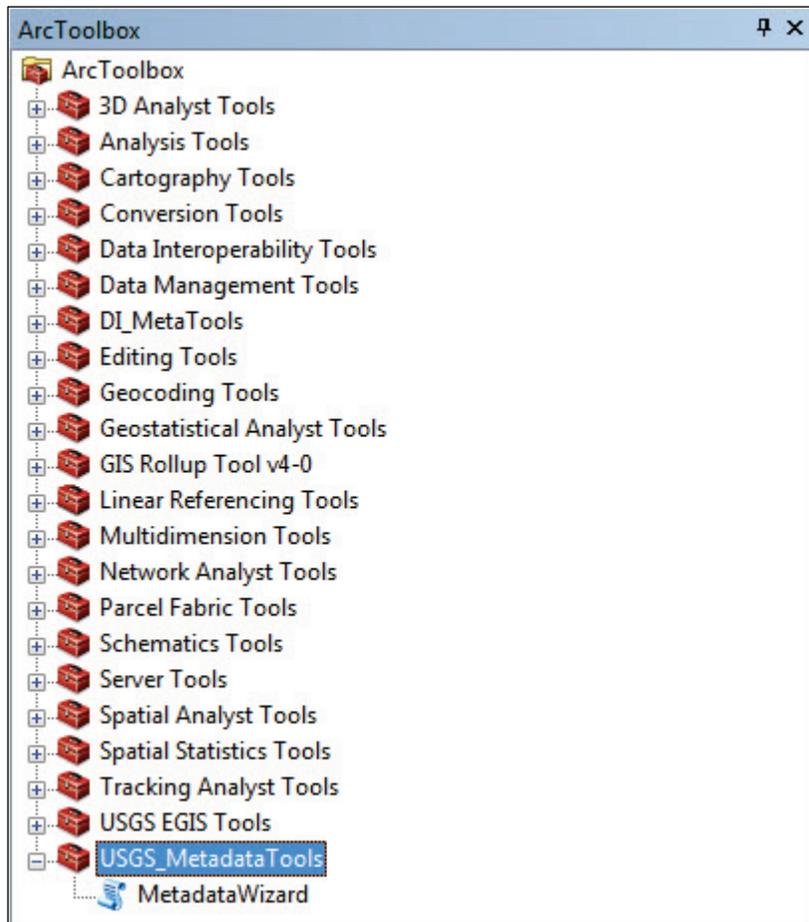


Figure 5. Screen capture showing the Metadata Wizard in the ArcToolbox pane.

Adding the toolbox in this way has the benefit of allowing the tool to persist in ESRI’s software so that it will be present upon future use of ArcCatalog.

Whether the tool is used from the Catalog Tree or in the ArcToolbox pane does not affect its performance. Although the tool can also be added to and used from the ESRI ArcMap application, it is recommended to be used from ArcCatalog due to the fact that all software testing was performed while using the tool from ArcCatalog. The Metadata Wizard was also designed as more of a data-management resource than a geoprocessing tool; data management and organization are generally the domain of the ArcCatalog application in ESRI’s GIS software suite.

Using the Metadata Wizard

After navigating to the USGS_MetadataTools toolbox in ArcCatalog, users simply double-click the “MetadataWizard” icon to run the tool. This will bring up the dialogue box for the tool (fig. 6). The tool (version 1.2.1 shown) has a total of four input parameters/options for use. These four parameters are discussed in depth in the section below.

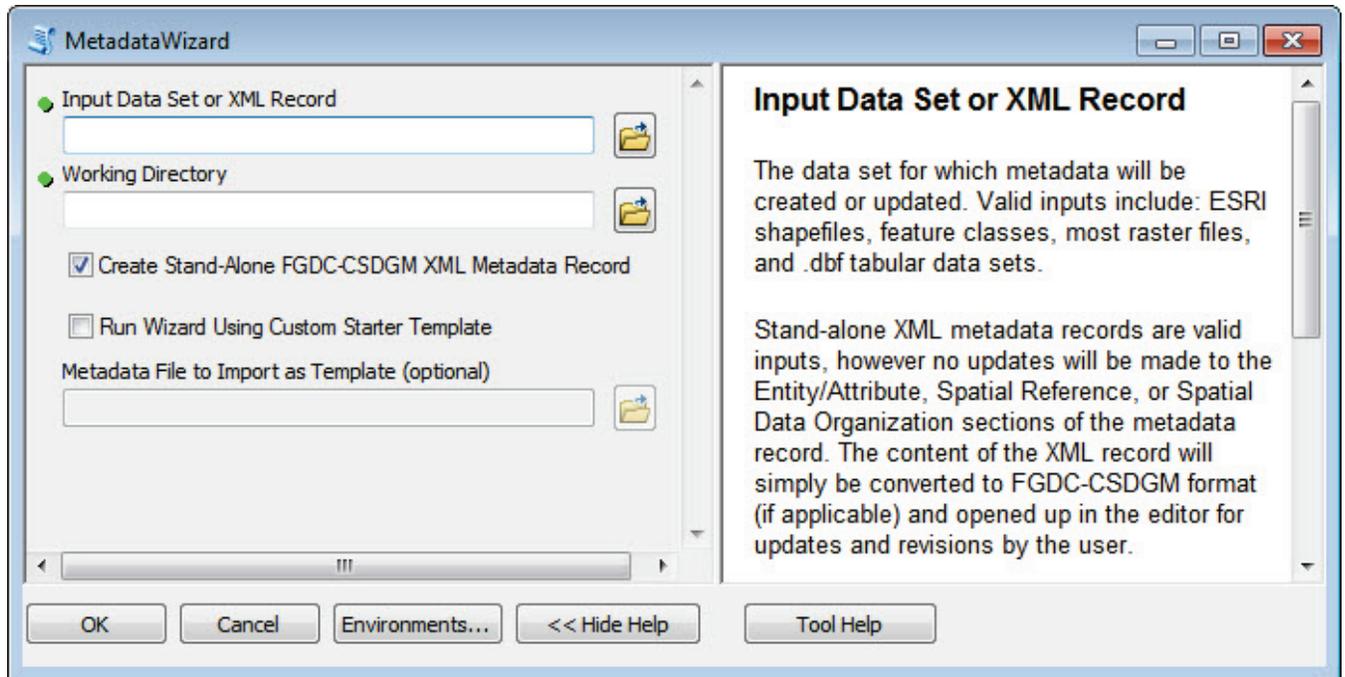


Figure 6. Screen capture of the Metadata Wizard tool dialogue box.

Parameter 1: “Input Dataset or XML Record”

The Metadata Wizard is currently capable of handling three basic types of file inputs: (1) geospatial datasets, including ESRI shapefiles, ESRI featureclasses from file and personal geodatabases, and raster files (in the .TIF, .IMG, .BMP, .GIF, .PNG, .JPG, and ESRI GRID formats); (2) stand-alone tabular data files (in the .dbf, ESRI INFO, or ESRI geodatabase table formats); and (3) XML metadata files.

Due to the way that the Metadata Wizard extracts specific properties from a unique geospatial dataset to pre-populate components of the metadata, ESRI feature datasets within a geodatabase are not supported in the current version of the tool (version 1.2.1). These types of ESRI data objects contain and define a set of multiple geospatial datasets and are thus incompatible with a workflow that extracts the properties and reads information from a single data entity (as the Metadata Wizard does). Users interested in documenting a feature dataset might consider a workflow that uses the Metadata Wizard to develop a record for a participating feature class (within the feature dataset in the geodatabase) and then modifying the record as necessary and importing that metadata record into the higher-level feature dataset using the ESRI Metadata Importer tool.

Parameter 2: “Working Directory”

The “Working Directory” is the folder location that the tool will use to save temporary outputs during runtime. A copy of the original, unmodified metadata associated with the input file at runtime is also saved to this location as a backup. For the tool to run properly, users should ensure that they have write access to this location. If a user opts to preserve a stand-alone XML file of the newly created or updated FGDC–CSDGM metadata record (parameter 3), the Metadata Wizard will also save this stand-alone XML file in this working directory.

Parameter 3: “Create Stand-Alone FGDC–CSDGM XML Metadata Record” (Optional Toggle)

This toggle should be activated when users want to create and keep a copy of the complete stand-alone XML metadata record in the working directory when the tool finishes running—by default this option will be selected. If the user specifies a dataset for parameter 1, versus an XML record, selecting the checkbox for parameter 3 will generate a metadata record for the dataset, as well as a stand-alone XML file. If a stand-alone XML file is provided as an input to the tool, this option will be automatically selected. This option can be a helpful feature for users needing stand-alone FGDC–CSDGM metadata records to accompany publications or for submission to an online metadata clearinghouse.

Parameter 4: “Run Wizard Using Custom Starter Template” (Optional Toggle)

Advanced metadata users can use this option to create a metadata record for the input dataset (parameter 1), using a template file as a starting point to pre-populate some or all of the metadata content. When this option is selected, the user specifies a file from which metadata will be imported. The metadata from this file is then imported into the input dataset (parameter 1), and serves as the starting point for the metadata editing process. The Metadata Wizard is distributed with an XML file (provided with the toolbox) that can be modified and used as a template when running the tool (the file is located/named as follows: “...\MetadataWizardToolbox_v10.x\CustomStarterTemplate.xml”). The “...” denotes the folder where the package is unzipped. Alternatively, users may specify a different template metadata file, which can be an XML file or another ESRI-supported dataset with associated metadata.

When this parameter is selected, a back-up copy of the tool input’s metadata (parameter 1) will be saved in the Working Directory; however, the tool input’s metadata will effectively be ignored in the Metadata Wizard and will instead be replaced by the contents of the Starter Template. Advanced users should consider modifying the Starter Template file (or generating another high-quality FGDC–CSDGM metadata record with standardized content) and using this optional parameter as a way to create metadata more efficiently for multiple datasets that share metadata content. Agencies or user groups can create a quality Starter Template and establish a metadata workflow that always uses the template as a starting point for metadata creation as a way to standardize metadata content and reduce entry errors.

Note that accessing the Metadata Wizard package from a shared storage device can result in conflicts or overwrites if multiple users edit the starter template metadata file (shipped with the tool) from that location. If the software is stored on a shared drive and users decide to modify the starter template file, users can copy the template file to a local location to avoid this problem. Alternatively, users can save the entire toolbox package locally and ‘add’ the Metadata Wizard to ArcCatalog from the local location. With either of these approaches, any changes made to the Starter Template in a local location will be unique to the user.

Software Workflow/Processing Steps

The processing steps performed by the tool vary slightly depending on the tool input provided by the user. Figure 7 illustrates the sequence of steps that the Metadata Wizard performs internally, based on the file type on which it is being run.

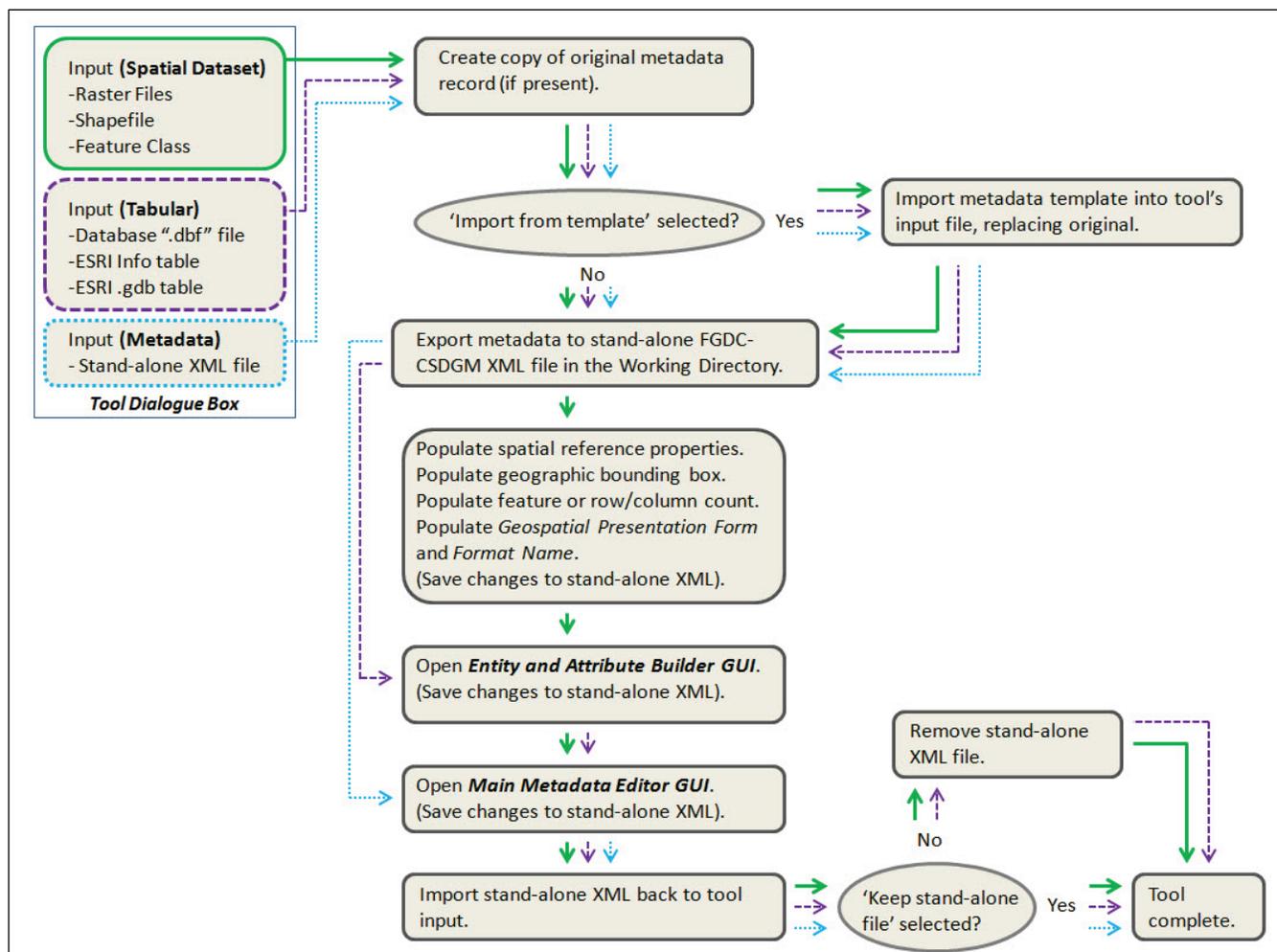


Figure 7. Processing sequence and operational steps performed by the Metadata Wizard.

As figure 7 shows, the tool will run through the complete set of steps outlined in the graphic when a user provides a spatial dataset as the tool input.

For stand-alone tabular data, the tool will bypass the routine to extract the spatial properties, as there is no inherent spatial reference associated with a stand-alone data table. The Entity and Attribute Builder graphical user interface (or GUI) will open, however, providing the user a semi-automated method of describing the columns (and the values they contain) that are present in the dataset. The Main Metadata Editor GUI will also open, allowing for editing of the rest of the FGDC–CSDGM metadata content.

Lastly, if the tool is run on a stand-alone XML file, the Metadata Wizard will again bypass the routine to extract the spatial properties (due to the fact that stand-alone XML files do not possess formal spatial properties). The tool will also bypass the Entity and Attribute builder; because the Entity and Attribute builder in the Metadata Wizard requires access to the table of an input as a data object in order to read in all the columns that are present, this step cannot be performed on a stand-alone XML file. However, the Main Metadata Editor form (the second GUI) will open, allowing a user to easily edit the other metadata content in the XML file. Any existing information in the Entity and Attribute section of the XML record will remain unchanged. Running the tool in this way against a stand-alone XML file

can be an efficient way to create a template file that contains default values and standardized text for accuracy assessments, legal disclaimers, contact information, and other relevant information for an organization. A user can then specify this template file when using the Metadata Wizard in the future to pre-populate content and streamline metadata creation.

Metadata Wizard Data Entry Graphical User Interfaces

Entity and Attribute Builder Graphical User Interface

After the initial tool dialogue box (fig. 6), the Entity and Attribute Builder GUI is the first pane that will be presented to a user when the Metadata Wizard is run on a spatial dataset or a stand-alone tabular dataset. For these types of data, where tabular information is present, the Metadata Wizard will traverse the columns found in the dataset and pull the column headings (and the values they contain) into the Entity and Attribute Builder form so that users can easily provide details about the dataset (fig. 8).

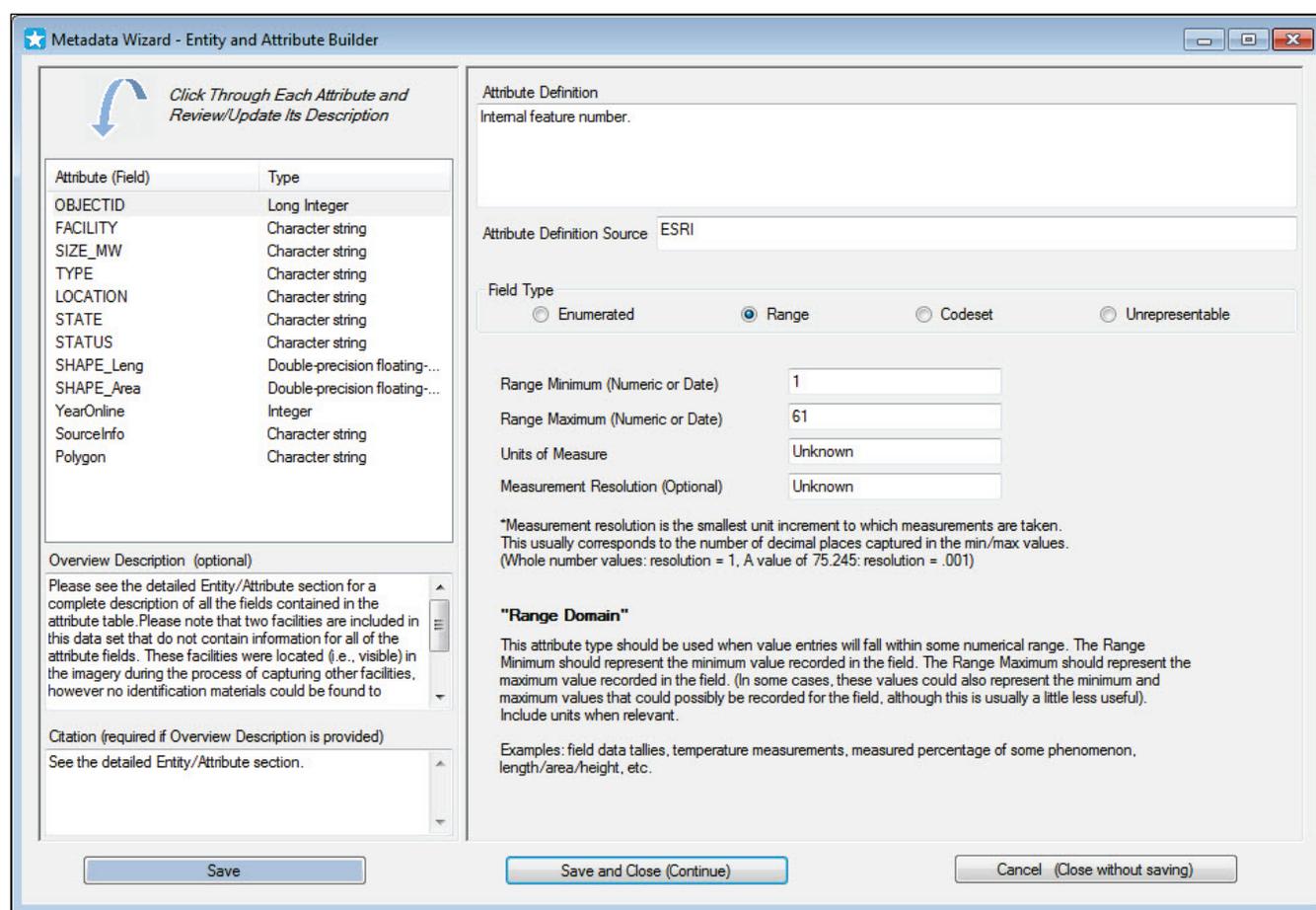


Figure 8. Screen capture of the Metadata Wizard's Entity and Attribute Builder GUI.

A user can define and explain each field in the tabular data, provide an overview, add an attribute definition source if applicable, and selectively toggle between different 'field types' for each column to define the type of information they contain (per the FGDC–CSDGM standard). The wizard

defines each field type to clearly indicate how the different types should be used. The tool will attempt to determine the proper type of field for each particular column based on the values encountered in the data, but users should review each column (and value) description(s) for accuracy.

Main Metadata Editor Graphical User Interface

For all types of file inputs, the Metadata Wizard will open the Main Metadata Editor GUI (fig. 9). For spatial datasets and stand-alone tabular data, this will be the second GUI pane that will open after the initial tool dialogue; for XML files, this will be the only pane that is opened.

Figure 9. Screen capture of the Metadata Wizard's Main Metadata Editor GUI, tab 1.

The Main Metadata Editor form allows a user to enter general information about the input dataset, the individual(s) or organization that produced it, instructions for appropriate use, details on the accuracy of the data, and information about how the data can be accessed.

The Main Metadata Editor GUI has four internal tabs that correspond to different sections of the FGDC–CSDGM metadata standard. Throughout the form, the wizard provides simple language and clear descriptions of the various elements and how they should be populated. The tool's design avoids the esoteric language of the metadata standard to make metadata authoring simple for users who may be unfamiliar with metadata creation or the FGDC–CSDGM requirements. Additionally, all of the elements in the Main Metadata Editor GUI include a 'help tip' to assist users in understanding how the element should be populated; pressing the F1 key (fig. 10) will bring up the help for each element in the form.

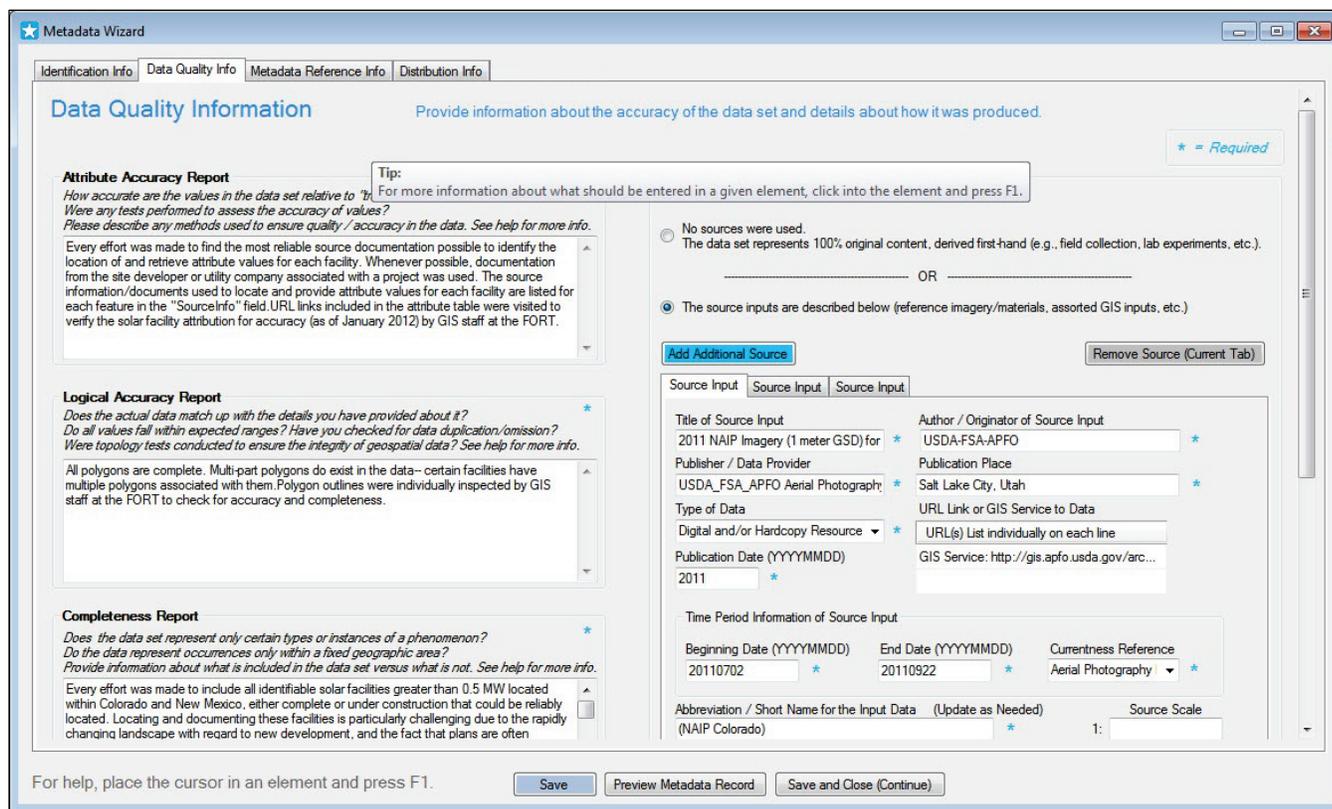


Figure 10. Screen capture of the Metadata Wizard's Main Metadata Editor GUI, tab 2.

The Metadata Wizard tool will extract any existing values from the input's metadata file, if they are present, and display them in the form. After the user updates the metadata content within the form and saves the results, the information will be saved to an XML output and re-associated with the tool's input dataset.

Another feature that the Main Metadata Editor form provides is the ability to preview a metadata record in an easy-to-read format (HyperText Markup Language, HTML). After making edits to a metadata record, users can read through their entire record in the Preview pane (fig. 11) by clicking the "Preview Metadata Button" shown in figures 9 and 10. Users can also easily print their metadata record from the Preview pane, which can be a useful feature for generating hard copies for review. Additionally, users can save the record in a format that is easy for others to read, for instance by printing to a PDF writer utility such as CutePDF or similar (however, users would need to have such a program installed separately on their machine).

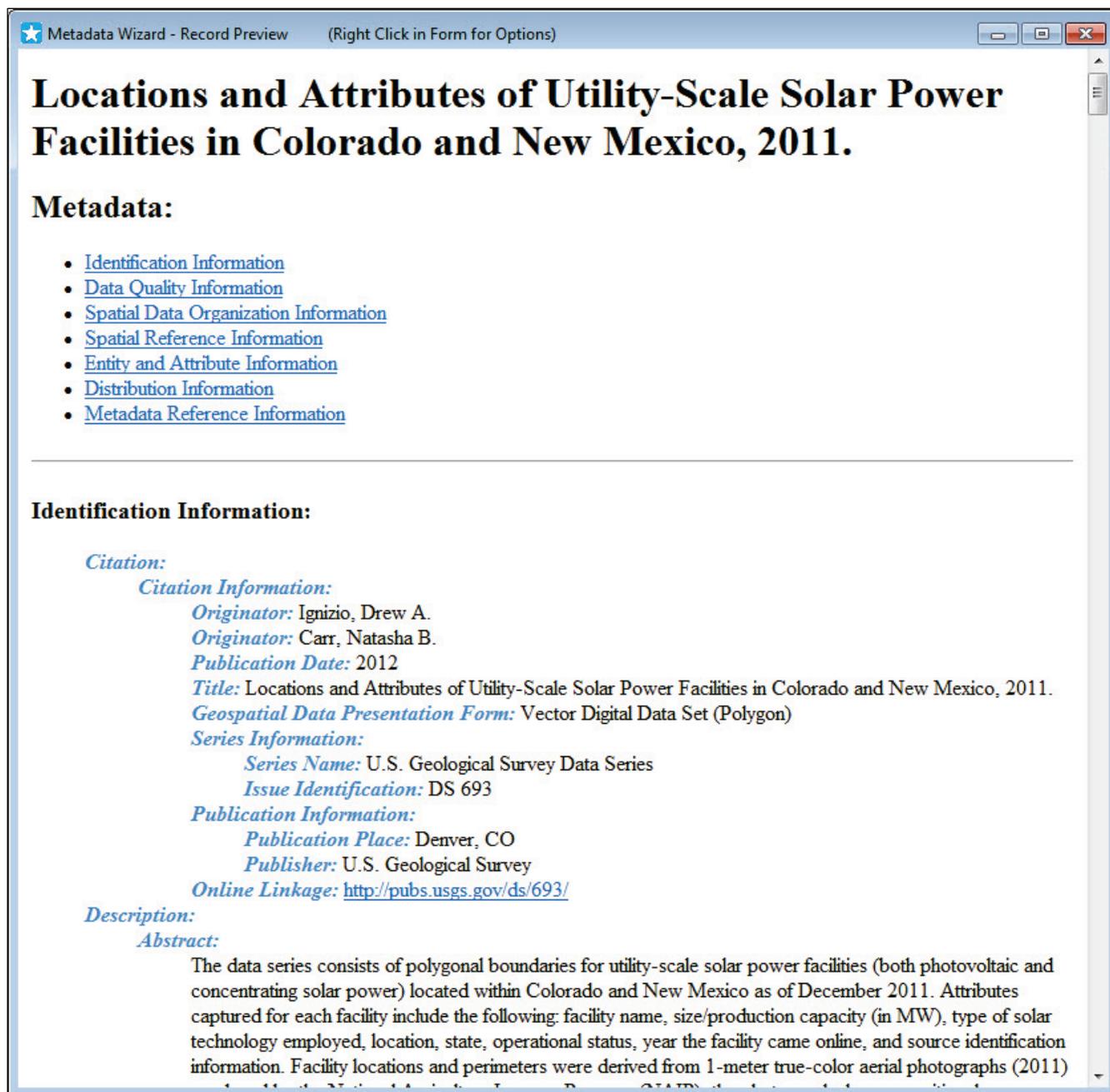


Figure 11. Screen capture of the Metadata Wizard's preview pane.

Tool Updates and Maintenance

As of the time of this publication, the Metadata Wizard tool is considered complete. Minor updates and software bug resolution may take place, as resources permit. The most recent version of the tool will be continue to be provided at www.sciencebase.gov/metadatawizard. Relevant version notes and software changes will be recorded and provided in the documentation made available with the tool download.

Data and Software Disclaimer

Although this program has been used by the USGS, no warranty, expressed or implied, is made by the USGS or the U.S. Government as to the accuracy and functioning of the program and related program material nor shall the fact of distribution constitute any such warranty, and no responsibility is assumed by the USGS in connection therewith.

Acknowledgments

The authors would like to express their gratitude to the USGS Community for Data Integration and the USGS Core Science Analytics and Synthesis program for funding to help support the development of the Metadata Wizard. Additionally we are grateful for the support and collaboration of Peter Schweitzer for his work to help integrate valuable functionality into the Metadata Wizard, as well as for his efforts in developing and maintaining the USGS Metadata Parser (MP) tool, which is an integral resource for the functionality of the software. Lastly we would like to thank the USGS FORT for the institutional support and resources that made this work possible.

Reference Cited

Executive Order 12906, published in the April 13, 1994, edition of the Federal Register, v. 59, no. 71, p. 17,671–17,674. [amended by Executive Order 13286, published in the March 5, 2003, edition of the Federal Register, v. 68, no. 43, p. 10,619–10,633].

ISSN 2331-1258
<http://dx.doi.org/10.3133/ofr20141132>