

XMLInput

Application Guide


Version 1.6

August 23, 2002

**Mid-Continent Mapping Center
1400 Independence Drive
Rolla, MO 65401
Richard E. Brown (reb@usgs.gov)**

(573) 308-3525

Table of Contents

OVERVIEW	4
GENERAL	4
ABOUT THE INTERFACE	5
ENVIRONMENT	6
FEATURES	6
WHAT'S NEW	6
STILL TO COME	6
DOWNLOADING THE APPLICATION	7
JAVA RUNTIME ENVIRONMENT (JRE) DOWNLOAD	7
XMLINPUT DOWNLOAD	7
INSTALLING THE APPLICATION	8
OVERVIEW	8
INSTALLING THE JAVA RUNTIME ENVIRONMENT (JRE)	8
INSTALLING XMLINPUT	8
CREATING A SHORTCUT TO XMLINPUT	8
RUNNING THE APPLICATION	8
THE MENU BAR	9
FILE PULL-DOWN MENU	9
TOOLS PULL-DOWN MENU	9
HELP PULL-DOWN MENU	10
THE TOOL-BAR	11
TEMPLATES	12
SELECTING TEMPLATES	12
ADDING & SAVING NEW TEMPLATES	13
CREATING NEW TEMPLATES	14
<i>From Existing Master Templates</i>	14
<i>From an Existing DTD</i>	14
<i>From "Scratch"</i>	15
EDITING	16
GENERAL	16
EDITING ELEMENT TEXT	16
ADDING COMMENTS	16
RENAMING ELEMENTS	18
MANIPULATING ELEMENTS WITH THE POP-UP MENU	19
<i>Expand</i>	19
<i>Collapse</i>	19
<i>Copy</i>	19
<i>Paste</i>	19
<i>Insert</i>	19
<i>Duplicate</i>	19
<i>Delete</i>	20
FILTERING OPTIONS DIALOG	21
OUTPUT ONLY FILTER OPTIONS	21
XML DATA FILTER OPTIONS 	21

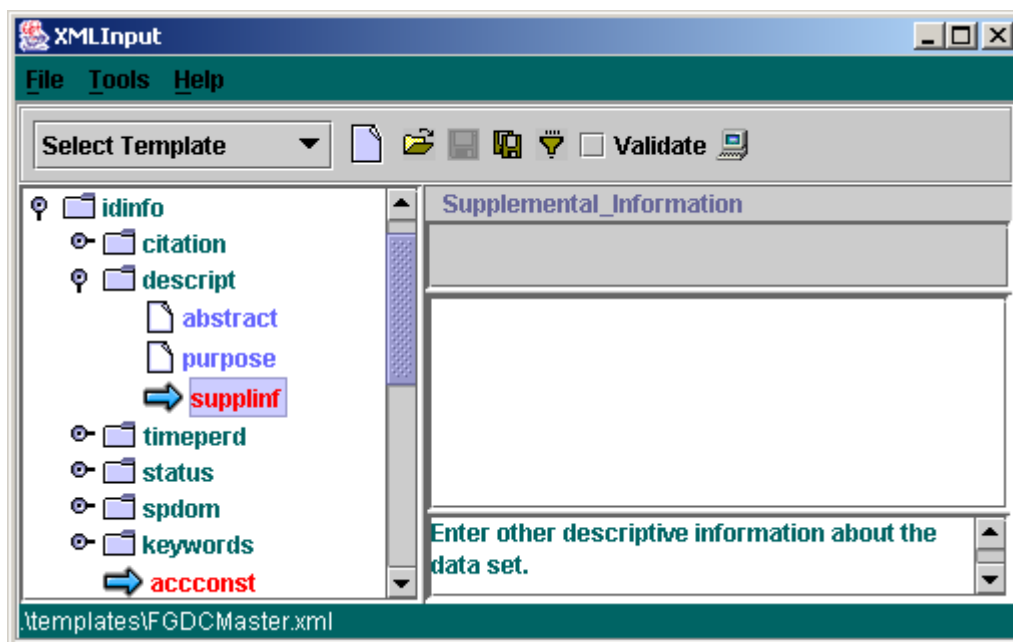
Overview

General

Data producers, researchers, librarians, and anyone who has a requirement to collect and exchange metadata can benefit from the power and flexibility of the *XMLInput* application.

XMLInput is based on the [Extensible Markup Language \(XML\)](#) that defines a universal text file *format* standard that can be used to visualize, store, and exchange data representing various requirements and standards for data content. This industry format standard assures users their data will be exchangeable with various customers between various systems.

XMLInput preserves the various data content requirements and standards as XML-based templates that can later be completed and saved to disk or transferred in the XML format.



The interface design makes data collection as intuitive as possible. A tree structure organizes the various metadata elements. Symbols and icons highlight template areas that require input. XML comments serve as user prompts to aid in describing the input required. XML attributes provide guidance to automated data integration applications. Multi-line entry fields allow cut/copy and paste for large blocks of text. New elements may be created, copied, pasted, and duplicated. And templates can be created and modified from existing templates.

The USGS has reviewed this tool for suitability in collecting subsets of the FGDC metadata standard from various data collection sources. A companion tool, currently under development, will make integration of this XML data with enterprise databases more automated.

About the Interface

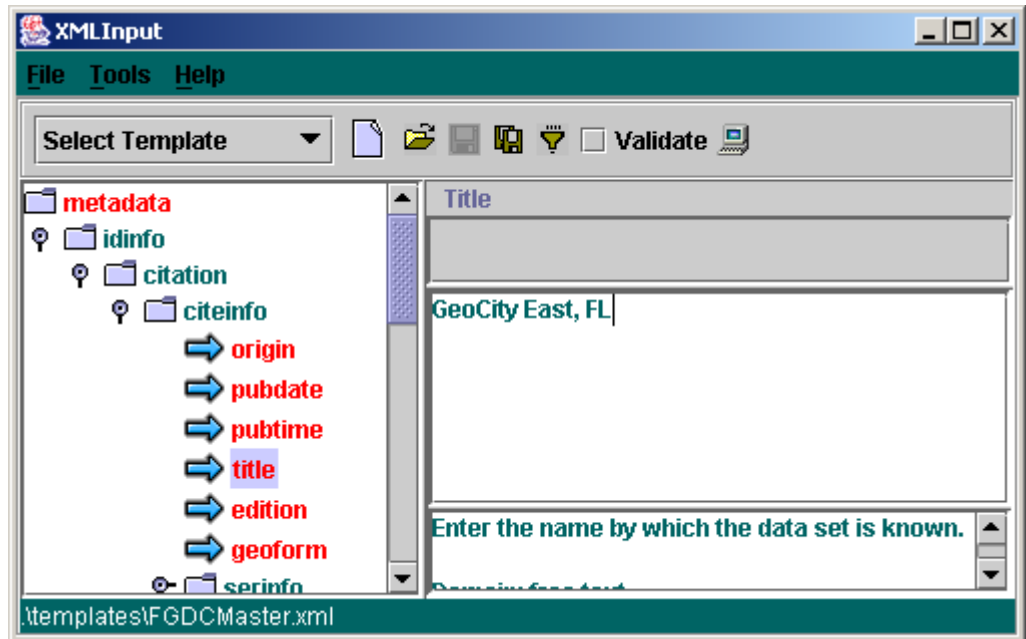
The XMLInput application can be divided into a number of components.

The **Menu Bar** occupies the line directly under the Window title (XMLInput). This bar is divided into three sections: File, Transfer, Help.

The **Tool Bar** occupies the next line and contains: the Select Template Box, New, Open, Save, SaveAs, Filter, Validate, and System viewer buttons.

The **Element Tree View** contains a hierarchical view of the underlying XML data structure which is similar to a Microsoft Windows *Explorer* view. This occupies the left side of the editing portion of the application

The **Element Label** is a gray area on the top right side of the interface that shows the longer name of the selected element.



The **Attribute Editing View** is directly below the Element Label on the right side of the application. Attribute *values* may be edited and new *attributes* inserted. Attribute deletes have not been implemented yet.

The **Element Editing View** occupies the middle section of the right side of the application that in the picture above contains the words "GeoCity East, FL". This is the area where element content is added. Copy/Cut/Paste keyboard combinations are available in this area for inserting multi-line text. Words are wrapped at the view borders when displaying multi-line text and a scrollbar appears when inputs exceed the displayable area.

The **Comment Editing View** occupies the lower section of the right side of the application. This area will contain any XML comments included as part of the selected element content. These are commonly used as user prompts to define the input required for a particular element.



Environment

XMLInput is a Java application. This allows use of this program on any platform (Windows and UNIX) that support Java JDK 1.3 or greater. Minimum system requirements have not been established at this time but the application performs excellently on systems with 500Mh processors and 128MB Ram.

Features

XMLInput has the following features:

- Multi-Platform – Java Based
- Variable Content XML-based Data Format (FGDC/GILS/Other)
- Ability to Create, Modify and Save XML-based [Templates](#)
- XML Validation (DTD & XML Schema)
- XML Element, Attribute and Comment Editing
- XML Element Modifications (Insert, Delete, Copy, Paste, Duplicate)
- XML Filter Options to filter out comments, attributes or empty elements.
- XML Output Options to modify the appearance of the resulting XML file.
- Email XML Submission of raw XML

What's New

DTD Template Creation

New templates can be created directly from a specified DTD. By selecting the desired DTD and root element, the application automatically builds all the non-optional elements specified by the DTD.

DTD-driven Element Insertion

When a DTD is specified in an existing XML file, the application automatically displays the possible sibling and child elements for a given context when using the insert function.

Still to Come

A number of enhancements are still in work for this application. A **Print** function is needed to overcome limitations inherent in most XML-based applications for obtaining formatted hard-copy output. **XML Schema** support will be added as open source libraries become more mature.



Downloading the Application

Java Runtime Environment (JRE) Download

XMLInput is a JAVA application that requires a current version of the Java 2 Runtime Environment (JRE).

Your system may already have a JRE installed. For Windows platforms this can be checked by



opening up the Control Panel and looking for a "Java Plug-in" icon.

If this icon is present you already have a JRE installed. Double click on this icon to make sure that it's at least 1.3 or higher. If you don't already have a JRE installed the download of the most current version should always be available at <http://java.sun.com>.

The JRE is platform dependent and all supported platforms should be available at the above site. Currently the [j2re-1 3 1 03-win-i.exe](#) JRE for Windows is available from the MCMC FTP site at:

<ftp://ftpmcmc.er.usgs.gov/release/java/>

XMLInput Download

XMLInput may be downloaded from the MCMC FTP site at:

ftp://ftpmcmc.er.usgs.gov/release/xmlinput/XMLInput1_6.zip

Clicking on the above link should start the download process by prompting you for a directory to save the file in. Choose any directory - but remember where you put it.

Installation instructions are included in the next session.



Installing the Application

Overview

Installation of *XMLInput* involves the following steps:

- Installing a current Java Runtime Environment (JRE),
- Extracting the *XMLInput* download file into a new directory,
- Creating a shortcut to the *XMLInput* JAR (.jar extension) file. (Optional)

Installing the Java Runtime Environment (JRE)

The Java Runtime Environment (JRE) download file is a self-extracting installation program that prompts the user for certain installation options. If you are uncertain about any of these options, the default options should be sufficient for the *XMLInput* application.

To begin installation, double-click on the JRE download file.

Installing XMLInput

The downloaded file (*XMLInput.zip*) is in compressed (ZIP) format and contains a number of files necessary for the application. The final directory structure should look like this:

```
./XMLInput
./XMLInput/img
./XMLInput/help
./XMLInput/templates
```

Extract the ZIP file to any directory making sure to extract all the files and retain the existing directory structure.

Creating a Shortcut to XMLInput

You may create a shortcut for the *XMLInput* application by right clicking on the *XMLInput.jar* file and selecting 'Create Shortcut' from the pop-up menu. The shortcut will be created in the same directory as the JAR file. Simply drag this to your Desktop or any preferred location.

Running the Application

To run the *XMLInput* application, simply double click on the *XMLInput.jar* file located in the application directory. If the *XMLInput* window does not appear you may need to create a batch file with the following contents:

```
java -cp xmlinput.jar xmlinput
```


The Menu Bar


File Pull-Down Menu

Then file pull-down menu contains common commands for working with the XML files.

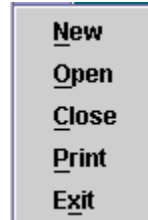
The **N**ew command creates an entirely new file starting with the user specified root element. See: [Creating New Templates](#) for further information.

The **O**open command opens an existing XML file located on the local file system. Templates can be opened from the application's templates directory or any other valid XML file may be opened. The File Selection Dialog allows the user to select a file.

The **C**lose command replaces the current XML file with an empty XML file with the root element named 'empty'. This element can be renamed and additional elements can be added without selecting the 'New' command.

The **P**rint command has not been implemented yet. This will be one of the first enhancements to the application and is included to let user's know the need for this functionality has been identified. For now, printing can be accomplished in a two ways. The XML file can be opened in another application that reads text and then printed from that application or the System  button on the Tool Bar may also be used to create a temporary file of the current XML which is then opened in the default system browser.

The **E**xit command closes the current session of *XMLInput*.



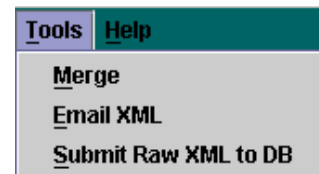
Tools Pull-Down Menu

The Transfer pull-down menu is provided as a convenience feature to automate the electronic transfer of XML data.

The **M**erge command allows multiple XML files to be merged into a single file. Details on the merge command are available in [Appendix A](#).

The **E**mail command sends the current XML file into a Mail Dialog box where the XML may be emailed to a specified person. The same thing can be accomplished by attaching the XML file to a email message and this implementation is only provided as a convenience.

The **S**ubit Raw XML to DB command will, when implemented, send the current XML file to a CLOB-like field of a specified database table. This command has not been implemented yet but should be included in a future release.





Help Pull-Down Menu

The Help pull-down menu is provided to allow access to information about the *XMLInput* application. Currently, only the User's Guide and an About dialog are available. The User's Guide is available in PDF format. The About box is




The Tool-Bar

The ToolBar provides quick access to common functions.




The Select Template Box stores a list of templates found in the applications 'templates' sub-directory. New XML templates may be copied into this directory and will appear in the Select Template Box the next time the application is invoked.

 The **New** File Button allows creation of a new XML file by specifying a root element name or by generating from a DTD.

 The **Open** File Button allows selection of XML files located on the current file system for opening into *XMLInput*.

 The **Save** Button saves the current XML file back to its original name. Note: This button becomes disabled when a template is loaded to avoid the possibility of overwriting the original template. See [Adding & Saving New Templates](#) for information on how to override this behavior.

 The **SaveAs** Button saves the current XML file to another name specified by the user.

 The **Filter** Button controls in-memory and output options which are described in the section, [Filtering Options Dialog](#).

The **Validate** CheckBox toggles XML validation during the initial opening of the file. When this box is checked, the parser that opens the specified XML file will perform validation based on the Document Type Declaration (DTD) or XML Schema specified in the XML file. Examples of each of these specifications are shown below.


DTD Specification

```
<!DOCTYPE metadata SYSTEM "metadata.dtd">
```

XML Schema Specification

```
<root  
  xmlns:xsi="http://www.w3.org/2000/10/XMLSchema-instance"  
  xsi:noNamespaceSchemaLocation="file:metadata2.xsd">
```

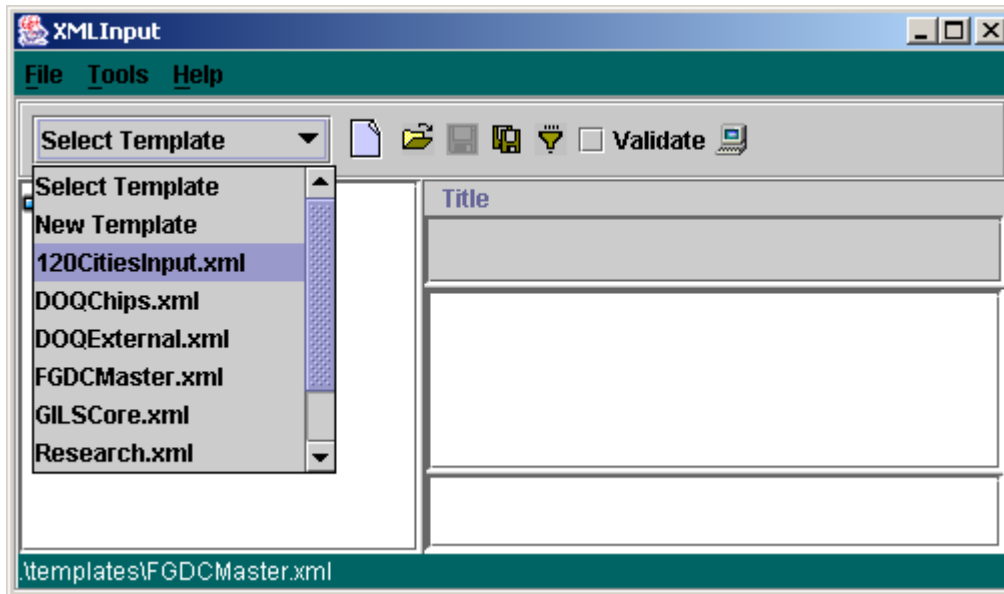
NOTE: This validation only occurs when the file is opened. In-memory validation has not been implemented yet.

 The **System** Button creates a temporary file of the current XML and attempts to open the file using the local system's default browser. This command has been tested on Windows platforms with Internet Explorer as the default browser. Results may vary on other platforms.

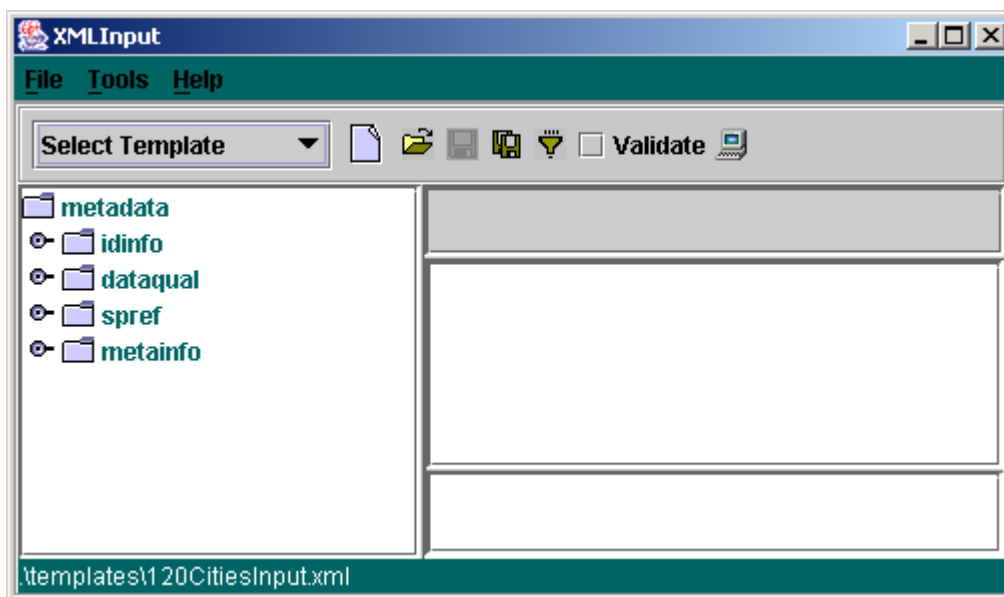
Templates

Selecting Templates



Templates may be selected from the drop-down selection box on the *XMLInput* interface as shown below.



The template is then loaded into the *XMLInput* application as shown below and is ready to be edited. *Note: The save button is disabled to prevent overwriting of the original template. To save the file use the "Save As" button to save the file to another name.*






Adding & Saving New Templates

Any suitable XML file may be used as a template by simply opening the file using the open  button. To include a new template in the Template Selection Box the files must be loaded in the "templates" sub-directory of the *XMLInput* application which is re-read each time the application is started. New templates may be copied to this directory using the native operating system commands or interfaces or the file may be opened in *XMLInput* and saved in the "templates" directory by using the "Save As"  button.

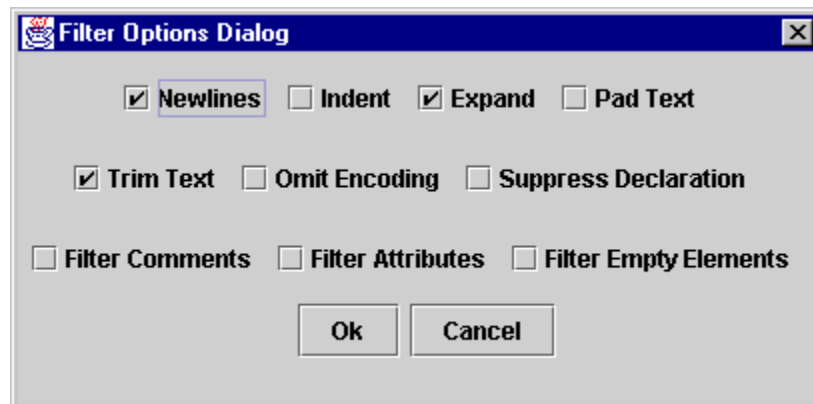
To prevent accidental overwriting of standard templates, templates opened through the Template Selection Box may NOT be saved directly to disk with the same name. Whenever a template is selected from the pull-down box, the "Save" button is disabled.

There are two ways to workaroud this behavior.

- Instead of using the selection box to open a file, use the "Open"  button to select and open an existing template file. The "Save"  button will remain enabled allowing the file to be saved to it's original location.
- Use the "Save As"  button to save the file and use the same name as the original template.

Either of the above methods will cause the "Save" button to remain enabled so subsequent saves can be done as usual.

Anytime the "Save" or "Save As" buttons are selected, the [Filter Options Dialog](#) is presented.



This dialog allows the user to filter certain aspects of the XML file during the save operation. Each of these options are discussed in more detail in the [Filtering Options](#) portion of this document.

Creating New Templates

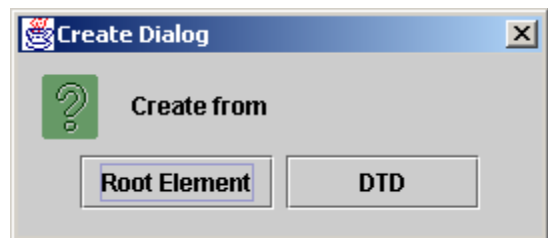
From Existing Master Templates

One way to create a new template is to [select](#) an existing template (FGDCMaster.xml or GILSCore.xml), [delete](#) any elements that are not needed and [save](#) the resulting template to disk for future use. Add comments as appropriate to assist user in accurately completing the required data.

From an Existing DTD

Creating a new template from an existing DTD assures the resulting template conforms to the defining document type. Selecting the 'New Template' option from the pull-down template menu or the 'New' button on the tool panel will display the **Create Dialog**.

Select the 'DTD' button to create a new template based on the existing DTD.

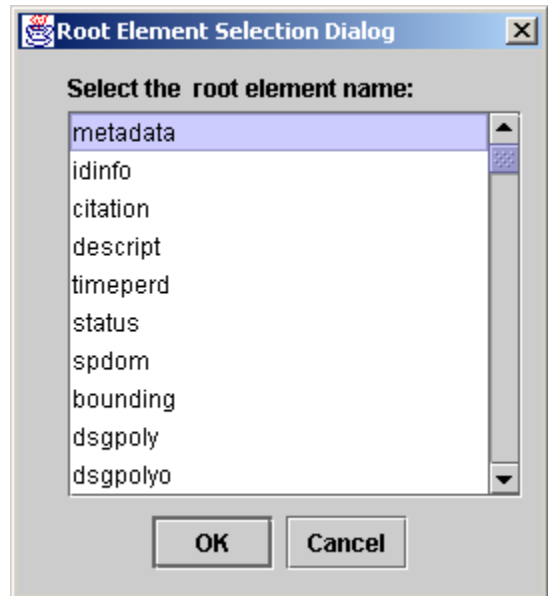


From the **File Selection** menu choose an existing DTD. The 'csdgm2.dtd' is the FGDC DTD. Also included with the *XMLInput* are DTDs for GILS (GILS.dtd) and the ESRI ArcXML (ArcXML.dtd) DTD. Any valid DTD may be selected.

After selecting the DTD file, the application displays the **Root Element Selection Dialog** which contains a list of elements found in the DTD. Select the root element for the DTD. (The order of the list is dependent on the DTD so the root element may not always be on the top of the list. For the FGDC DTD select the 'metadata' element and press the OK button.

A new XML document will be created within XMLInput with all mandatory elements. (Note: Some elements may be mandatory and be part of a list of choices where only one value is possible. By default the application creates the first element found in the list of choices. If this is not the desired element the element must be deleted and the desired element added using the DTD-driven insert function.)

Additional elements may be added to the XML by using the DTD-driven insert function.

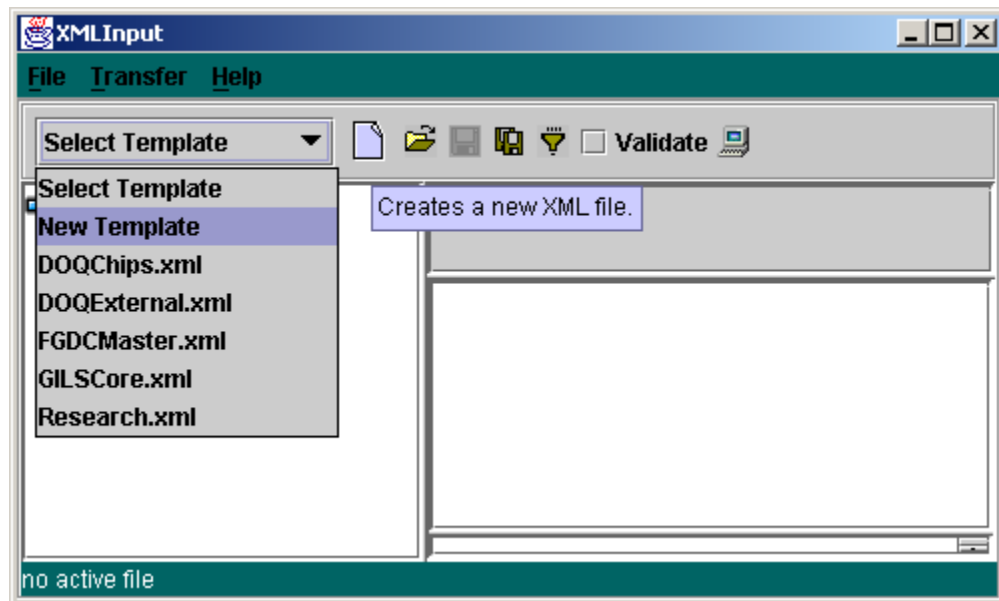


From “Scratch”

Templates may also be created from scratch with only a user supplied root element.

Note: This option is NOT recommended because there are no constraints in place to constrain elements based on context but does allow rapid prototyping of new XML structures.

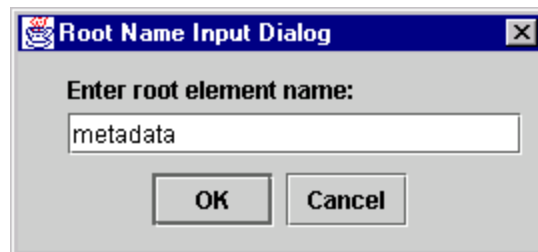
New templates may be created by selecting the “New Template” option from the Select Template Box or the ‘New’ button on the Toolbar.



The Create Dialog should appear. Choose the create from ‘Root Element’ button.



The Root Name Input Dialog will appear asking for the name of this root element.



With the root element named, the user can continue building the new template using the [Insert](#), [Duplicate](#) and [Delete](#) commands found in the Element Tree View [Pop-Up Menu](#).

EDITING

General

Word wrapping occurs in all text areas to make text fit side-to-side in the editing area. A scrollbar will appear when text extends past the bottom of the editing area.

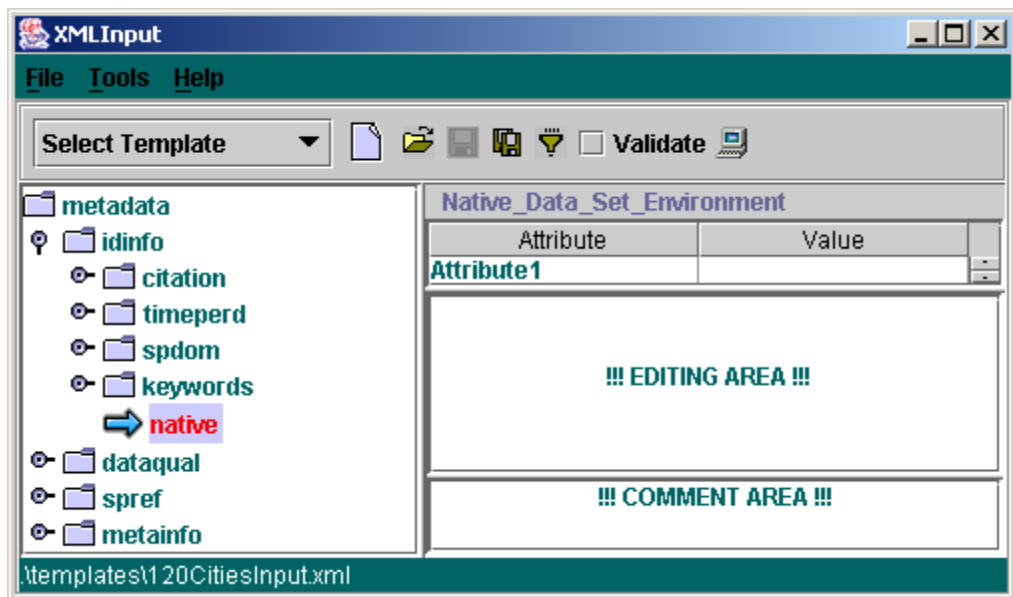
Line returns may be added to create spaces between paragraphs but should generally not be added at the end of each line.

Large blocks of text can be copied into the text areas by using the standard system Copy and Paste key sequences (for example: Ctrl-C to copy and Ctrl-V to paste.)

Text modifications are not applied to the current XML while the text area is still selected. TO APPLY TEXT MODIFICATIONS, CLICK OUTSIDE OF THE AREA BEING EDITED. This can be done by either selecting another element or clicking in another text or attribute area.

Editing Element Text

Element text can be added or modified by selecting the element in the Element Tree View on the left side of the application and entering text into the Element Text Area on the top right.



Adding Comments

Comments may, likewise, be added by selecting the element and entering text into the Comment Area. The comment area can be used to help guide the user in correctly filling out the data. Information about the format of the required text (for instance, MM-DD-YYYY) would be helpful along with information about possible domains. For FGDC data it has been recommended that users include the description found in the Content Standard for Digital Geospatial Metadata such as:

1.2.2 Purpose -- a summary of the intentions with which the data set was developed.

Type: text

Domain: free text

XMLInput

NOTE: TEXT MODIFICATIONS ARE NOT APPLIED TO THE CURRENT XML WHILE THE TEXT AREA IS STILL SELECTED. TO APPLY TEXT MODIFICATIONS, CLICKED OUTSIDE OF THE AREA BEING EDITED. THIS CAN BE DONE BY EITHER SELECTING ANOTHER ELEMENT OR CLICKING IN ANOTHER TEXT OR ATTRIBUTE AREA.

Editing Attributes

Attribute editing is still limited in version 1.6. Attribute *values* may be modified and new attributes inserted. Attributes cannot be deleted at this time and “smart inserts” that look at the DTD for valid attributes has not been implemented.

To add an attribute, click the right mouse button over the element editing area and select “Insert Attribute” from the pop-up menu. The Input dialog will appear asking for the name of the attribute.

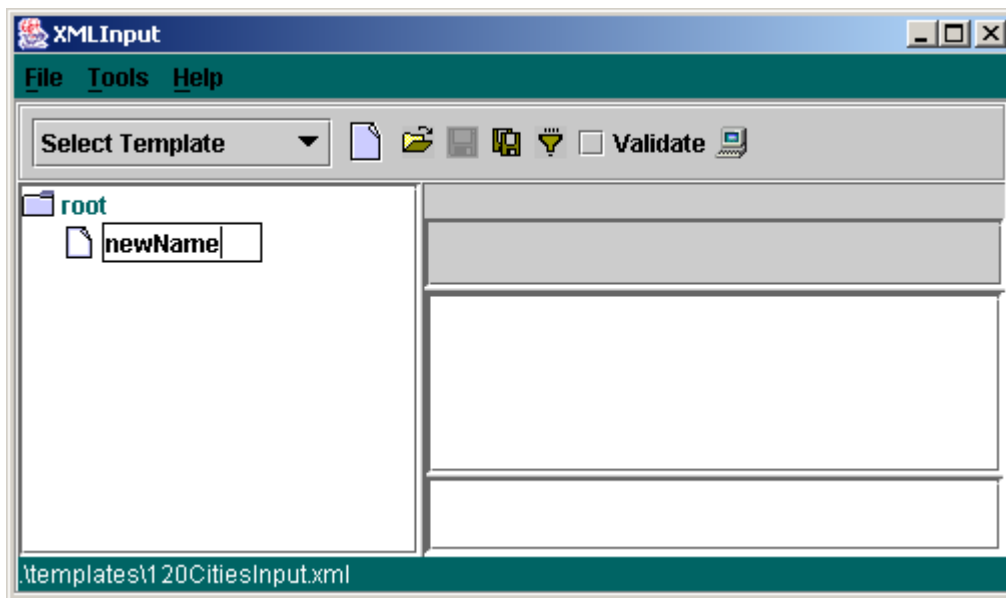
Look for enhancements to this functionality in future releases. To remove attributes use a plain text editor or another XML editor such as Microsoft’s XML Notepad available at:

<http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnxml/html/xmlpaddownload.asp>

Renaming Elements

Elements can be renamed by pressing and holding the left mouse button down when selecting an element in the element tree. This will activate the insert mode (shown below) on the element and the name can then be modified as required. When the name has been changed, press the ‘Enter’ key to apply the changes. To cancel the changes and revert to the original element name, select another element without pressing the ‘Enter’ key.

Element names may not contain spaces, special characters (<, >, !, &) or begin with numerals.



NOTE: NAME CHANGES WILL NOT BE APPLIED UNLESS THE 'ENTER' KEY IS PRESSED AFTER MODIFYING AN ELEMENT NAME. CLICKING OUTSIDE OF THE SELECTED ELEMENT WILL REVERT BACK TO THE ORIGINAL ELEMENT NAME.

Manipulating Elements with the Pop-up Menu

The popup menu is available by right-clicking anywhere within the Element Tree View.

Expand

The Expand option expands the selected element so that all child elements are visible. If no element is selected, the expansion begins at the root (topmost) element and all elements will be made visible.

Collapse

The Collapse command works similar to Expand except all child elements of the selected element are collapsed (closed) so that only the selected element remains visible. If no element is selected, the root node is assumed.

Copy

The Copy command copies the selected element and all its children to memory. The copied element can then be inserted into the XML tree at any location (other than root) using the Paste command.

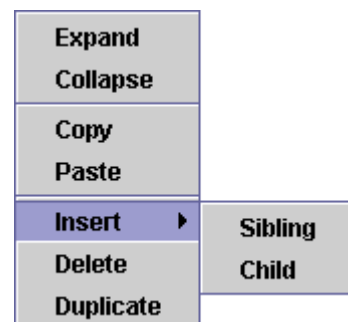
Paste

The Paste command inserts the element placed in memory using the Copy command immediately after the currently selected element.

Insert

There are two methods for inserting elements. *Insert Sibling* creates an element directly after the selected element and on the same indentation level. *Insert Child* creates an element nested within the selected (parent) element. Siblings are always added directly after the selected element. Children are added after the last child if the selected element already has child elements.

When a DTD is defined for the current XML, inserts become context-aware. Sibling and child menus will point to valid elements for the selected context. By default, elements inserted with DTD input includes not only the element selected but all mandatory child elements.



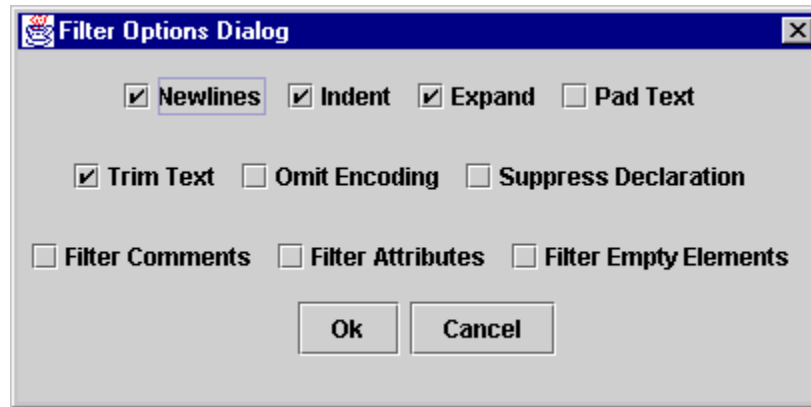
Duplicate

The *Duplicate* command makes a copy of the currently selected element (and all children and content) and inserts it directly after the selected element.

Delete

The *Delete* command removes the currently selected element and all its content including children and all descendents.

Filtering Options Dialog



Certain software applications that import XML do not rigidly support the full XML Specification. A specific flavor of XML is required for successful ingestion into these packages. The Filter Options Dialog is provided to allow control of many of the options that affect the final format of the XML file. Some of these options only apply to the resulting XML file produced by the “Save” or “Save As” commands. A few options apply also to the in-memory XML representation.


NOTE: LOSS OF DATA CAN OCCUR USING SOME OF THESE FILTER OPTIONS. THE USER MUST MAKE SURE HE UNDERSTANDS THE IMPLICATIONS OF EACH OPTION BEFORE SAFELY DEVIATING FROM THE DEFAULTS!

Output Only Filter Options

The Output Only Options apply only to files as they are saved. These options affect the appearance (Newlines, Indent, Expand, Pad Text, Trim Text) or the declarations (Omit Encoding, Suppress Declaration) of the resulting XML text file. The in-memory representation is not affected by these options. Whitespace, and newline characters are the only unrecoverable data which can be lost with these options.

XML Data Filter Options

The XML Filter Options (filter comments, filter attributes, and filter empty elements) apply to both output and in-memory uses. **THESE OPTIONS MUST BE USED WITH EXTREME CAUTION, AS DATA MAY BE IRRECOVERABLY LOST!**

Unlike the Output Only options, these options can apply to the in-memory representation of the XML file currently being edited. This is only true when the “Filter” button  has been selected from the XMLInput toolbar. Use of these options when saving a file will only affect the resulting XML output file – the in-memory representation will not be affected.

On the next page is a table showing descriptions for each of these options.

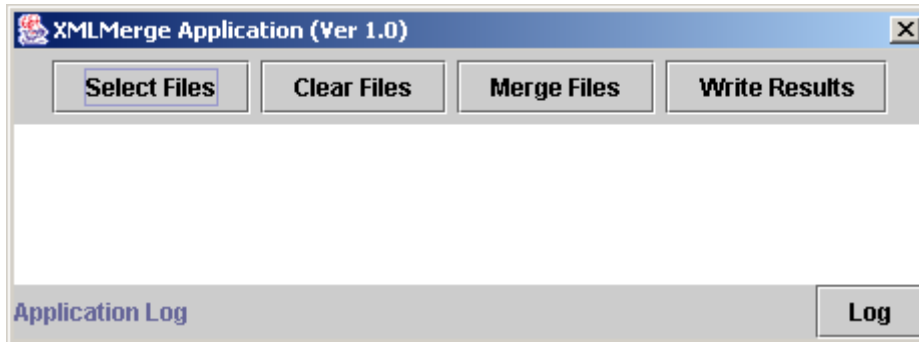
Table 1 - Filter Options Summary

OPTION	DESCRIPTION	DEFAULT
<i>OUTPUT ONLY OPTIONS</i>		
Newlines	If checked, the output document will contain newlines after each element. Unchecking this option will save space but be less human readable.	TRUE
Indent	If checked, the output document will be indented resulting in a more readable document. The default is set to false to make the resulting file smaller.	TRUE
Expand	If checked, empty elements are expanded from the compressed designation <tagName/> to the more lengthy <tagName></tagName>.	TRUE
Pad Text	If checked, text immediately preceded by or followed by an element will be "padded" with a single space.	FALSE
Trim Text	If checked, all extra whitespace will be trimmed from the element text. Checking this will generally produce a cleaner output but does remove some of the original document's whitespace.	TRUE
Omit Encoding	If checked, the XML declaration (<?xml version="1.0" encoding="UTF-8"?>) will be omitted in the output file.	FALSE
Suppress Declaration	If checked, the XML declaration (<?xml version="1.0"?>) will be suppressed in the output file.	FALSE
<i>XML FILTER OPTIONS (IN-MEMORY & OUTPUT)</i>		
Filter Comments	If checked, will remove all comments. Comments are not understood by some XML validation software.	FALSE
Filter Attributes	If checked, will remove all attributes.	FALSE
Filter Empty Elements	If checked, will remove all empty elements (elements that do not contain text).	FALSE

APPENDIX A – MERGING XML FILES

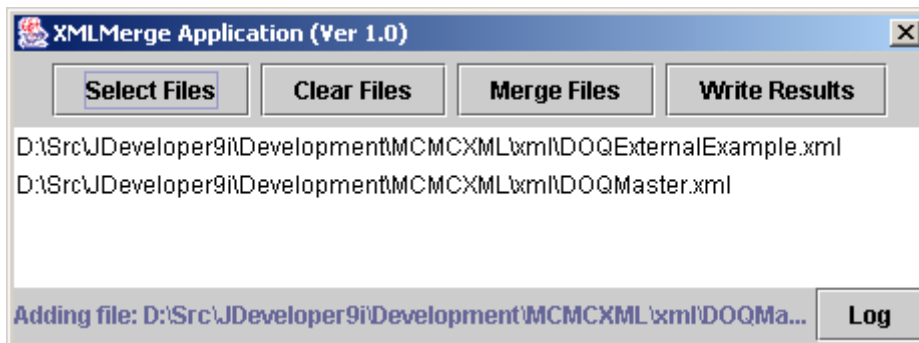
As of version 1.6, a tool was added to the XMLInput to allow merging of similar XML files. A separate application called XMLMerge performs this function and is integrated into the XMLInput 'Tools' menu. The intended purpose of this tool is to allow merging of general metadata that applies to a group of products with file-specific metadata that is different for each file.

The interface looks like this:



Selecting Files

Users may select two or more files using the 'Select Files' button. As files are selected their pathnames appear in the list box as shown below. Multiple files can be selected from the same directory by holding down the 'Cntr' key while selecting each file. Multiple files may also be selected by repeatedly using the 'Select Files' button. All files may be removed from the list by pressing the 'Clear Files' button.



The root element of each file must be the same. If a DTD is specified in the first XML file and can be located on the system, this DTD will be used to assure conformance with the document structure. If no DTD is specified or present results may vary.

Merge Files

Once the files have been selected use the 'Merge Files' button to begin the merge process. Clicking on the 'Log' button displays the process messages.

Write Results

XMLInput

Use the 'Write Results' button to save the merge results to a new filename.