Sample rst2pdf doc

version

Colin Talbert

September 28, 2017

Contents

MetadataWizard User Guide

The Metadata Wizard 2.0 is a user-friendly tool for creating, editing, and validating CSDGM/FGDC metadata. It is based on design principles of auto-capturing content from the data set being documented, pre-populating a record with reasonable defaults, and lowering the expertise level needed to make a high quality metadata record. The Metadata Wizard 2.0 provides an efficient and pleasant metadata experience for a wide range of users from scientists and data managers. It is being built and maintained by the data management team of the USGS Fort Collins Science Center with input from others in the USGS.

It was engineered as an open-source evolution of the original MetadataWizard to allow for use outside of the ESRI ecosystem.

Installation

Windows Installation

Download a copy of the pre-compiled Windows installer from ScienceBase:

metadatawizard_setup_2.0.0.exe.

Double click on downloaded file to begin installation **does not** require elevated (admin) privileges. Most users will want to select the defaults as they step through the installation procedure in the window that pops up. When the installation finishes, the application can be launched from the duck icon on the user's desktop or start menu.

Installing from Source

Currently the only option for installation on Mac or Linux systems is to install the application from source. | Instructions for installing pymdwizard from source are intended for someone with at least a basic familiarity with Python and Python installations, but should be possible for someonewith minimal experience to follow.

1. Install Anaconda or Miniconda.

   While the code for the MetadataWizard is 2-3 compatible, I would recommend installing the Python 3 version (64x), as 2.7 is being phased out.

2. Open the Anaconda command window.

3. Add the conda-forge channel:

```
$ conda config --add channels conda-forge
```

4. Install the optional conda developer tools

```
$ conda install conda-build
```

5. Create a pymdwizard environment:

```
$ conda create –name pymdwizard python=3.5 pyqt=5.6.0*
```

6. Activate this environment: |

   (on Windows)

   ```
   $ activate pymdwizard
   ```

   (on Linux, Mac)

   ```
   $ source activate pymdwizard
   ```

7. Install git:

   ```
   $ conda install git
   ```

8. CD to the directory you want to install the actual wizard in:

   Your directory below could be different depending on operating system or organization

   ```
   $ cd c:/projects
   ```

9. clone our pymdwizard project:

   ```
   $ git clone https://github.com/talbertc-usgs/fort-pymdwizard.git
   ```

10. CD into our project folder:

    ```
    $ cd fort-pymdwizard
    ```

11. Install the rest of our requirements:

    ```
    $ conda install –yes –file requirements.txt
    ```

12. Add our git folder to the pythonpath:

    ```
    $ conda develop C:/projects/fort-pymdwizard
    ```

13. Launch the Wizard:

    ```
    $ python pymdwizard/gui/MainWindow.py
    ```

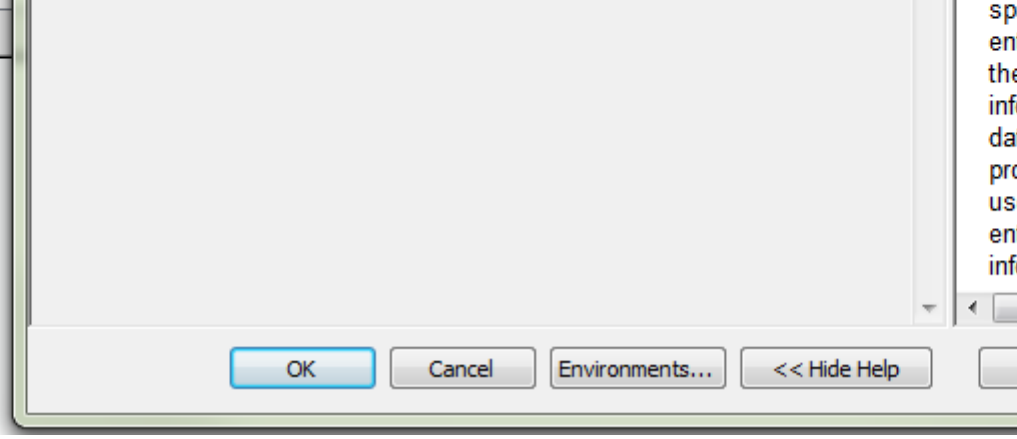
Installing ArcMap Toolbox

Users who want to access the Metadata Wizard from ESRI in the same manner as the previous generation of the tool will need to manually add the Metadata Wizard toolbox to their ArcToolbox. When using the tool from ESRI some additional geospatial data types are supported for the auto-generation of spatial and entity and attribute info. Also the access to existing metadata tied to a geospatial dataset and the saving back out of the edited metadata is facilitated.

The Windows installer provides all of the necessary components as part of the install package. Follow the steps below to add the toolbox to ArcToolbox

1. Find the Metadata Wizard installation directory. If you installed into the default location this will be either C:\Users***your user name here***\AppData\Local\Programs\MetadataWizard or ‘C:\Program Files\Metadata Wizard’ if you installed with elevated priveledges. Note that the ‘AppData’ directory in this path is hidden by default, you will either need to manually type it into the folder path window in Windows Explorer or change your settings to display these hidden directories, see: https://support.microsoft.com/en-us/help/14201/windows-show-hidden-files

Usage



Reporting Bugs and Enhancement Ideas

*Before submitting an issue please make sure you application is up to date by following the steps here.

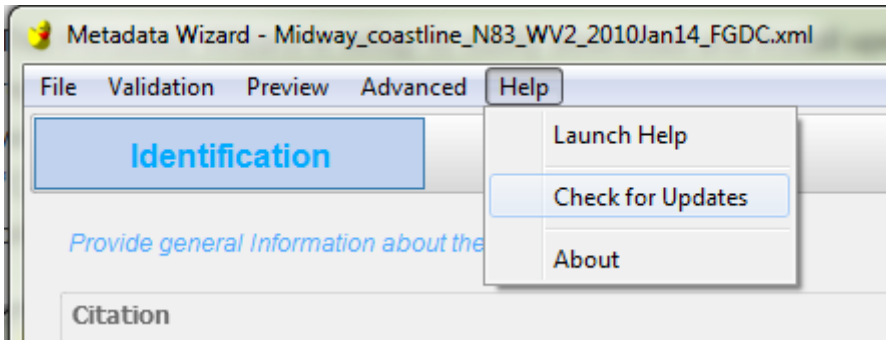
If you encounter any problems while using the Metadata Wizard or have ideas for how to improve the application please submit them as issues directly to the GitHub repository's issues page.

To submit an issue on that page, click the 'Issues' item above to see the list of issues that are currently open. If the issue or idea you have hasn't been added to that list click the 'new issue' button.

Join the USGS Organization on GitHub Adding or commenting on GitHub issues requires a GitHub account. If you do not have an account yet and are a USGS employee follow these steps to *create a free account*

Getting Software Updates

The Metadata Wizard is setup to easily download and install updates directly from this GitHub repository. For this to work the application must be installed in a location that the current user has write access to. If the application was installed with elevated or admin privileges (i.e. in the "C:\program files" directory) this will not work unless the application is being run with elevated privileges when updating.



*To update the application click 'Check for Updates' in the 'Help' menu item at the top of the application.** A message box will pop up notifying you of the result of update. If an update was made, you'll need to restart the application to see it.*

Technical note: Updates are made using simple git commands to **fetch** and **merge** the master branch from this GitHub repository. The commands to execute this operation are stored in a file called "update_wizard.bat" in the application's install folder. Other issues that might prevent this update from working include:

- Edited or otherwise changed files in the local installation directory (merge conflict)
- Internet connectivity issues
- Firewalls or other security measures

Users with knowledge of GIT can troubleshoot these problems directly in the project's GIT repository.

Usage

File Management, New, Open, Save, Save as….

The Extensible Markup Language (XML) file format is the only currently supported format for the metadata used by the Metadata Wizard. The XML format is one of the most commonly used formats for CSDGM metadata and is the format expected by many repositories and tools. By convention the metadata file will be named identically to the data file it describes with the addition of a .xml extension. For example data.csv would have a metadata file called data.csv.xml.

Within the Metadata Wizard XML files can be created, opened, saved, or saved as a new name using a dropdown file menu similar to many desktop applications. In the file browser window that pops up for these functions only files with a .xml extension will be visible.

When first getting started making metadata click the File -> New option and navigate to a directory wher you want to save metadata output. This will create a new .xml with the content contained in the default metadata template as well as the current date in the metadata date element. See Changing-Your-Template for more information about updating your template to streamline future metadata creation.

If you would like to modify or extend an existing record click File -> Open and browse to a local .xml file. Make sure to click File -> Save as… if you want to save the output to a new file. You can also drag-and-drop an xml file onto the application from a file browser to open it.

A list of the last 10 files opened will be stored in the File -> Recent Files: menu item. Clicking on a file name in that menu list will open that record.

When closing the application or switching records a users will be prompted to save their changes. It's a good practice to save periodically when editing a record by clicking File -> Save.

File Management in this new version of the Metadata Wizard is quite a bit different the previous ESRI toolbox based version.

Using the Metadata Wizard

Sections

The Metadata Wizard is set up as set of forms/tabs you'll enter information into about your data set. These sections are broken down categorically into sections that roughly correspond to the sections in an FGDC metadata record.

- **Identification** : *What the data is, who created it, associated publications, when it was created, etc.*
- **Data Quality** : *How was the data created (methodology), what sources were used, what quality control was used.*
- **Spatial** : *Where is the data geographically, what spatial projection or reference is it in, how is it organized spatially.*
- **Entity and Attribute** : *How is the data organized, what is in each of the columns, what units are they in, etc.*
- **Distribution** : *How would someone obtain the data.*
- **Metadata Reference** : *Who made the metadata record, when, what standard was used.*

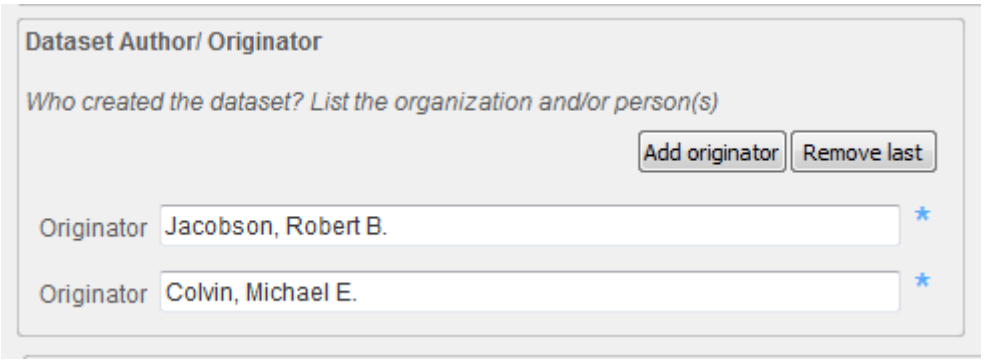
Filling out a record

Cycle through each of the sections above by clicking on the blue labels across the top of the application. On each one, enter appropriate content into the text boxes. There are often hints above each that help to identify the appropriate content. Items marked with an * are required and must be filled out. | |

Expanding a collapsed element

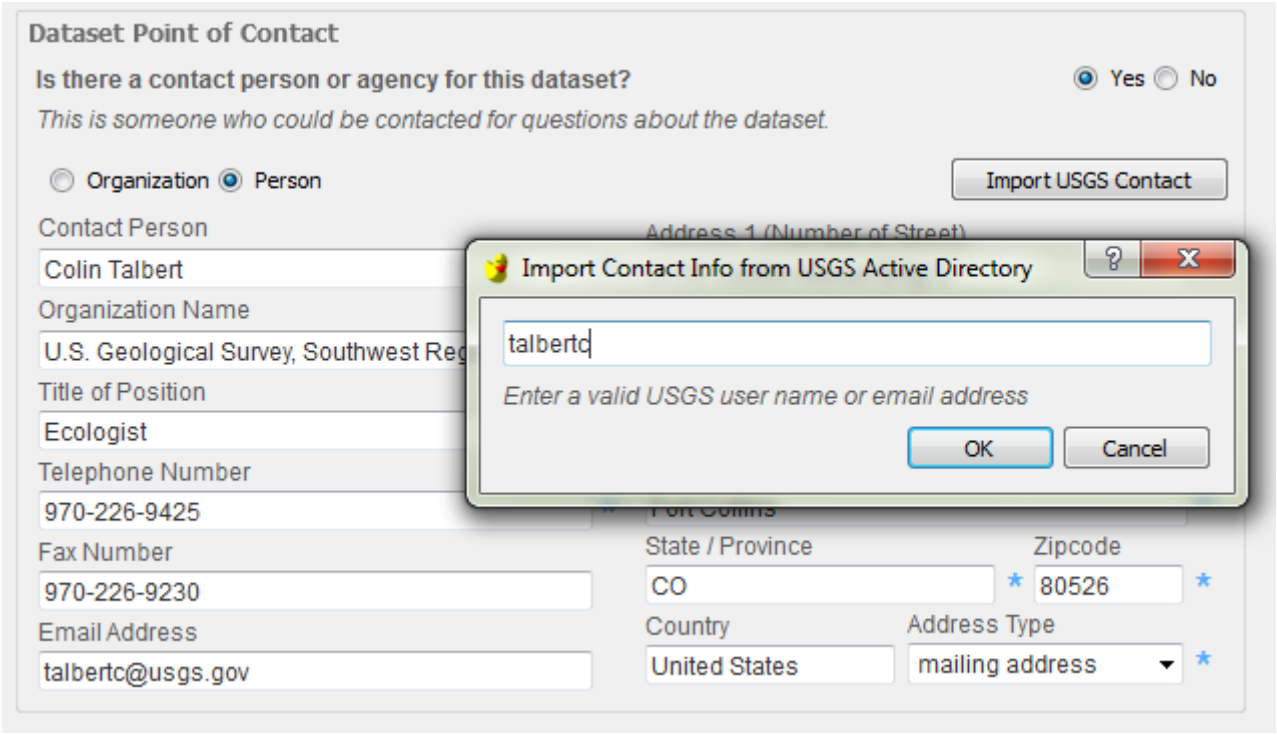
Some compound items are hidden until you check the 'Yes' radio button next to the question asking if they are needed for that record. For example, if your data is part of a series you will want to record the series name and number in the Citation on the Identification tab. These will not be visible until you check Yes to the question 'Is this dataset part of a series?'.

Adding or removing items in a list (repeating elements)



Items that can be repeated multiple times such as authors, keywords, online linkages, etc., have buttons for adding additional sections or removing the last one.

Using tools to auto-populate compound element



Some elements contain convenience tools for populating their content which are launched by clicking the button in that section. For example contact information for USGS users can populated by clicking the 'Import USGS Contact' button in a contact section, and entering their Active Directory username.

Other tools include adding keywords from controlled vocabularies, populating a citation section from a active DOI, generating a taxonomy section from the Integrated Taxonomy Information Service (ITIS)

Spatial Content

Many data have an inherent spatial context that needs to be captured in the metadata. At a minimum one should record the geographic location the data were collected, generally called the extent or bounding box.

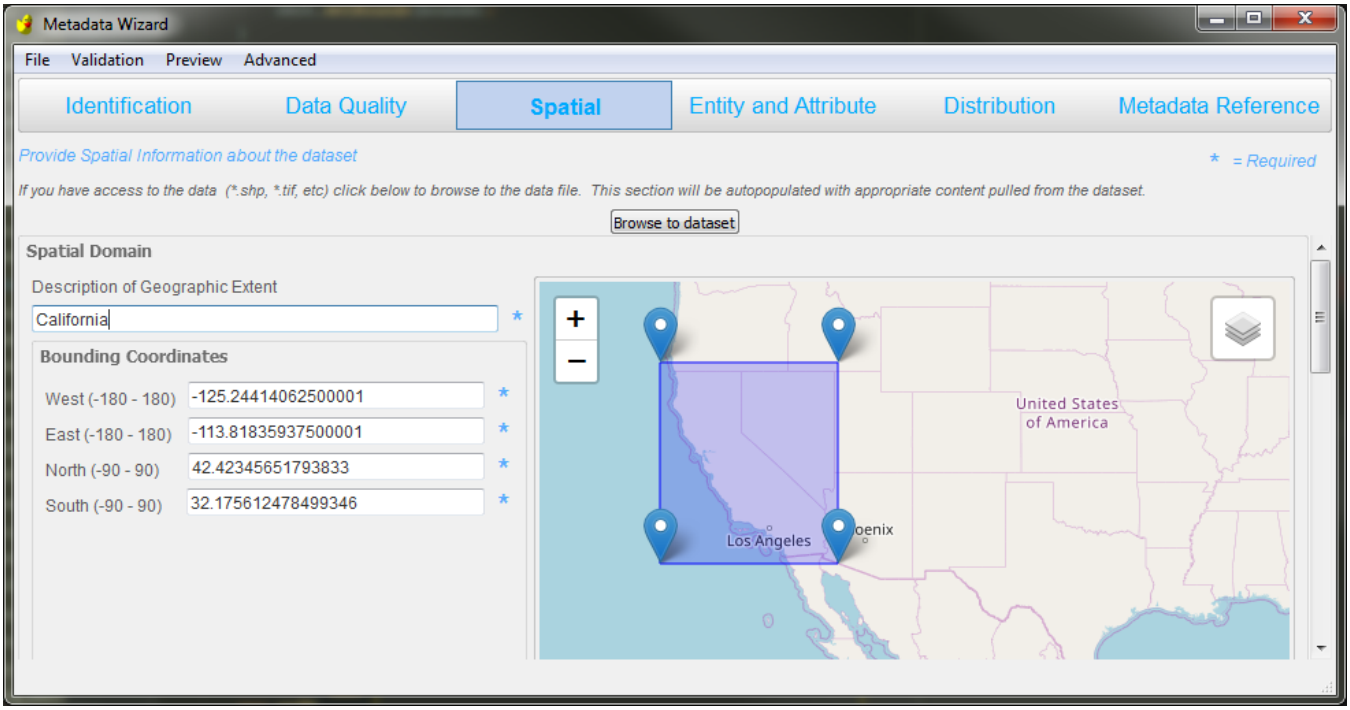
If the data are in a spatial data format, for example a shapefile or raster, you will also need to document the spatial reference or projection and how the data are organized spatially.

Setting the bounding coordinates manually.

If your data are in a non-spatial format, a .csv for example, you can manually specify the bounding box interactively in the Spatial Tab. By default new records in the Metadata Wizard start with an extent set to the whole world. This can be changed by editing the East, West, North, South coordinates in the text boxes to the left of the map. The map will update to show the extent as specified in the text boxes.

If you don't know your bounding coordinates or would like to interactively edit them you can do so in the map to the right of the text boxes. Click and drag one of the corner markers to resize the box. The mouse wheel or +/- buttons in the upper left of the map can be used to zoom the map. Clicking and dragging anywhere other than the corner markers can be used to recenter the map. As you zoom the map in and refine your extent additional detail will be displayed on the map to help locate your study area. The layer button in the upper right can be used to switch to a satellite imagery background.

When using the Biological Data Profile don't forget to update the Description of Geographic Extent to match the bounding coordinates displayed, e.g. Southwest United States, Colorado, Key West Florida.



Interface for editing the spatial extent

Auto-populating spatial content from data

If your data are in a spatial data format the information for this section can be auto-populated by pointing to the appropriate file. The currently supported file types are:.shp, .tif, .jpg, .bmp, .img, .png, .grd. In addition to extracting the data's extent the coordinate system/projection and spatial data organization will be extracted. For this functionality to work the data must have a defined projection. Note that some less common projections might display incorrectly, so do check that the values imported are appropriate

How to fill out the Entity and Attribute tab

The Entity and Attribute section is one of the most useful sections for data re-use. It allows users to unambiguously determine the appropriate way to interpret the data. Included are items such as what units the data are in, measurement resolution, what abbreviations stand for, and where you could look up the categorical values used.

This tools facilitates a robust Entity and Attribute generation by auto-populating much of this content by introspecting the data being documented directly. Currently CSV, Excel, and Shapefile formats are supported, with additional geospatial formats available when calling the MetadataWizard from the ArcToolbox.

Given the diversity of data that can be described in a single Entity and attribute metadata section, here are some options for how to fill out this section:

Data described elsewhere

The data are fully described in a data dictionary or other document that can be referenced in this record.

Leave the ‘Detailed’ tab empty but fill out the ‘Overview’ with a description of where one can obtain the appropriate data dictionary or other document. Make sure that the document being referenced is going to be persistent and easily available to all users of this data. If there is some question about this, consider including a copy of the data dictionary with this data/metadata.
***

Data in proprietary/complex format

The data are in a proprietary format (Genbank format data), organized in a complex but non-tabular format (folder with individual files that adhere to a naming convention), or are easily described (A photo scan of a historic map)

Leave the ‘Detailed’ tab empty but fill out the ‘Overview’ with a description of the data format, naming conventions, or contents
***

Single tabular dataset

The data are tabular in nature (CSV, shapefile attribute table, Excel worksheet) and in a ‘tidy format’ where each variable forms a column, each observation forms a row, and each observational unit forms a table

On the ‘Detailed’ tab click the ‘Browse to Dataset’ button on the left and navigate to your data file. Appropriate content will be extracted from your file to fill the form. Sequentially fill out each of the columns on the right, by providing a definition of what’s in that column. For each column also provide the Column Contents which will be one of:
1) Enumerated (Categorical/Factor data for which a definition will need to be provided for each unique value in the column.
2) Range for numeric data, a min, max, and the units will be required
3) Codeset for when the values can be obtained from a definitive source which is specified
4) Unrepresentable for every other case where you must provide a free text description of how to interpret the values in this column.
***

Multiple tabular datasets

The data consists of multiple distinct tabular datasets, for example several related CSVs or multiple sheets in an Excel workbook.

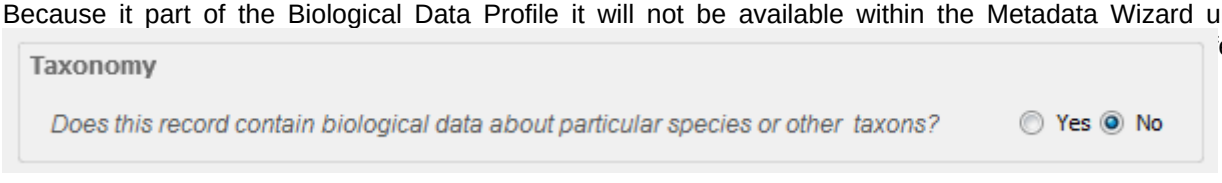
Click the ‘Add Detailed‘ button below

Generating a Taxonomic Information Section

In an FGDC record the Taxonomy section is used to record information about the species (or other level) being studied.

Finding the taxonomy section

Because it part of the Biological Data Profile it will not be available within the Metadata Wizard unless ‘FGDC



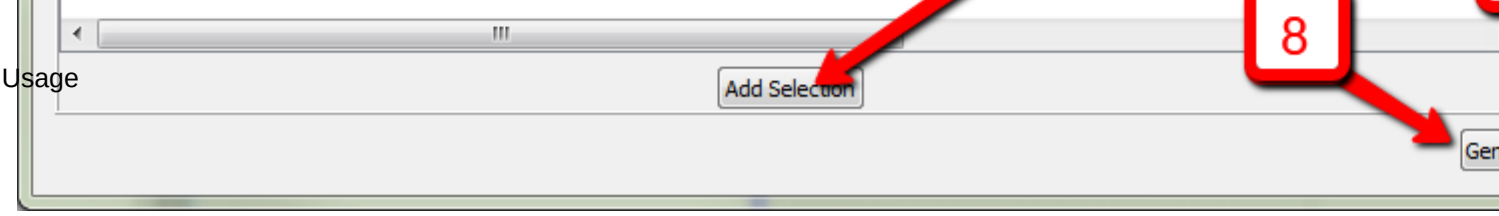
erence Tab.

Tab (halfway

Using the taxonomy section builder

Within the Metadata Wizard it is intended that users will be auto-generating a Taxonomy Section using the built in tool which pulls data from webservices hosted by the Integrated Taxonomic Information System (ITIS)

To launch the Taxonomy builder click the button ‘Add Items from ITIS’



1. Start by entering a scientific or common name into the ‘Search Term’ box. The term used can be from any level of the scientific heirarchy (order, family, species, sub-species, etc.).
2. Make sure that ‘Search Type’ either ‘Scientific Name’ or ‘Common Name’ matches the term used.
3. Click ‘Search ITIS’ to get a list of search results which will be displayed below.
4. Select one or more items from this list.
5. Click ‘Add Selection’ to add the selected item to the list of taxonomic items that will be included.
6. Steps 1-5 can be run as many times as needed to complete the list of items to include. If any item needs to be removed from this list, select the item and click the ‘Remove Selection’ button below.
7. If you would like to include the common names in your taxonomy section check this box.
8. Click the ‘Generate Taxonomy Section’ button to create a taxonomy section based on the the list in the ‘Items to include:’ box above.

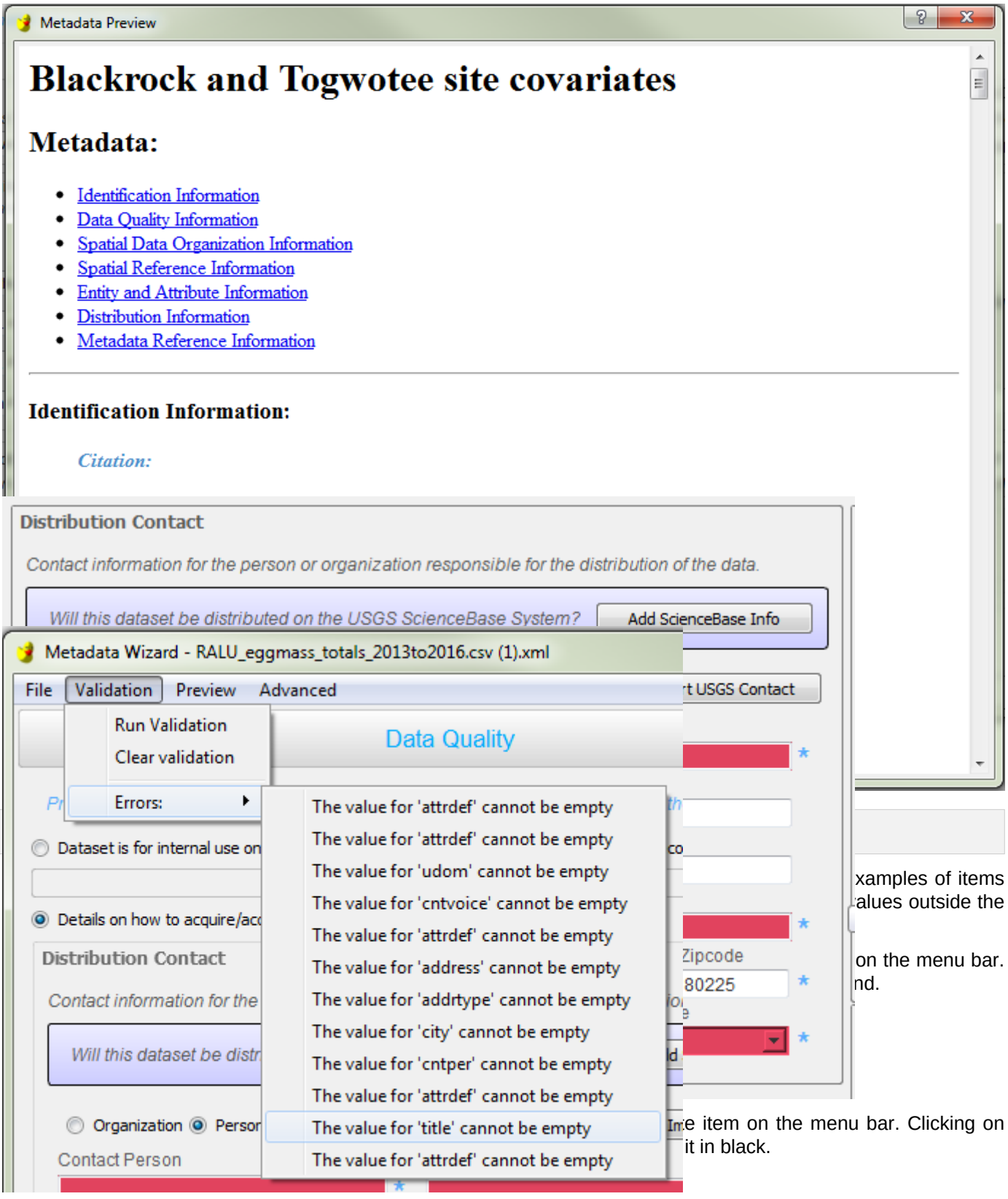
Limitations

Be aware that currently the Taxonomy section in the Metadata Wizard does not contain the ‘Taxonomic System’ section of the standard, nor is it intended to allow for creating a record manually from scratch.

Previewing a Record

Sometime it can be helpful to look over, or read an entire record. Within the Wizard there is a build in Preview window that can be accessed at any time by selecting the “Preview” item in the Preview menu.

The window that opens will contain a copy of the entire metadata record rendered using the default stylesheet.



xamples of items
alues outside the

on the menu bar.
nd.

e item on the menu bar. Clicking on
it in black.

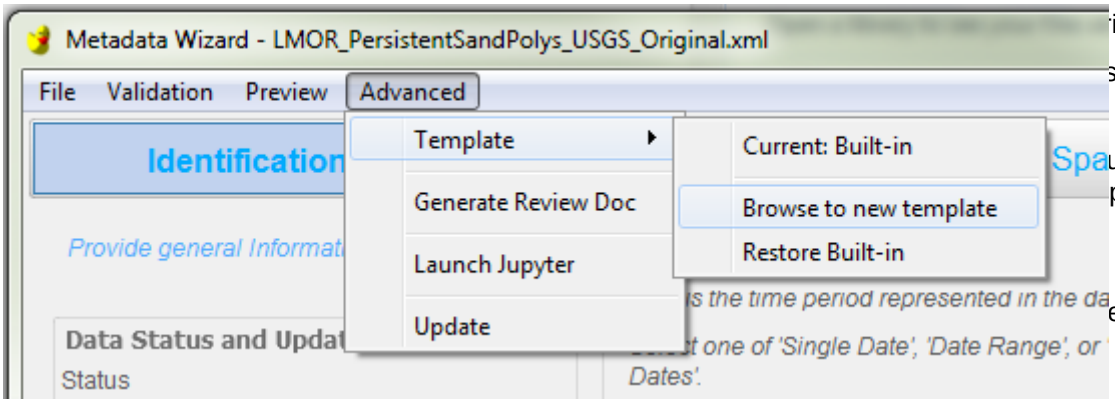
Keep in mind that to find some errors you might need to scroll up or down in the application or select the appropriate tab.

Advanced Usage

Changing Your Template

When you select File->New the Wizard is opening a ‘template’ file and saving it to a new location. This template is just an FGDC record that contains default content which forms a solid starting point for a blank new record. For example the point of contact, distribution contact, disclaimers, etc. will often remain the same for most records created by an individual. Having to update these individually every time a new record is started can be inconvenient. Keep in mind that anything specified in the template can be overwritten individual when edit a new record, they are just the defaults that will pre-populate the application.

The default template record used by the Metadata Wizard is part of the standard installation but can be changed to suit the particular needs of individuals or organizations.



itten or inaccessible.

s. Sections to consider

ution Information, ISO
port, Logical Accuracy

e to the xml file saved

BrowseToTemplate.png

The template selected will remain active after closing and reopening the application. If at any time you want to revert to the original Metadata Wizard template click the Advanced -> Template -> Restore Built-in item. The current temple file name will be displayed at the top of the Advanced -> Template menu,

Indices and tables

- **genindex**
- **modindex**
- **search**