



Tools to facilitate metadata

Inigo San Gil
LTER Network Office
Albuquerque, NM





Metadata tools -- Outline

- Introduction, broad remarks
- Tools:
 - Editors
 - Databases
 - Custom tool
 - Comparative matrix
- Metadata standard crosswalks





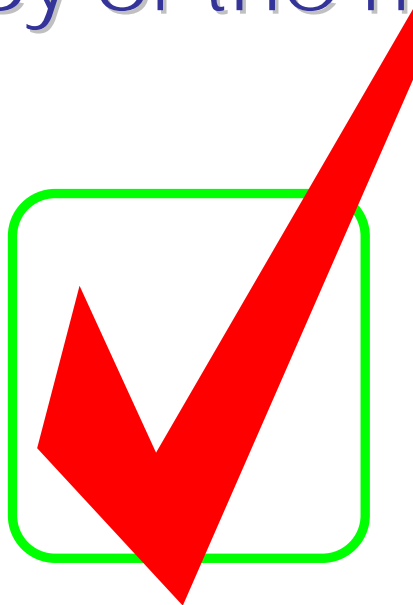
Introduction





On Metadata Quality

No tool can check the
accuracy of the metadata












Selecting A Tool

Consider ...

- *Cost*
- *Platform*
- *User support*
- *Capabilities*
- *Data-entry method*
- *Level of compliance*



Metadata Tool Matrix

	Open Source Proprietary	On-line Stand Alone	Platform
XML Spy	Home ed - \$0 Professional ~ \$1000 Enterprise ~ \$3000	Stand Alone	
Morpho	Free	Stand Alone On-line*	Any
oxygen	Academic \$50 Professional \$200	Stand Alone	Any
Notepad, Word		Stand Alone	
Emacs, vi, ...	Free	Stand Alone	Unix 
Databases	SQLs { \$0 - \$\$\$} Oracle {\$50* - \$1000+} MSAccess 	Stand Alone On-line	Mixed
NBII tools	Free	Stand Alone On-line	Mixed 



Morpho – Editor & Entry tool

Refer to the next talk by Will Tyburczy





XML Editors -- XMLSpy

XMLSpy interface showing a schema design view. The main workspace displays a hierarchical structure of XML elements and attributes. The left pane shows a list of examples, including 'eml', 'dataset', and 'attributes'. The right pane shows a list of components, including 'element', 'access', 'attribute', 'attributeList', 'citation', 'dataTable', 'doc:description', 'doc:example', 'doc:lineage', 'doc:module', 'doc:moduleDocs', 'doc:summary', 'doc:tooltip', 'ds:dataset', 'eml', 'methods', 'otherEntity', 'party', 'physical', 'projectionList', 'prot:protocol', 'researchProject', 'spatialRaster', 'spatialReference', and 'spatialVector'. The bottom pane shows a list of details, including 'name', 'isRef', 'minOcc', 'maxOcc', 'type', 'content', 'derivedBy', 'mixed', 'nullable', 'block', 'form', and 'id'.

The main workspace displays a hierarchical structure of XML elements and attributes. The left pane shows a list of examples, including 'eml', 'dataset', and 'attributes'. The right pane shows a list of components, including 'element', 'access', 'attribute', 'attributeList', 'citation', 'dataTable', 'doc:description', 'doc:example', 'doc:lineage', 'doc:module', 'doc:moduleDocs', 'doc:summary', 'doc:tooltip', 'ds:dataset', 'eml', 'methods', 'otherEntity', 'party', 'physical', 'projectionList', 'prot:protocol', 'researchProject', 'spatialRaster', 'spatialReference', and 'spatialVector'. The bottom pane shows a list of details, including 'name', 'isRef', 'minOcc', 'maxOcc', 'type', 'content', 'derivedBy', 'mixed', 'nullable', 'block', 'form', and 'id'.

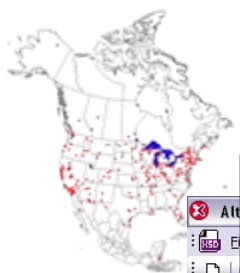
The main workspace displays a hierarchical structure of XML elements and attributes. The left pane shows a list of examples, including 'eml', 'dataset', and 'attributes'. The right pane shows a list of components, including 'element', 'access', 'attribute', 'attributeList', 'citation', 'dataTable', 'doc:description', 'doc:example', 'doc:lineage', 'doc:module', 'doc:moduleDocs', 'doc:summary', 'doc:tooltip', 'ds:dataset', 'eml', 'methods', 'otherEntity', 'party', 'physical', 'projectionList', 'prot:protocol', 'researchProject', 'spatialRaster', 'spatialReference', and 'spatialVector'. The bottom pane shows a list of details, including 'name', 'isRef', 'minOcc', 'maxOcc', 'type', 'content', 'derivedBy', 'mixed', 'nullable', 'block', 'form', and 'id'.



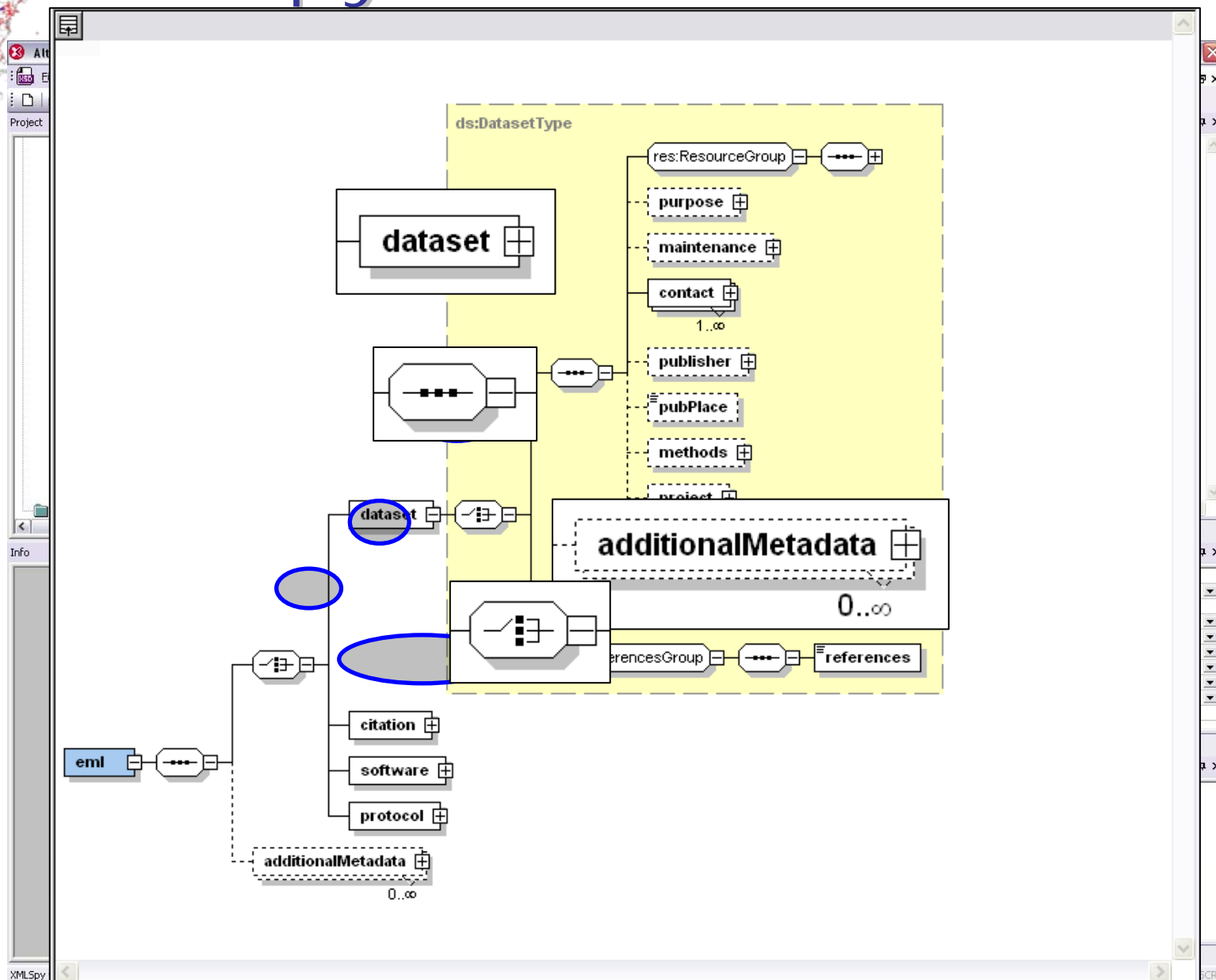
Editors -- XMLSpy

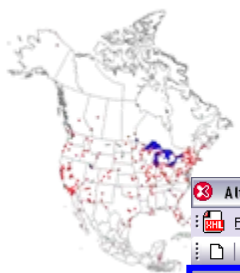
- Why an XML editor?
 - To create, edit, correct, or simply view XML
 - Schema inspection
 - Schema validation
- Working with XML to create basic EML
- Validation of EML





XMLSpy -- Schema View





XMLSpy -- Document View

Altova XMLSpy - [knb-lter-sev.1.1.xml]

File Edit Project XML

Project

- myEML *
- XML Files
 - knb-lter-sev.1.1.xml
- XSL Files
- XQuery Files
- HTML Files
- DTD/Schemas
 - C:\eml-2.0.1
- Entities

Info

Element Model

Model	dataset
choice	

Text Grid Schema/WSDL Authentic Browser

knb-lter-sev.1.1.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XMLSpy v2005 sp1 U (http://www.xmlspy.com) by Mark Servilla (UNM Dept of Biology, LTER Network Office) -->
<eml:eml xmlns:eml="eml://ecoinformatics.org/eml-2.0.1" xmlns:ds="eml://ecoinformatics.org/dataset-2.0.1" xmlns:doc="eml://ecoinformatics.org/documentation-2.0.1" xmlns:cit="eml://ecoinformatics.org/literature-2.0.1" xmlns:prot="eml://ecoinformatics.org/protocol-2.0.1" xmlns:res="eml://ecoinformatics.org/resource-2.0.1" xmlns:sw="eml://ecoinformatics.org/software-2.0.1" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="eml://ecoinformatics.org/eml-2.0.1 C:\eml-2.0.1\eml.xsd" packageid="K" >
  <dataset>
    <title>Seville LTER NPP Quad</title>
    <creator>
      <individualName>
        <givenName>Este</givenName>
        <surName>Muldaiv</surName>
      </individualName>
      <organizationName>New Mexico State University</organizationName>
      <organizationName>Department of Biology</organizationName>
      <address>
        <deliveryPoint>
          <city>Albuquerque</city>
          <administrativeArea>New Mexico</administrativeArea>
          <postalCode>87131</postalCode>
          <country>USA</country>
        </deliveryPoint>
        <electronicMailAddress>este@nmsu.edu</electronicMailAddress>
      </address>
    </creator>
    <abstract>
      <section>
        <para>Net primary production (NPP) is the rate at which plants and other autotrophs produce organic matter through photosynthesis, minus the amount of carbon consumption and fixation. NPP is a key component of the carbon cycle and is an important indicator of ecosystem health. The Seville LTER NPP Quad is a spatial and temporal response of the NPP at the Seville LTER to monitor net primary production in a grassland dominant grassland, juniper-savannah. NPP is an important component in estimating NPP, the described change in plant mass, including any above ground biomass production (ANPP) and below ground biomass production (BNPP). In addition, volumetric measurements of biomass and volume.
      </para>
      </section>
    </abstract>
    <keywordSet>
      <keyword>ANPP</keyword>
      <keyword>biomass</keyword>
      <keyword>Seville</keyword>
      <keyword>LTER</keyword>
    </keywordSet>
    <contact>
      <individualName>
        <surName>
      </individualName>
    </contact>
  </dataset>
</eml:eml>
```

Elements

- alternateIdentifier
- shortName
- title
- references
- creator

Attributes

Entities

Ent	amp
Ent	apos
Ent	gt
Ent	lt
Ent	quot

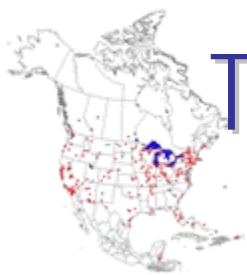
project

- access
- dataTable
- spatialRaster
- spatialVector
- storedProcedure
- view
- otherEntity

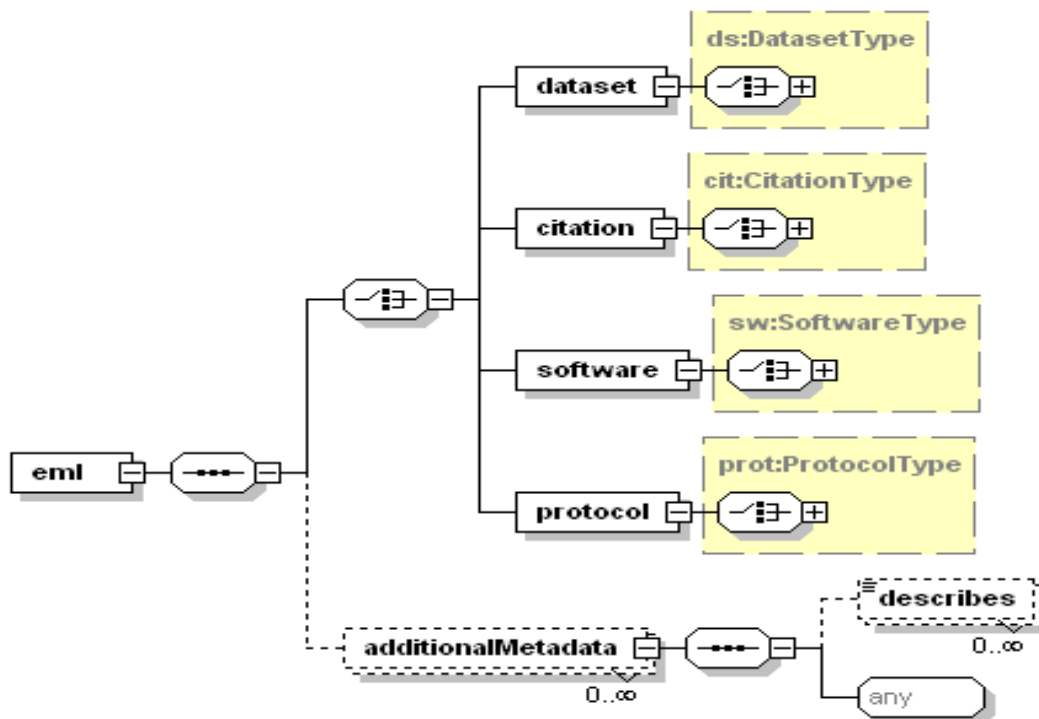
Seville LTER NPP Quad

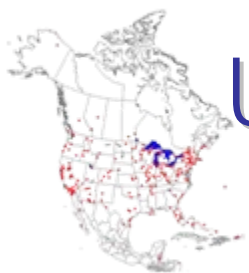
Net primary production (NPP) is the rate at which plants and other autotrophs produce organic matter through photosynthesis, minus the amount of carbon consumption and fixation. NPP is a key component of the carbon cycle and is an important indicator of ecosystem health. The Seville LTER NPP Quad is a spatial and temporal response of the NPP at the Seville LTER to monitor net primary production in a grassland dominant grassland, juniper-savannah. NPP is an important component in estimating NPP, the described change in plant mass, including any above ground biomass production (ANPP) and below ground biomass production (BNPP). In addition, volumetric measurements of biomass and volume.

NSF

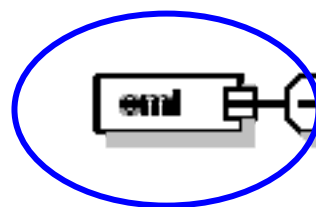
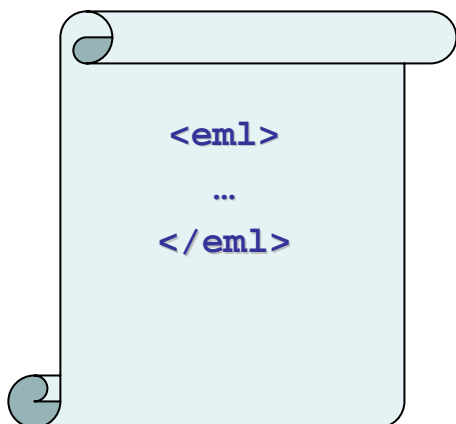


Top level EML Schema structure

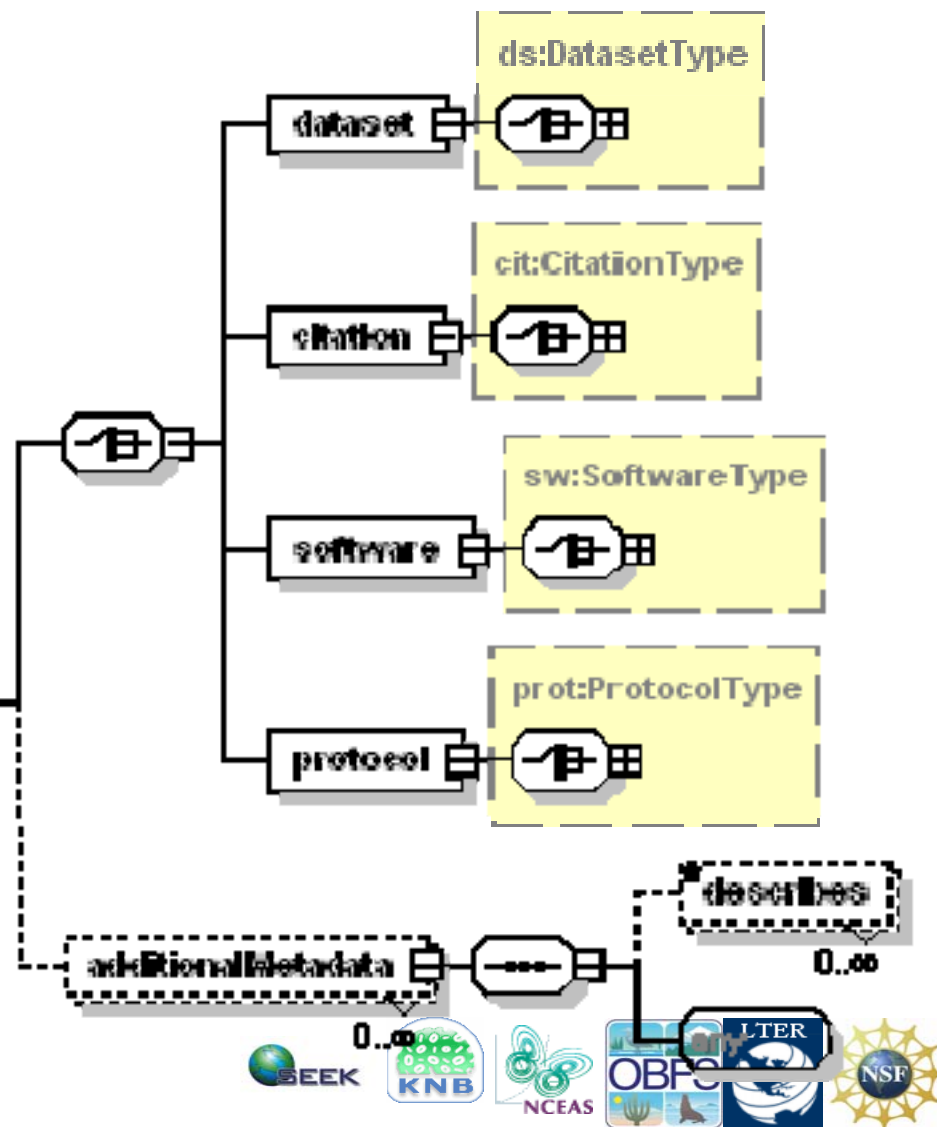




Understanding XMLSpy diagrams



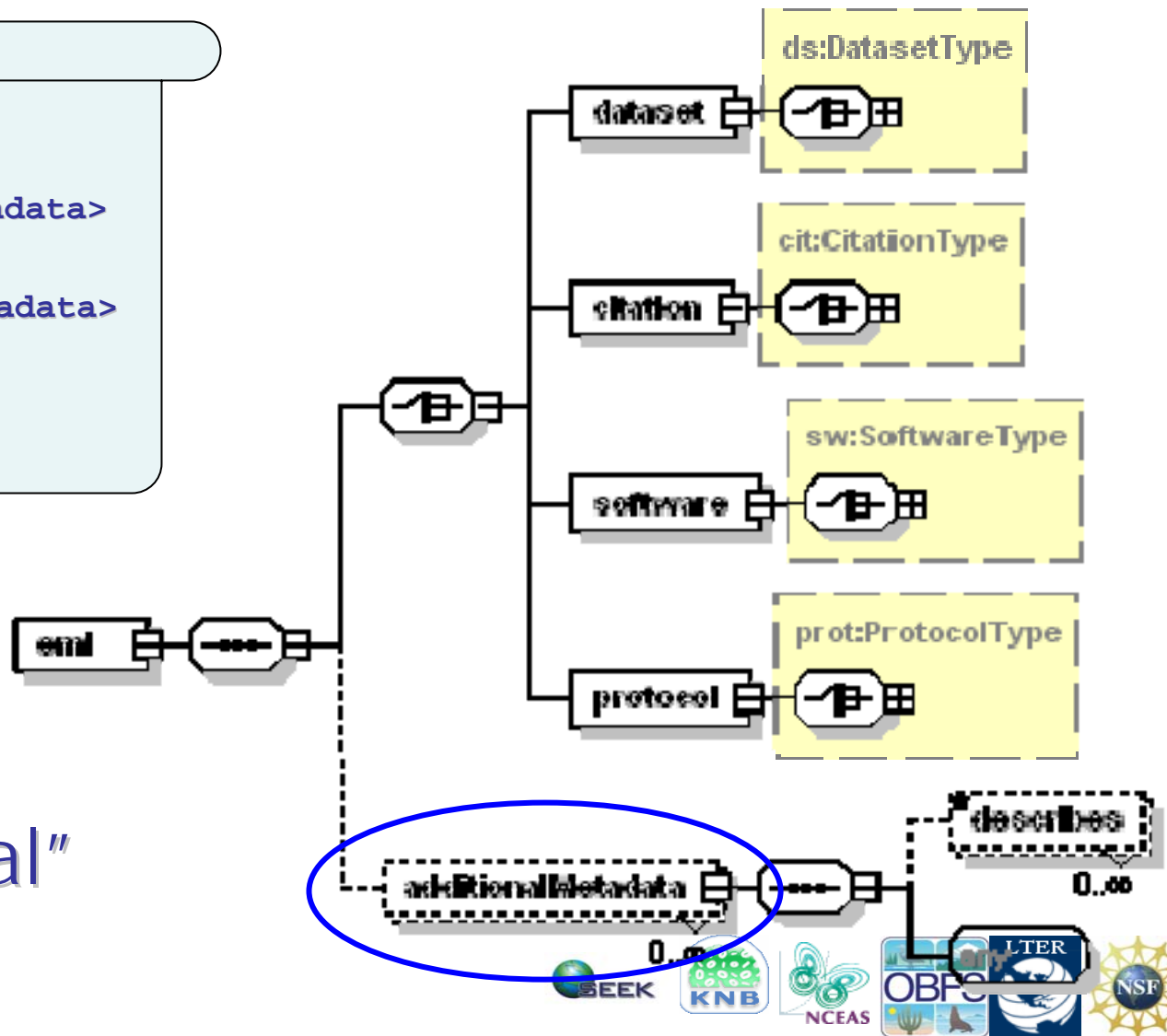
"root
element"



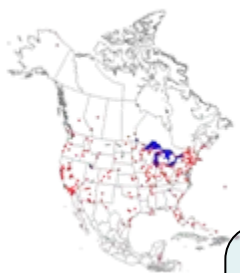


XMLSpy -- Diagrams

```
<eml>  
...  
<additionalMetadata>  
...  
</additionalMetadata>  
</eml>
```

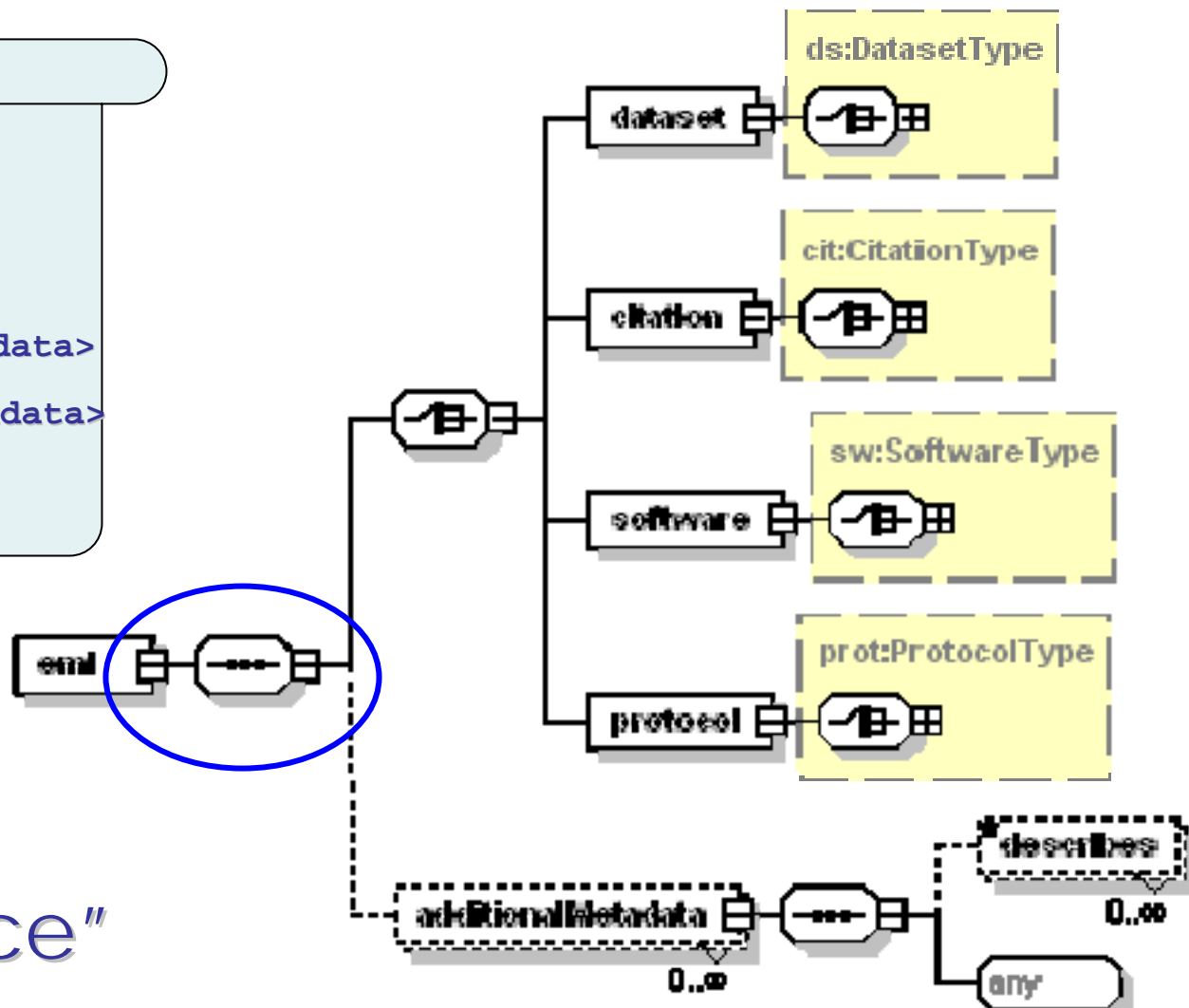


"optional"



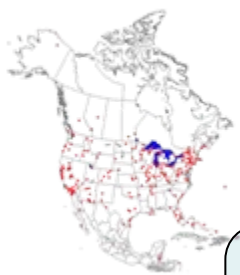
XMLSpy -- Diagrams

```
<eml>
  <dataset>
    ...
  </dataset>
  <additionalMetadata>
    ...
  </additionalMetadata>
</eml>
```



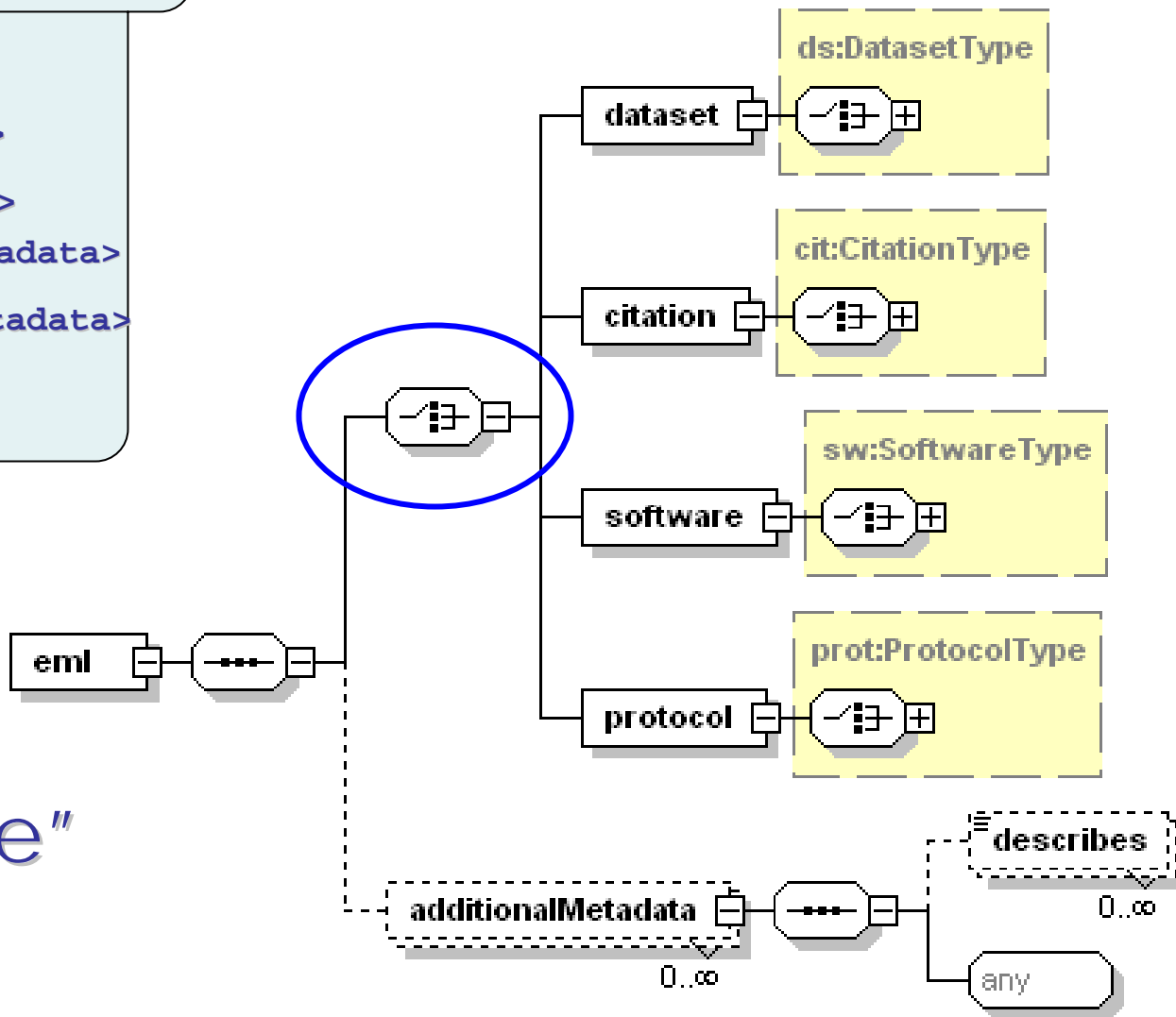
"sequence"



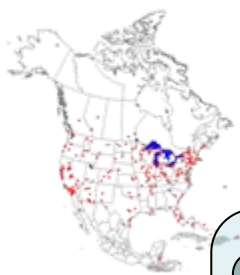


XMLSpy -- Diagrams

```
<eml>
  <dataset>
    ...
  </dataset>
  <additionalMetadata>
    ...
  </additionalMetadata>
</eml>
```

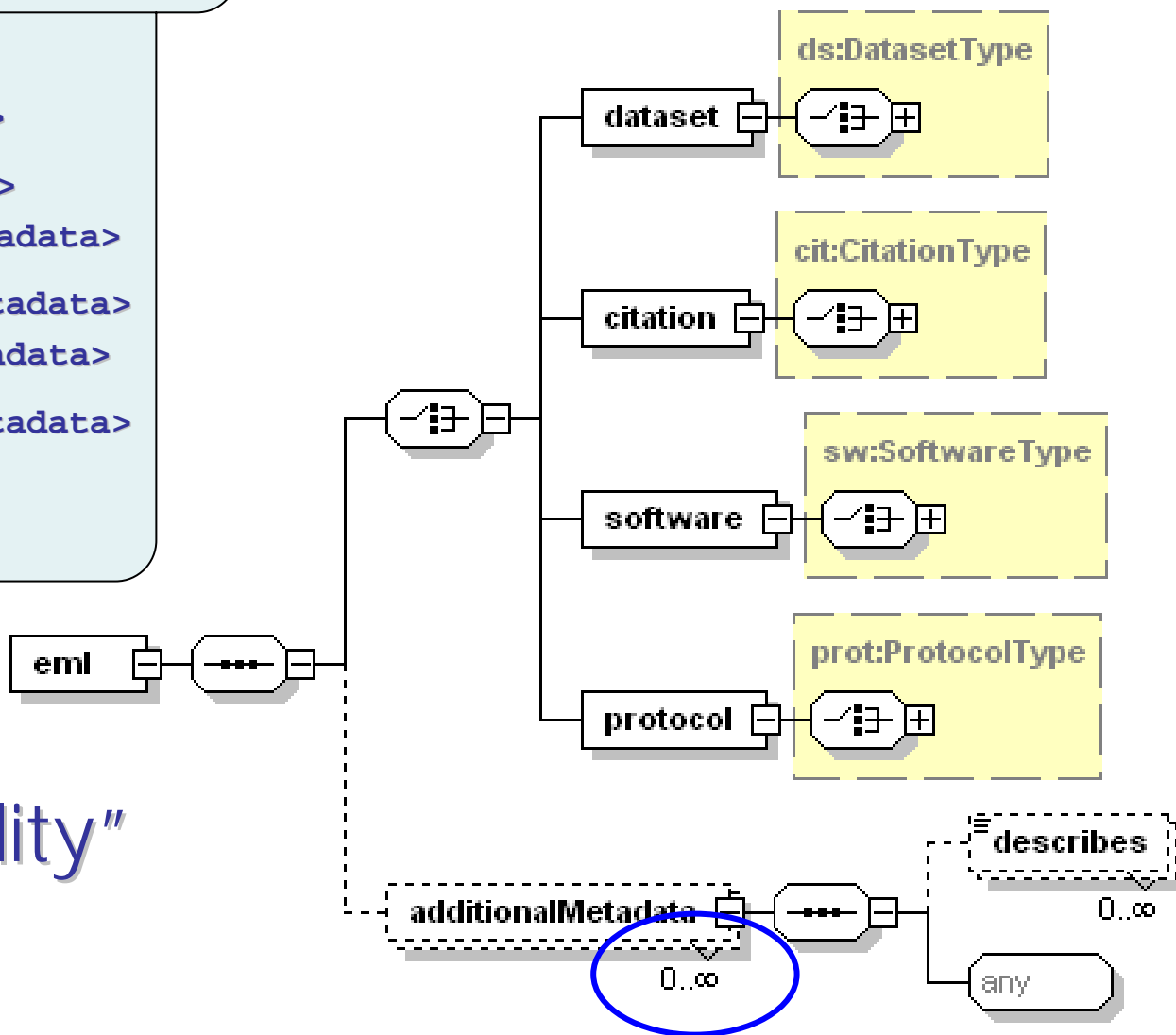


"choice"

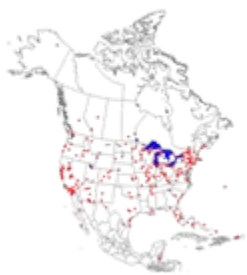


XMLSpy -- Diagrams

```
<eml>
  <dataset>
    ...
  </dataset>
  <additionalMetadata>
    ...
  </additionalMetadata>
  <additionalMetadata>
    ...
  </additionalMetadata>
</eml>
```



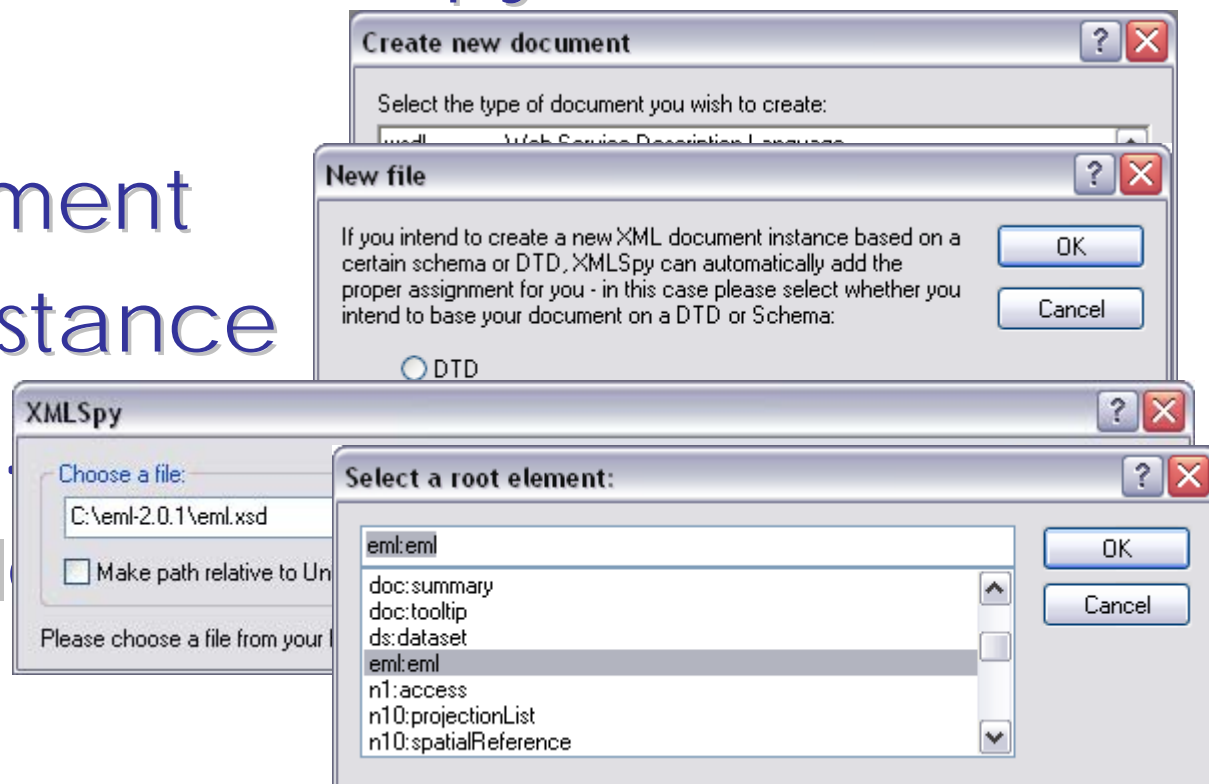
"cardinality"



Begin A New EML Document



- Double-click on XMLSpy icon
- File – New
- XML Document
- Schema Instance
- C:\eml-2.0\eml.xsd
- EML root element





Vestige of an EML Document

[Untitled1.xml *]

File XML DTD/Schema Schema design XSL/XQuery Authentic Convert View Browser WSDL SOAP Tools Window Help

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <eml:eml xmlns:eml="eml://ecoinformatics.org/eml-2.0.1" xmlns:ds="eml://ecoinformatics.org/dataset-2.0.1"
  xmlns:doc="eml://ecoinformatics.org/documentation-2.0.1" xmlns:cit="eml://ecoinformatics.org/literature-2.0.1"
  xmlns:prot="eml://ecoinformatics.org/protocol-2.0.1" xmlns:res="eml://ecoinformatics.org/resource-2.0.1" xmlns:sw
    ="eml://ecoinformatics.org/software-2.0.1" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="eml://ecoinformatics.org/eml-2.0.1
    C:\eml-2.0.1\eml.xsd" packageId="" system="">
3   <dataset>
4     <title/>
5     <creator>
6       <individualName>
7         <surName/>
8       </individualName>
9     </creator>
10    <contact>
11      <individualName>
12        <surName/>
13      </individualName>
14    </contact>
15  </dataset>
16 </eml:eml>
```

Elements

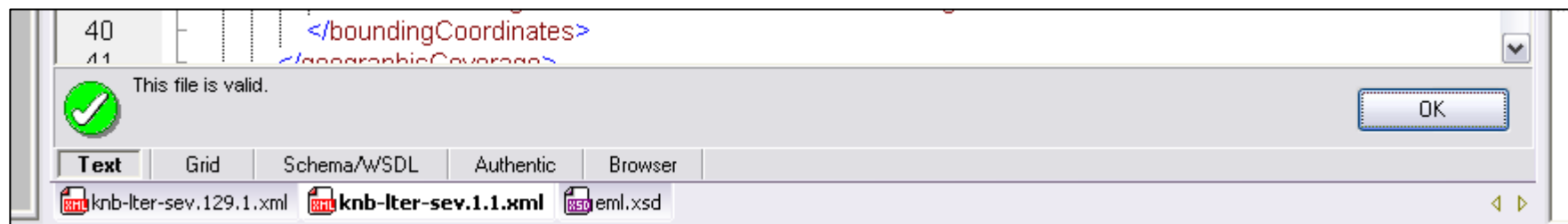
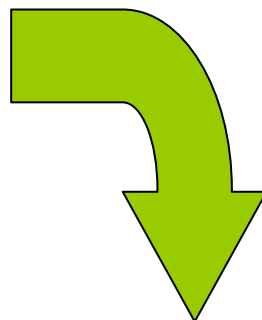
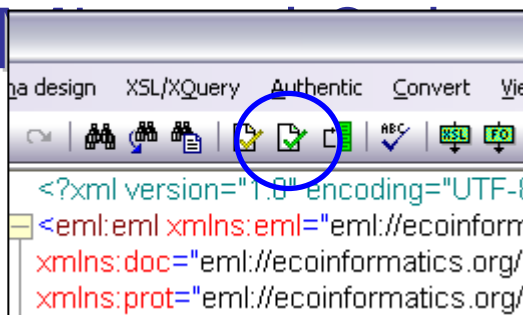
- cit:citation
- doc:descriptio
- doc:example
- doc:lineage
- doc:moduleD
- doc:summary
- doc:tooltip
- ds:dataset
- eml:eml
- n10:projection
- n10:spatialRe
- n11:spatialVe
- n12:storedPro
- n13:text
- n14:view
- n1:access
- n2:attribute
- n2:attributeLis
- n3:dataTable
- n4:otherEntity
- n5:methods
- n6:party
- n7:physical





Validation

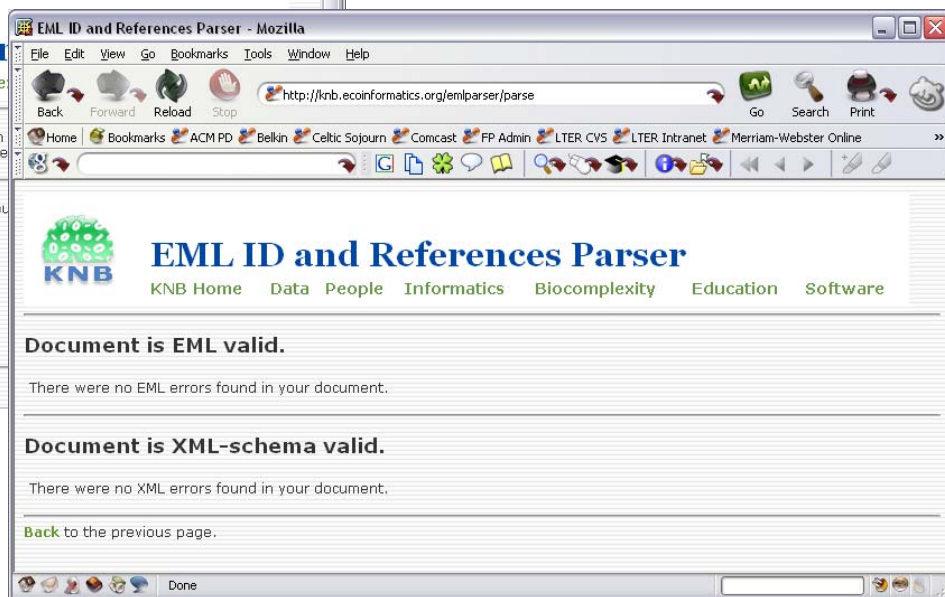
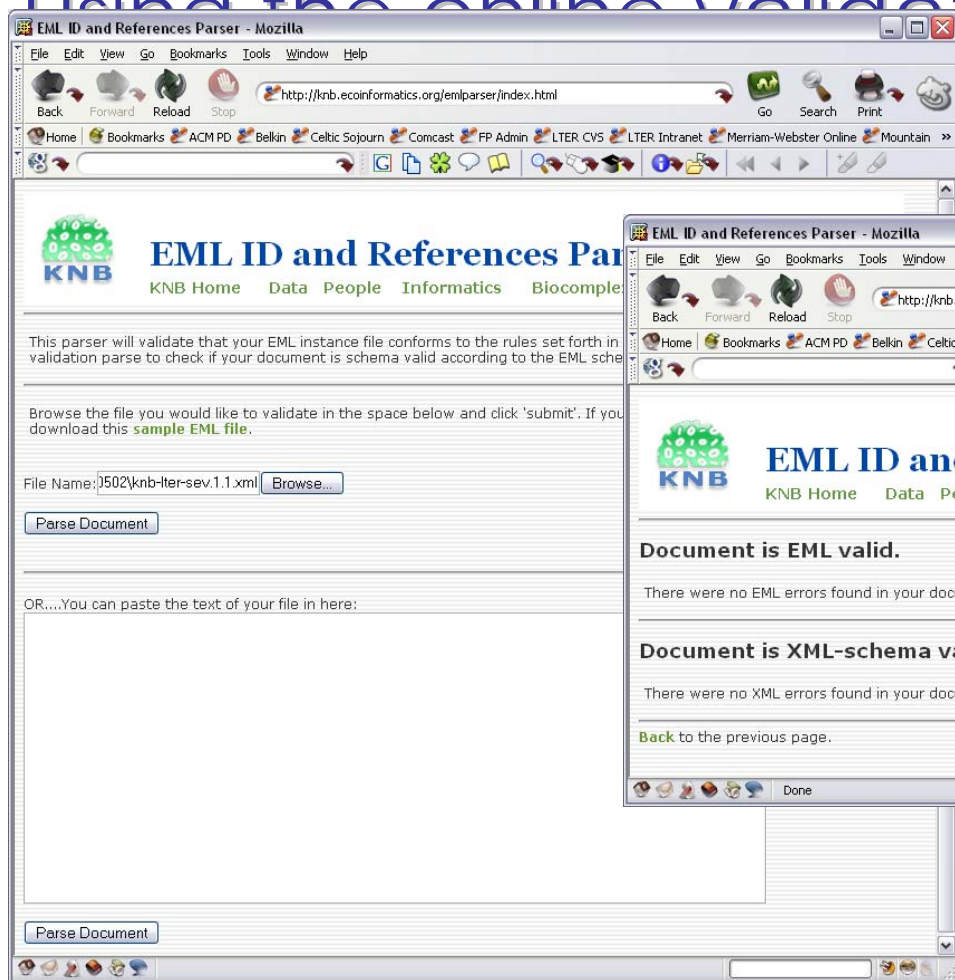
- Using XMLSpy to check for well-formed XML Schema validation





Validation cont.

- Using the online validation service at <http://knb.ecoinformatics.org/emlparser/index.html>



XML Editors – oXygen

The screenshot displays the oXygen XML Editor interface with the Logical Model View of an XML Schema. The main workspace shows a hierarchical diagram of the schema structure. The left pane shows the Project Explorer with a file named 'newProject.xpr'. The bottom-left pane shows the Outline view, listing the schema components and their namespaces. The bottom-right pane shows the Full Model View, displaying the XML Schema definition (XSD) code for the 'eml' element. The right pane shows the Attributes view, listing the attributes of the selected element.

Project Explorer: newProject.xpr

Outline:

- http://www.w3.org/2001/XMLSchema
- "eml://ecoinformatics.org/documentation-2.0.1"
- "eml://ecoinformatics.org/dataset-2.0.1"
- "eml://ecoinformatics.org/literature-2.0.1"
- "eml://ecoinformatics.org/software-2.0.1"
- "eml://ecoinformatics.org/protocol-2.0.1"
- "eml://ecoinformatics.org/resource-2.0.1"

Full Model View:

```
<?xml version='1.0' encoding='UTF-8'>
<eml>
  <appinfo>
    <doc:tooltip>Ecological Metadata
    <doc:summary>A collection of EML metadata
    <doc:description>The "eml" element allows
  </appinfo>
  <dataset>
    <DatasetType>
      <appinfo>
        <doc:tooltip>Dataset
        <doc:summary>DatasetType is the base type for
        <doc:description>DatasetType is the base type f
      </appinfo>
      <res:ResourceGroup>
        <purpose>
          0..1
        </purpose>
        <maintenance>
          0..1
        </maintenance>
        <contact>
          1..∞
        </contact>
        <publisher>
          0..1
        </publisher>
        <pubPlace>
          0..1
        </pubPlace>
        <methods>
          0..1
        </methods>
      </res:ResourceGroup>
    </DatasetType>
  </dataset>
</eml>
```

Attributes:

Name	Value
doc:description	eml://ecoinformatics.org/documentation-

Model:

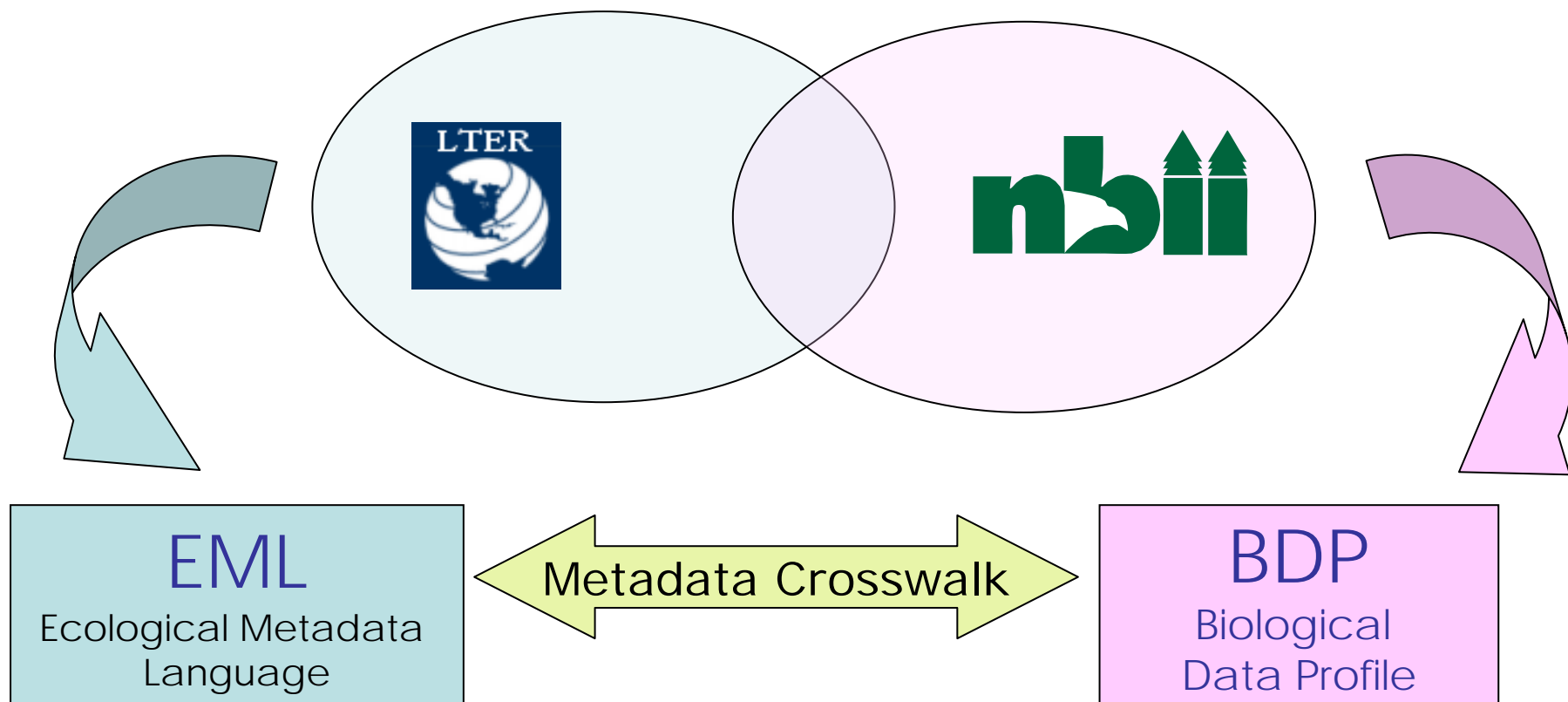
doc:description
Namespace : eml://ecoinformatics.org

Components:

An XML Schema annotation is not available for the current selection.



Metadata Crosswalks





NBII Tools – TKME Editor

Select
elements
from list

Enter
responses in
window

7 Tkme: J:/Metadata/clearinghouse/farwest/fre00176.txt

File Edit View Add Snippets Help

Metadata

- Identification_Information
- Citation
 - Citation_Information
 - Originator
 - Publication_Date
 - Title**
- Geospatial_Data_
- Description
 - Abstract
 - Purpose
 - Supplemental_Information
 - Time_Period_of_Content
 - Time_Period_Information
 - Range_of_Dates/Times
 - Beginning_Date
 - Ending_Date
 - Currentness_Reference
- Status
 - Progress
 - Maintenance_and_Update
- Spatial_Domain
 - Bounding_Coordinates
 - West_Bounding_Coordinate
 - East_Bounding_Coordinate
 - North_Bounding_Coordinate
 - South_Bounding_Coordinate

Mount Rainier Lentic Amphibian Survey: Amphibian Habitat

Tkme: Tk Metadata Editor by Peter Schweitzer (USGS)





Metadata Collection tool

Enter data
Section by
Section

Section 1 Identification_Information (barrier.shp)

Section Help Section Example

1.1 Citation

8 Citation_Information

8.4 Title: Western Oregon Stream Barriers

8.5 Edition:

8.1 Originators: (separate with a semi-colon) presty Ecosystem Research (CFER)

8.2 Publication Date (YYYYMMDD) 2001

8.8.1 Publication_Place:

8.8.2 Publisher:

8.10 Online_Linkage:

8.11.4 Larger_Work_Title:

8.11.1 Larger_Work_Originator:

8.11.2 Larger_Work_Publication_Date:

8.11.8.1 Larger_Work_Publication_Place:

8.11.8.2 Larger_Work_Publisher:

8.11.8.10 Larger_Work_Online_Linkage:

Retrieve Citation Information
Retrieve data from existing file.

Save Citation Information
Create new file.

Append Citation Information
Add new data to existing file.

Default file is citation.dbf.

Exit Clear Form Continue Section 1

Responses can be saved
and managed as .dbf files





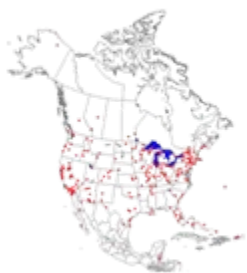
Metavist 2005 – USDA Forest Service

North Central Research Station

<http://ncrs.fs.fed.us/pubs/viewpub.asp?key=2737>

- File based metadata records
- Form entry
- Utilizes xml tags, schema, and stylesheets
- Supports CSDGM and Biological Data Profile





Metavist 2005 – USDA Forest Service

Metavist — Schweitzer_oceanography.xml

File Options Help

Identification Data Quality Spatial Data Org Spatial Reference Entity & Attribute Distribution Metadata Ref

Basic Info Spatial Domain Keywords Taxonomy Access Analytical Tools Miscellaneous

Citation for the data set

Author(s) Schweitzer, Peter N. Publication Date 1993

Title Modern Average Global Sea-Surface Temperature Edit Citation

Abstract

The data contained in this data set are derived from the NOAA Advanced Very High Resolution Radiometer Multichannel Sea Surface Temperature data (AVHRR MCSST), which are obtainable from the Distributed Active Archive Center at the Jet Propulsion Laboratory (JPL) in Pasadena, Calif. The JPL tapes contain weekly images of SST from October 1981 through December 1990 in nine regions of the world ocean: North

Purpose

The purpose of this data set is to provide paleoclimate researchers with a tool for estimating the average seasonal variation in sea-surface temperature (SST) throughout the modern world ocean and for estimating the modern monthly and



MetaScribe – NOAA Coastal Services Center

<http://www.csc.noaa.gov/metadata/metascribe/>

This software tool takes advantage of the fact that, within a collection of records for a given data type, the records are generally very similar in content, with only a few fields or phrases changing from one record to the next.

MetaScribe delivers the greatest savings when used to create a collection of similar records.

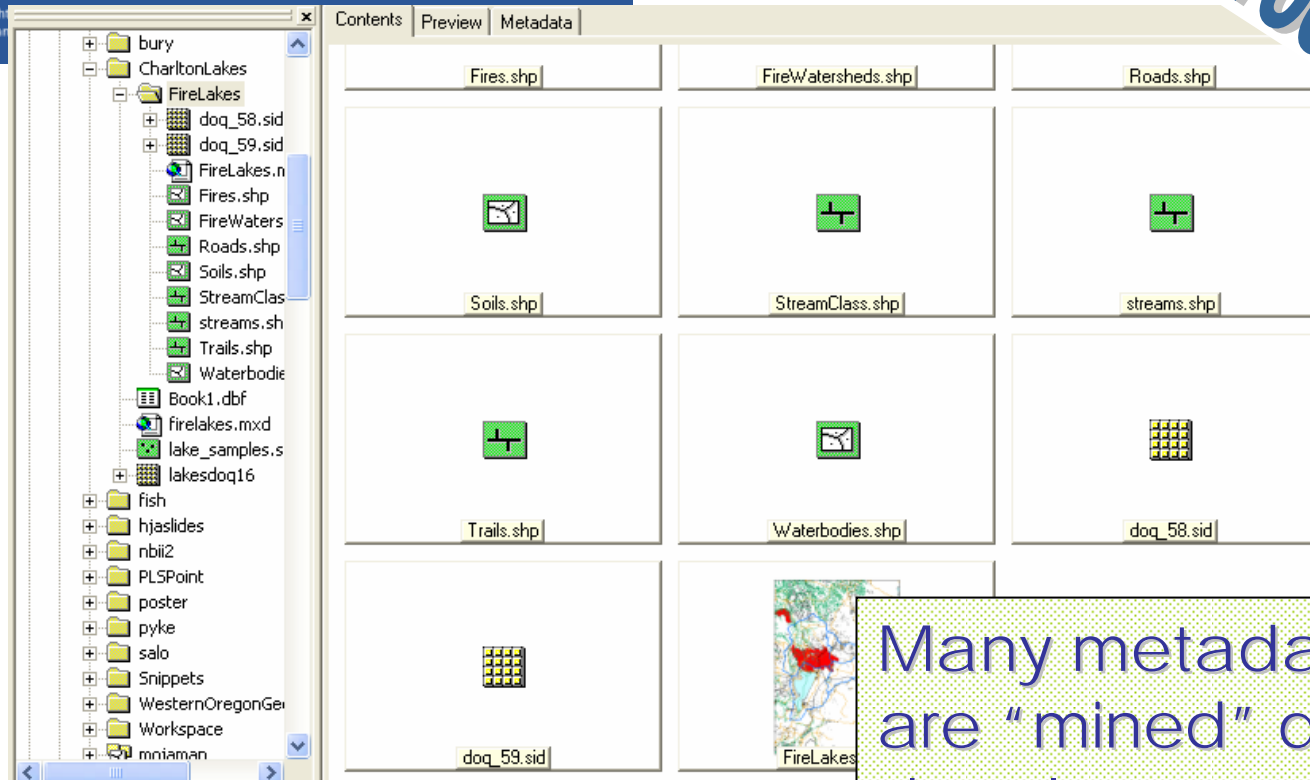
Think template.



ArcGIS 9

ArcCatalog™

Metadata Mining Tools



Many metadata elements are "mined" directly from data layer properties.





Menu

Metadata Tools

Stylesheets

The Catalog

File Edit View Go Tools Help

Location: S:\Projects\GIS\Workspace\larsen\lakesdb\Lakes.mdb\lk_snv205

Stylesheet: FGDC

Contents Preview Metadata

Western Mountain Lakes - Sierra Nevada

Data format: Personal Geodatabase Feature Class

File or table name: lk_snv205

Coordinate system: Geographic

Theme keywords: lakes, limnology, mountain lakes, natural lakes

Abstract: This dataset is an ArcGIS polygon shapefile of mountain lakes in Sierra Nevada ecoregion of the western United States. The data are in geographic coordinates based on the GRS1980 spheroid and datum NAD83.

FGDC and ESRI Metadata:

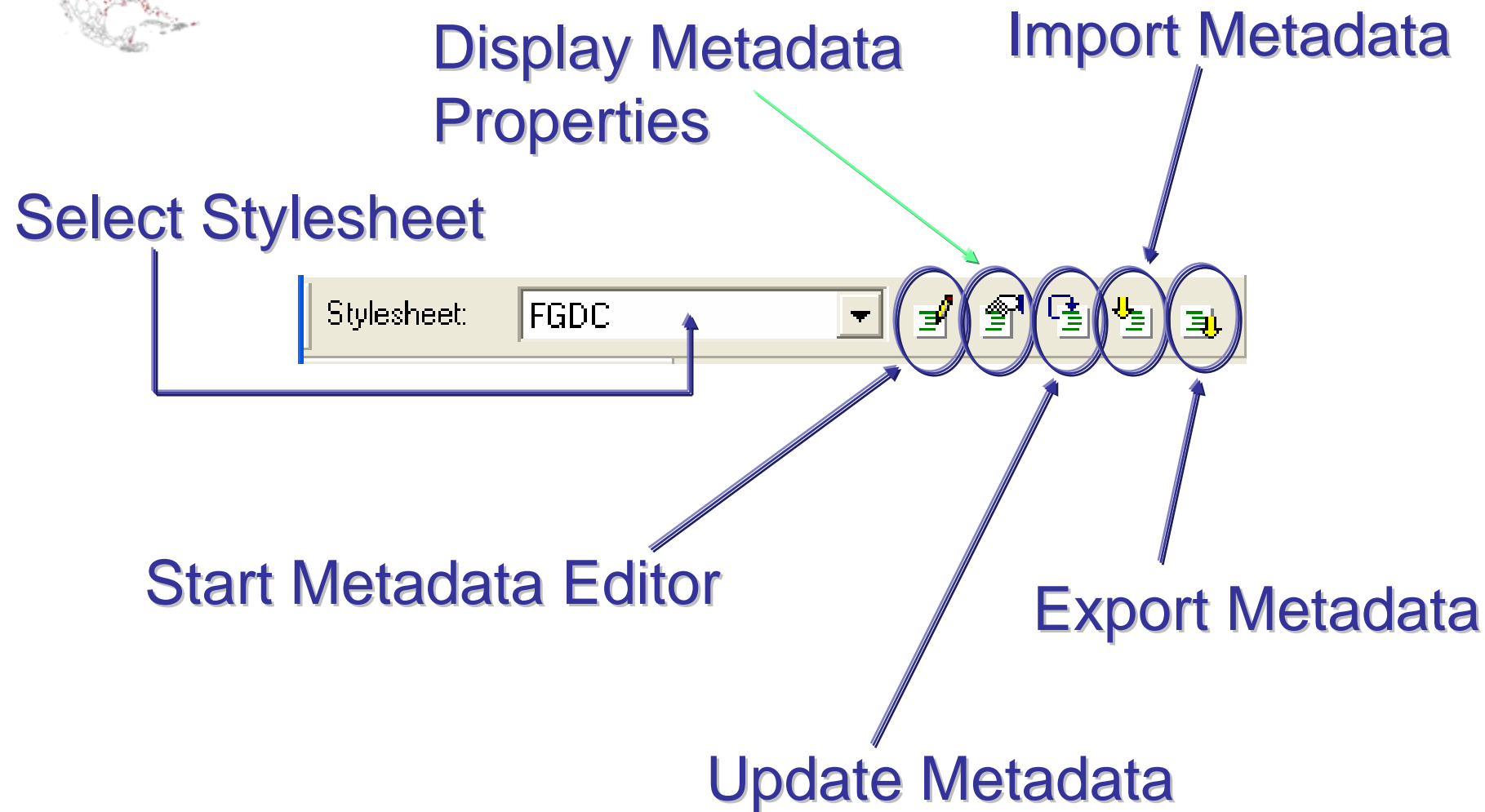
- [Identification Information](#)
- [Data Quality Information](#)
- [Spatial Data Organization Information](#)
- [Spatial Reference Information](#)
- [Entity and Attribute Information](#)
- [Distribution Information](#)
- [Metadata Reference Information](#)

Metadata elements shown with blue text are defined in the Federal Geographic Data Committee's (FGDC) [Content Standard for](#)





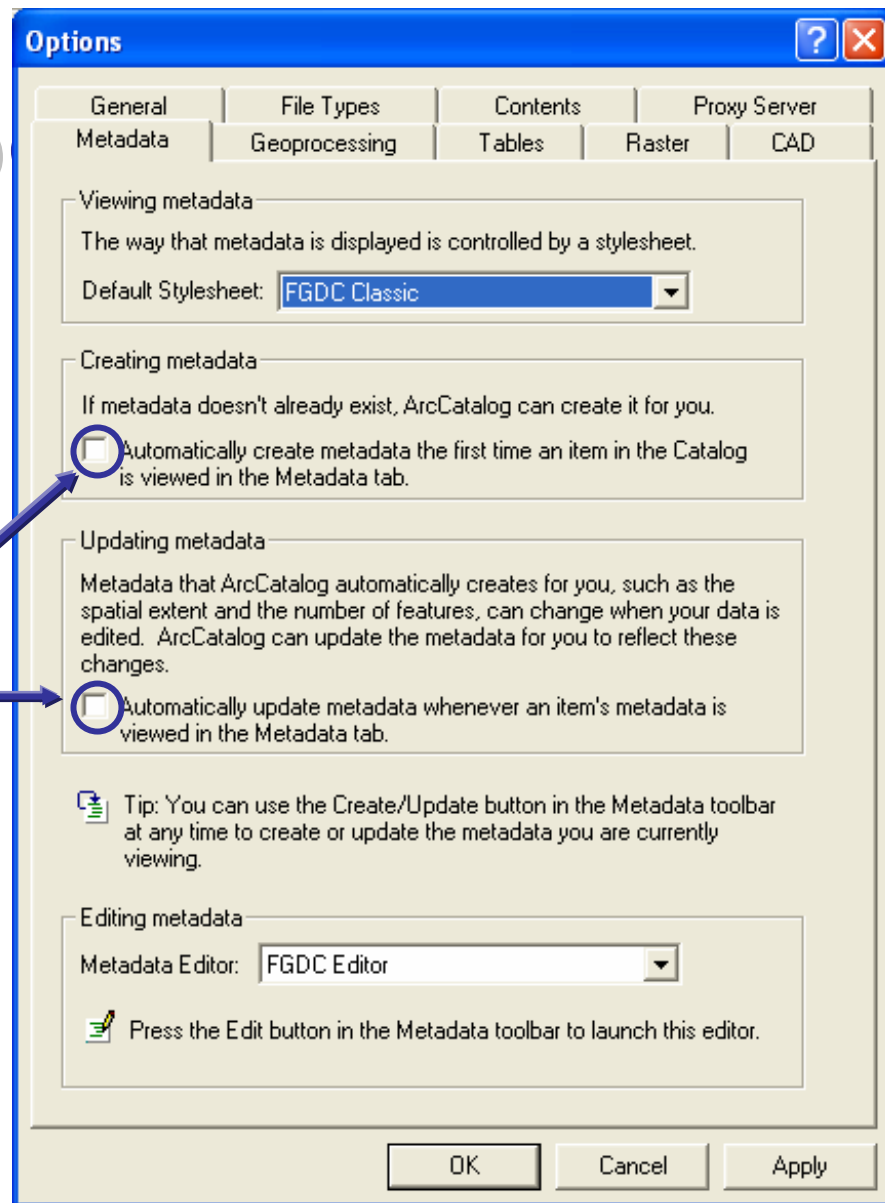
ArcCatalog Metadata Tools





Before you b

- In ArcCatalog
Tools>Options>Metadata
- Choose FGDC Classic as
the Default Stylesheet
- Uncheck boxes to
automatically create
and update metadata
- Select FGDC Editor as
the Metadata Editor





Even more FGDC (usgs) Tools



[Tools for creation of formal metadata](#)

cns



[Tools for creation of formal metadata](#)

mp



[Tools for creation of formal metadata](#)

Xtme





What is CNS?

- Pre-processes metadata records
- Formats record based on element names
- Profile requires configuration file
- CNS output can be read by the metadata parser or MP tool

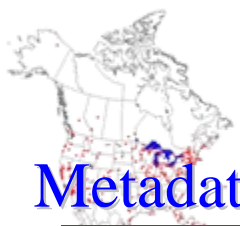




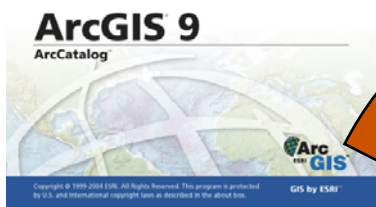
What is MP?

- Compiler to parse formal metadata
- Checks syntax against the FGDC metadata standard
- Configuration file for Profile
- Checks structure and values of elements, not content
- MP reads ASCII text or FGDC xml files!!
 - Uses output from CNS program





Metadata Editor



Export

Clearinghouse

Output Files

Parser



Revise Errors

Pre-Parsing

CNS

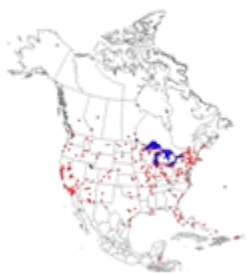




MP output

- layername.txt ASCII text file
- layername.html HTML format for display
- layername.sgml SGML format for search
- layername.xml XML format for metadata crosswalk to other applications
- Mperr.txt Metadata Parser error file





Working with MP Errors

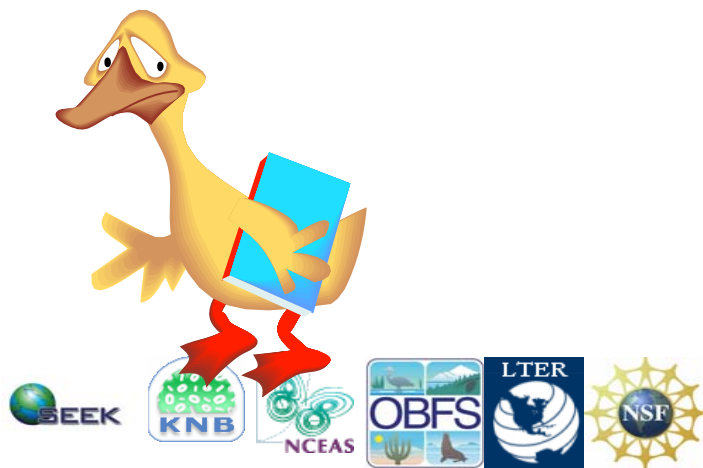
1. Examine error file
2. Note “acceptable errors” – they won’t disappear
3. Return to editor to correct errors
4. Run MP again
5. Once your record is error-free, it is complete



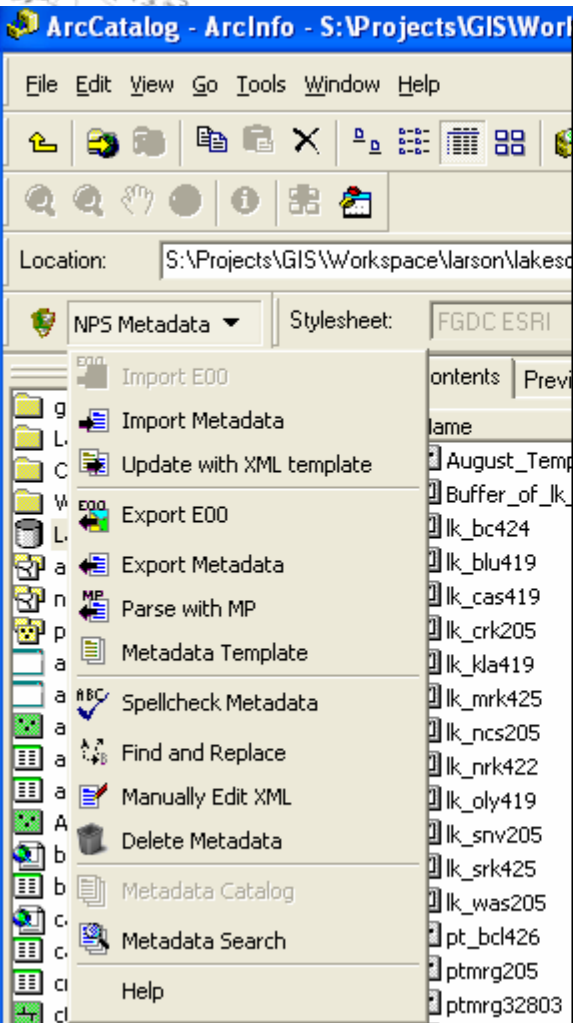


Notes about MP

- Use MP as a guide, not a rigid arbiter
 - Note error severity, set priorities for fixes
 - Be aware of “acceptable errors”
 - MP does not check content
- Only human beings can author quality metadata



NPS Metadata Tools for ArcCatalog



The National Park Service Midwest Region GIS

Metadata Tools - Extension for ArcGIS

Overview

Managing metadata can be a daunting task. With the proliferation of GIS data, however, the need to effectively document and track this data is more important than ever. While this task is considerably easier in ArcGIS's ArcCatalog, there are still a limited number of tools for managing large sets of metadata. The NPS Metadata Tools Extension expands on these tools:

Tools

Parsing

Use MP (USGS's metadata parser) within ArcCatalog to examine metadata errors and use extensions to the FGDC standard. Create an HTML version of the syntax error file to ease correcting errors.

Import metadata

Batch import from SGML or text metadata files or from XML templates.

Export metadata

Export metadata to SGML, text, HTML, or XML using ArcCatalog's or MP's exporter. **NOTE:** Use the **Parse with MP** tool for creating HTML pages to post to the NPS GIS Clearinghouse.

Fix ArcGIS 8.x errors

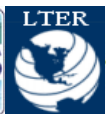
Fixes several ArcGIS 8.x metadata errors including table attributes.

Other

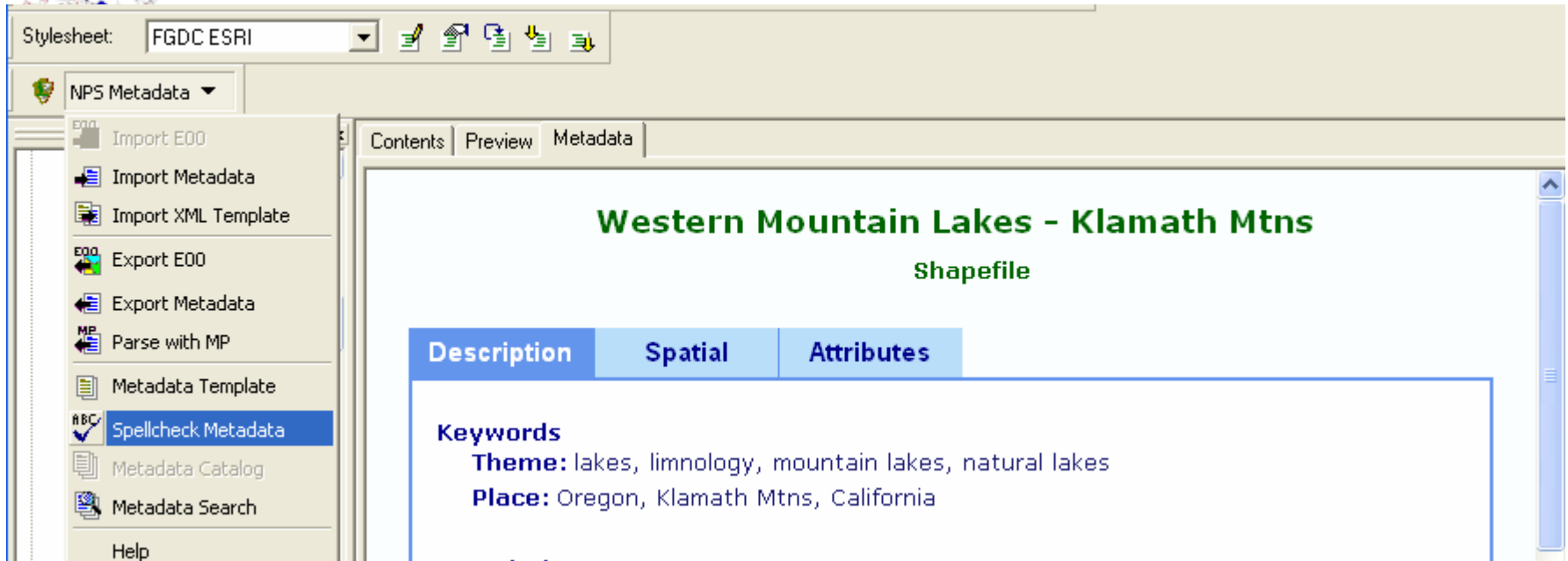
Search metadata fields easily.

Spell check metadata using Word's spellchecker.

http://www.nature.nps.gov/im/units/mwr/gis/metadata/metadata_tools.htm



Spellchecking



Select your data set in ArcCatalog and run Spellcheck Metadata from the NPS Metadata menu



The Mountain Lakes of Western North America database is being developed to facilitate **comparative** analyses of study results gathered by a wide variety of investigations throughout the region.

idinfo/timeperd/current
publication date

idinfo/status/progress
In work

idinfo/status/update
As needed

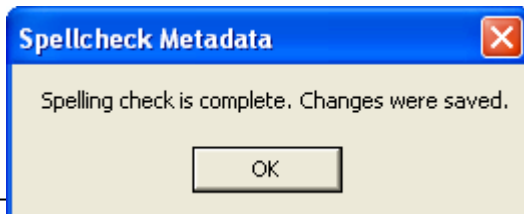
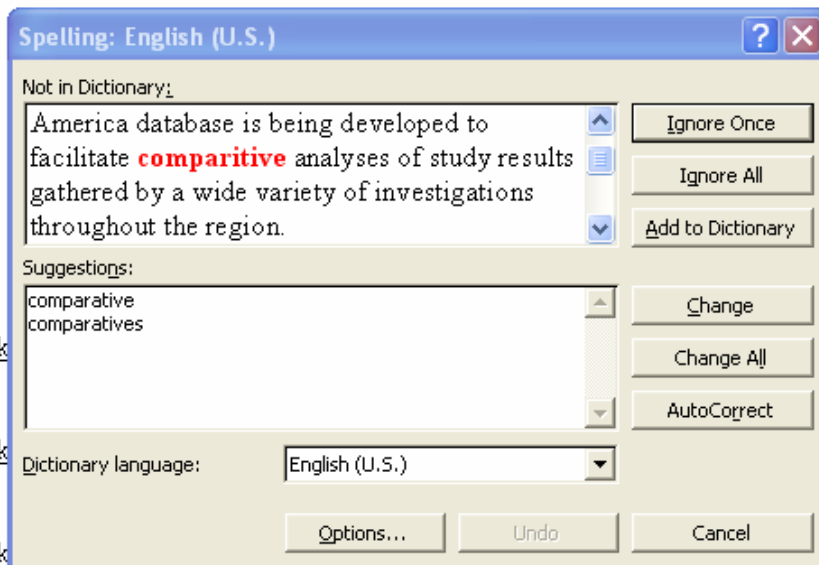
idinfo/keywords/theme/themek
None

idinfo/keywords/theme/themek
lakes

idinfo/keywords/theme/themek
limnology

idinfo/keywords/theme/themekey[2]
mountain lakes

idinfo/keywords/theme/themekey[3]
natural lakes

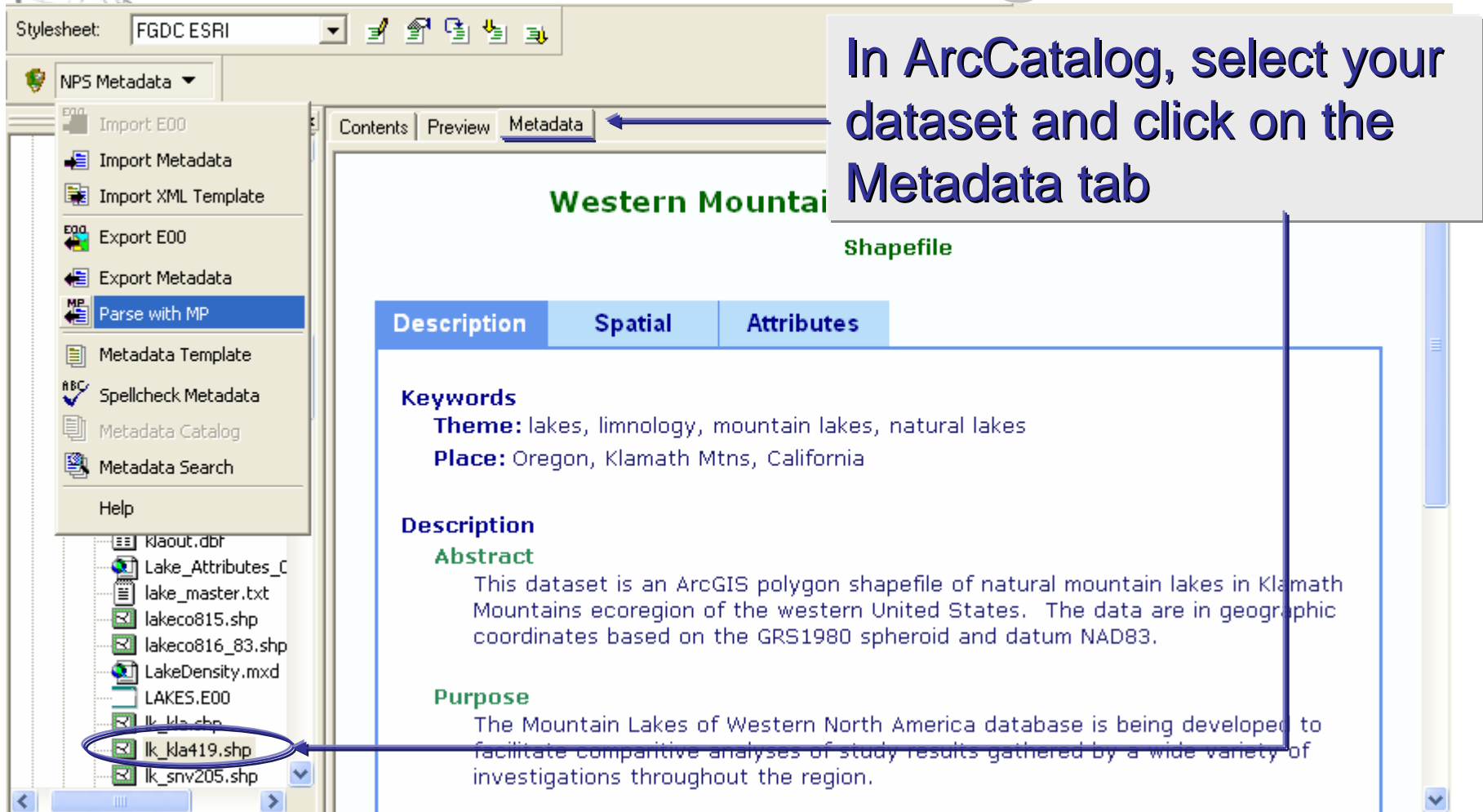


ArcCatalog
examines the
entire
metadata
record using
the Microsoft
Word spell
checking tool

Spelling corrections are written directly to
the <layer>.xml file



MP in ArcCatalog



The screenshot shows the ArcCatalog application window. On the left, the 'NPS Metadata' menu is open, with 'Parse with MP' highlighted. Below the menu, a list of files is shown, with 'lk_kla419.shp' circled. On the right, the 'Metadata' tab is selected for the 'Western Mountain Shapefile'. The metadata content includes 'Keywords' (Theme: lakes, limnology, mountain lakes, natural lakes; Place: Oregon, Klamath Mtns, California), 'Description' (Abstract: This dataset is an ArcGIS polygon shapefile of natural mountain lakes in Klamath Mountains ecoregion...), and 'Purpose' (The Mountain Lakes of Western North America database is being developed to facilitate comparative analyses...).

Stylesheet: FGDC ESRI

NPS Metadata

- Import E00
- Import Metadata
- Import XML Template
- Export E00
- Export Metadata
- Parse with MP**
- Metadata Template
- Spellcheck Metadata
- Metadata Catalog
- Metadata Search
- Help

Contents Preview **Metadata**

Western Mountain Shapefile

Description	Spatial	Attributes
Keywords Theme: lakes, limnology, mountain lakes, natural lakes Place: Oregon, Klamath Mtns, California		
Description Abstract This dataset is an ArcGIS polygon shapefile of natural mountain lakes in Klamath Mountains ecoregion of the western United States. The data are in geographic coordinates based on the GRS1980 spheroid and datum NAD83.		
Purpose The Mountain Lakes of Western North America database is being developed to facilitate comparative analyses of study results gathered by a wide variety of investigations throughout the region.		

Run MP from the NPS
Metadata menu



MP Settings

Browse to location of
MP program files

Browse to the location
configuration file for
profiles or extensions

Metadata Parser (MP) Settings

MP executable: C:\USGS\tools\bin\mp.exe Browse...

Configuration file: <use to specify MP extensions> Browse...

Output folder: <by default, same folder as dataset> Browse...

Generate for each XML file

- ☐ HTML file (*.html)
- ☐ HTML FAQ output (*.faq.html)
- ☐ SGML file (*.sgml)
- ☐ Text file (*.txt)
- ☐ XML file (*.xml)
- ☒ Syntax error file (*.err)
- ☒ HTML error file (*.err.html)
(uses 'err2html.exe' program)

☒ Show syntax error report for all datasets

☒ Strip ESRI XML tags before parsing

☐ Reorder major elements before parsing

Help OK Cancel

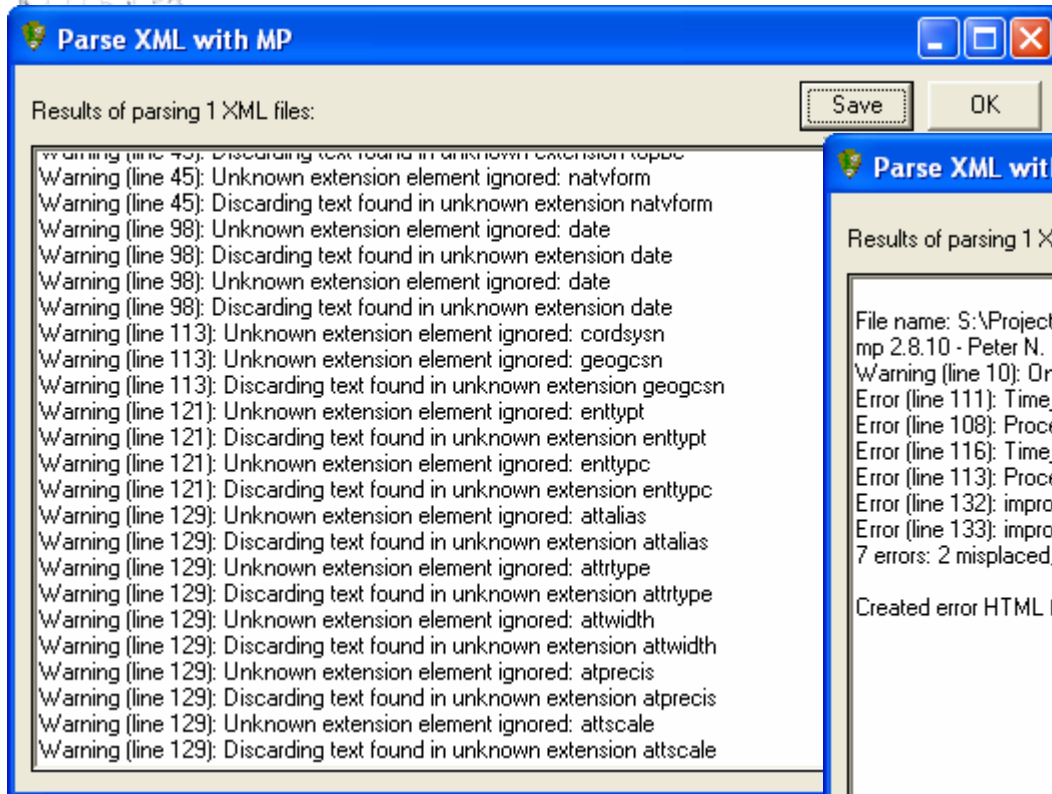
Output
Files

On-the-fly
error report

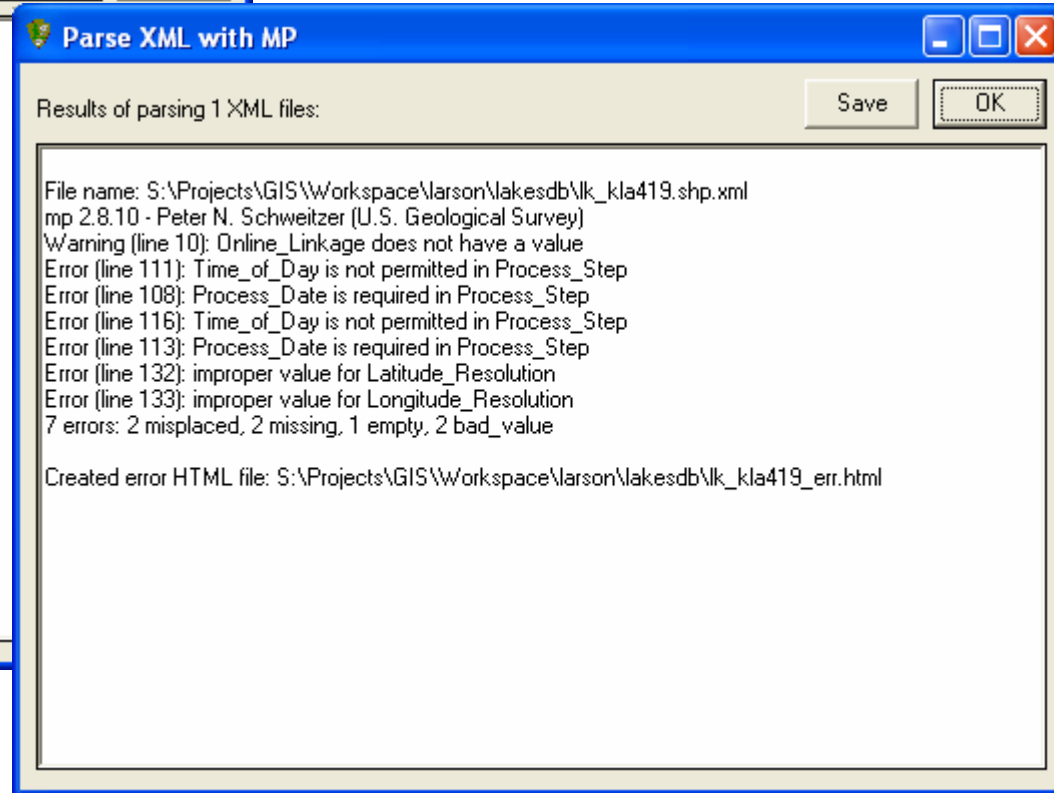
Creates easy-to-read
error file

Automatic clean-up
of non-standard
elements

ESRI XML Tags



Non-standard metadata
elements generate lots
of warnings in MP



Stripping ESRI xml tags
eliminates warnings



MP Error File (html)

mp 2.8.10 - Peter N. Schweitzer (U.S. Geological Survey)

7 errors: 2 misplaced, 2 missing, 1 empty, 2 bad_value

Type	Description or line numbers	Line(s) (or count)
Severity 5: Misplaced elements		
Error	Time of Day (9.1.2) is not permitted in Process Step (2.5.2)	111 116
Severity 3: Missing elements		
Error	Process Date (2.5.2.3) is required in Process Step (2.5.2)	108 113
Severity 2: Empty elements		
Warning	Online Linkage (8.10) does not have a value	10
Severity 1: Elements with improper values		
Error	improper value for Latitude Resolution (4.1.1.1)	132
Error	improper value for Longitude Resolution (4.1.1.2)	133

- Errors listed by severity

- Use line numbers to find elements in <layer>.xml file

- Use metadata editor to correct errors

- Run MP again

- When metadata record is “error-free”, output in FGDC txt, sgml, xml, and html formats





Links

- Ecoinformatics: <http://www.ecoinformatics.org>
- LTER <http://www.lternet.edu>
- Morpho <http://knb.ecoinformatics.org/software/morpho/>
- XMLSpy <http://www.altova.com>
- oXygen <http://www.oxygenxml.com>
- Stylus Studio <http://www.stylusstudio.com>
- CNS <http://geology.usgs.gov/tools/metadata/>
- MP <http://geology.usgs.gov/tools/metadata/>
- Err2html <http://geology.usgs.gov/tools/metadata/>
- TKME <http://geology.usgs.gov/tools/metadata/>
- MetaScribe <http://www.csc.noaa.gov/metadata/metascibe/>
- Metavist <http://ncrs.fs.fed.us/pubs/viewpub.asp?key=2737>
- ArcCatalog <http://www.esri.com/software/arcgis/about/desktop.html>
- Natl Park Service Tools for ArcCatalog
http://www.nature.nps.gov/im/units/mwr/gis/metadata/metadata_tools.htm
- MySQL <http://www.mysql.com>
- Open Source stuff <http://www.freshmeat.net>
- FGDC metadata online validator: <http://geo-nsdi.er.usgs.gov/validate.php>
- EML metadata online validator: <http://knb.ecoinformatics.org/emlparser/index.html>
- Mother of all search engines: <http://www.google.com>





Key people

- Morpho: Sid Garg sgarg@nceas.ucsb.edu, Matt Jones: jones@nceas.ucsb.edu
- CNS, MP, TKME george_lienkamper@usgs.gov
- XMLSpy Basic questions, Mark Servilla, Inigo San Gil, Deep troubles support@altova.com
- oXygen support@[oxygenxml.com](mailto:support@oxygenxml.com)
- Metavist <http://ncrs.fs.fed.us/pubs/viewpub.asp?key=2737>
- ArcCatalog





Thanks!

Mark Servilla

George Lienkamper

Sam Romanello





How to use the crosswalk

- Within XMLspy/oXygen – use XSL transformation option – follow instructions
- Using standalone “xalan” XSL interpreter





Using TKME for the Biological Data Profile

- Editing BDP not supported in ArcCatalog
- TKME is a public domain tool
 - Fully compliant with CSDGM and BDP
 - Simple interface
 - Supports Templates through “Snippets”
 - Developed and supported by Peter Schweitzer of USGS
- Contained on your workshop CD
- Available for download:
<http://geology.usgs.gov/tools/metadata/>



Entering metadata elements into TKME

Tkme: V:/Metadata/workshops/Finley/BDPMetadataTemplate.met

File Edit View Add Snippets Help

Metadata

Identification_Information

Citation

Citation_Information

Originator

Publication_Date

Title

Geospatial_Data_Presentation_Fo

Description

Abstract

Purpose

Time_Period_of_Content

Currentness_Reference

Status

Progress

Maintenance_and_Update_Frequency

Spatial_Domain

Description_of_Geographic_Extent

Keywords

Theme

Theme_Keyword_Thesaurus

Theme_Keyword

Access_Constraints

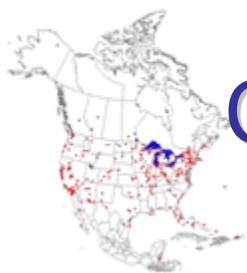
Use_Constraints

BDP Metadata Template

Metadata element names are listed on the left side of screen.

Element value is entered on the right side of screen.





Compound Elements

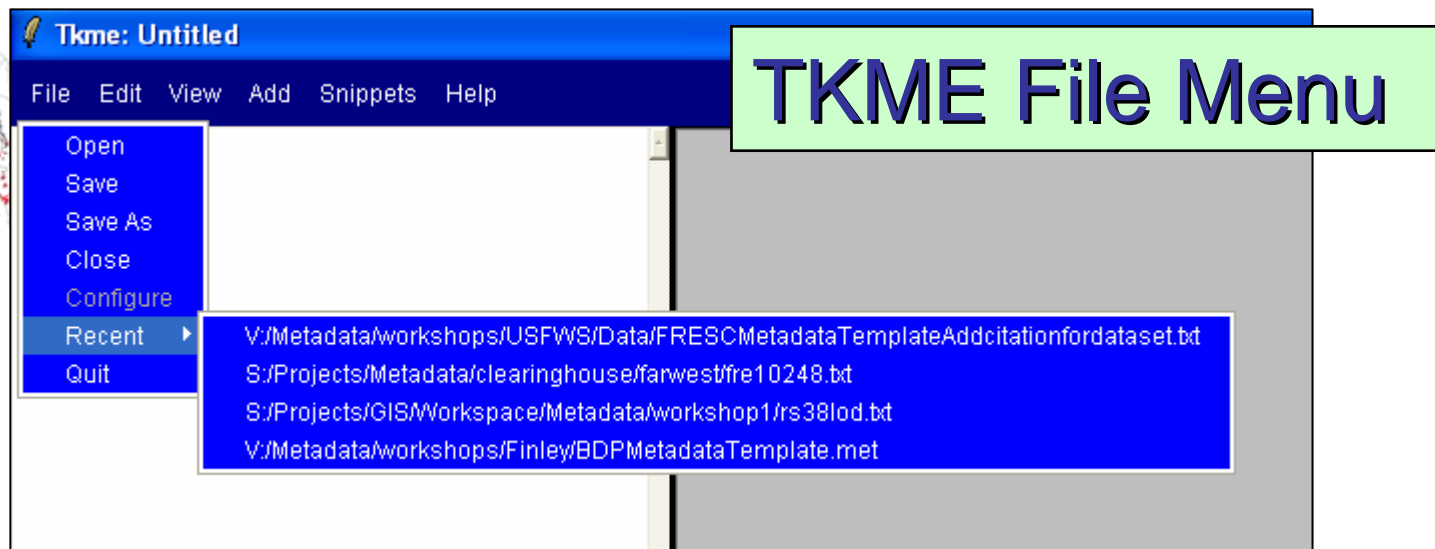
Tkme: V:/Metadata/workshops/Finley/BDPMetadataTemplate.met

File Edit View Add Snippets Help

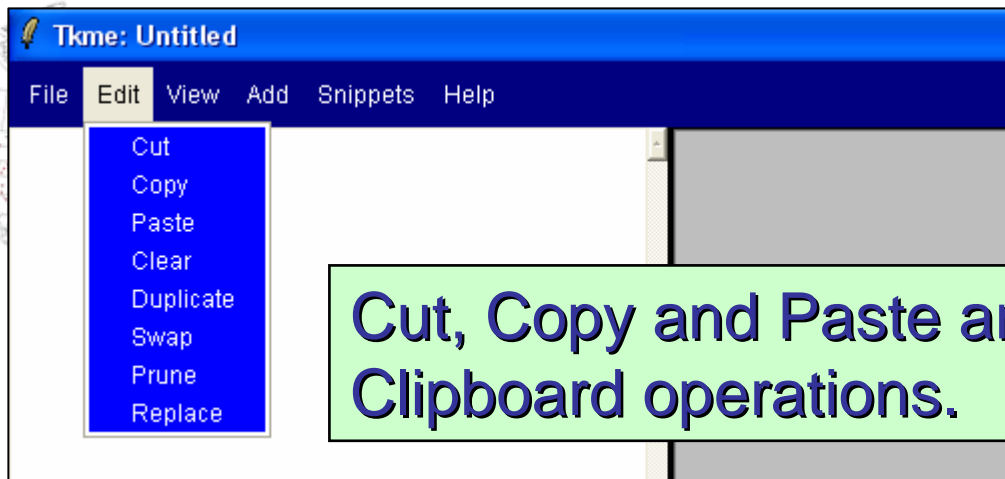
Metadata	Identification_Information
Identification_Information	
Citation	
Citation_Information	
Originator	
Publication_Date	
Title	
Geospatial_Data_Presentation_Fo:	
Description	
Abstract	
Purpose	
Time_Period_of_Content	
Currentness_Reference	
Status	
Progress	
Maintenance_and_Update_Frequency	
Spatial_Domain	
Description_of_Geographic_Extent	
Keywords	
Theme	
Theme_Keyword_Thesaurus	
Theme_Keyword	
Access_Constraints	
Use_Constraints	

Compound elements used to structure hierarchy are indicated by a grey screen with no data entry.





- Open, Save, Save As, Close, and Quit are typical Windows operations.
- Configure – specify a configuration file that contains information about profiles, extensions, and output format
- Recent – displays the most recently edited files

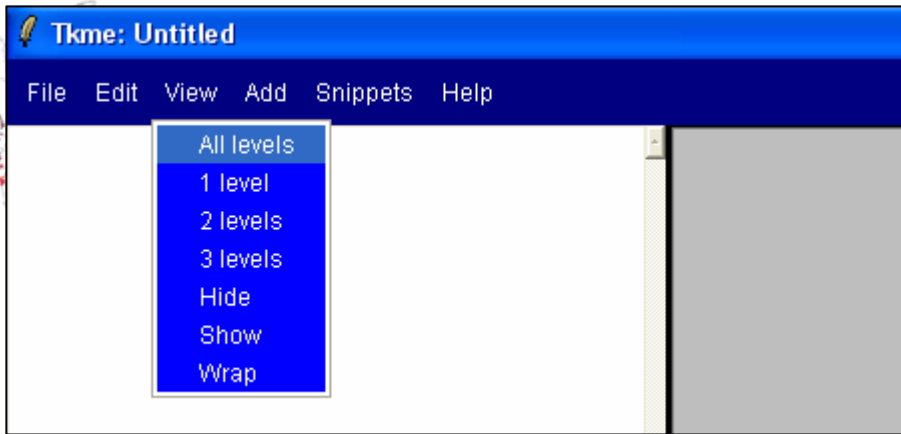


TKME Edit Menu

Cut, Copy and Paste are typical Windows Clipboard operations.

- Clear – removes text belonging to a subtree
- Duplicate – makes and places a copy of the current element
- Swap – switches the order of the current subtree with the preceding sibling
- Prune – deletes all subtrees that do not contain text
- Replace – performs search and replace function





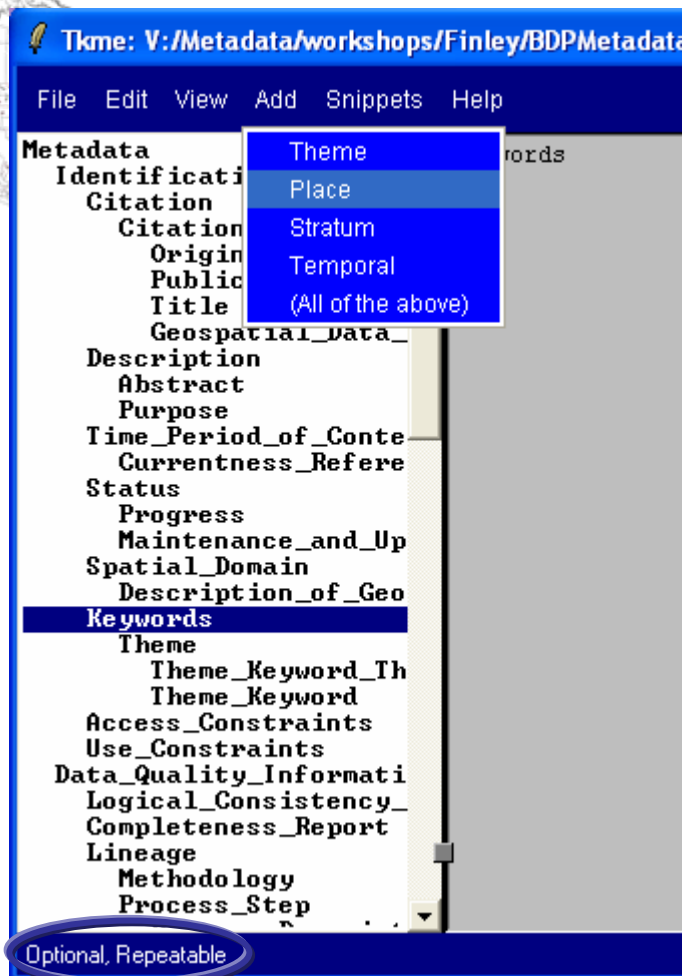
TKME View Menu

- All levels – shows the entire hierarchy
- 1 level – shows only the sections of the metadata standard
- (2,3) levels – shows two or three level below sections
- Hide/Show – hides or shows the children of the current element
- Wrap – turns on word-wrap in the value window



TKME Add Menu

Wherever the cursor is placed in the hierarchy, the Add menu displays elements that may be added at that level. Clicking will add the element. Use Add again to insert children of the new element.



Note: Reporting requirement



About Snippets

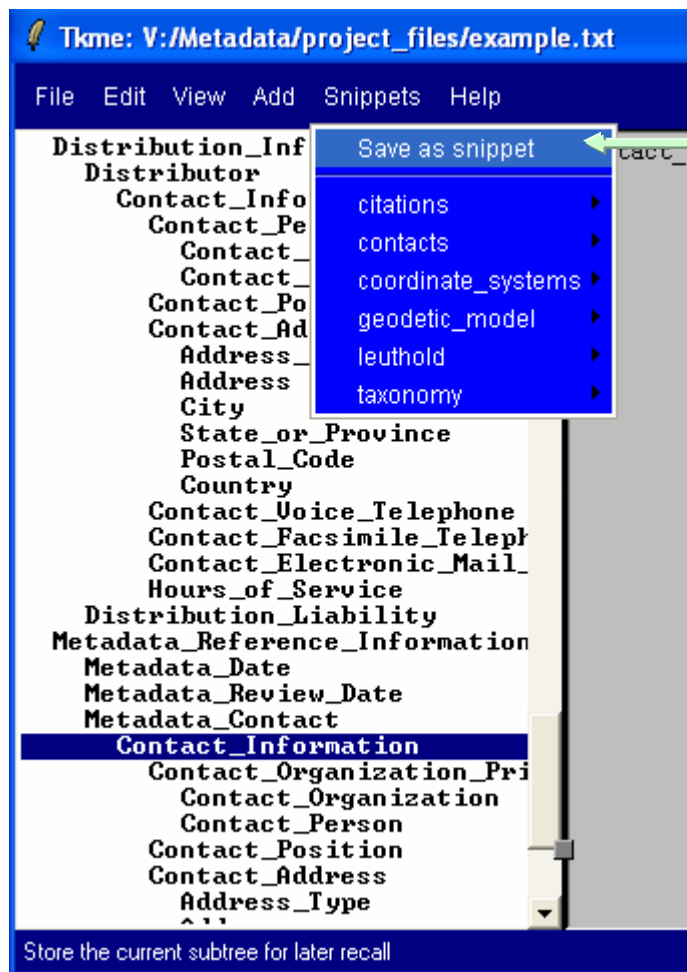
- Snippets are portions of a metadata record that may be copied and reused within a record or in other metadata records: *Think Templates*
- For BDP records, snippets are especially helpful for managing *Citations*
- See
<Programs\USGS\tools\doc\opinion\announce\snippets.html>
on CD for more details about snippets





Saving Snippets

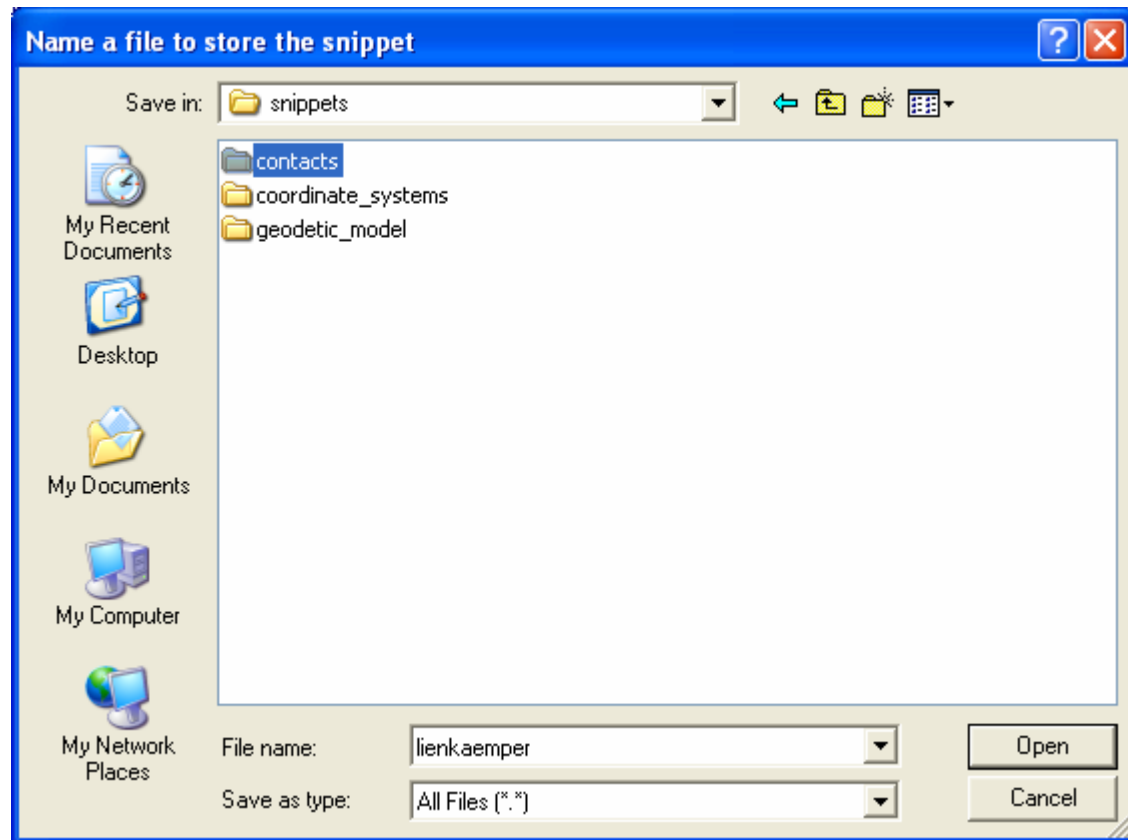
Highlight an element that heads a section you want to save as a snippet



Select Save as snippet from the Snippets menu



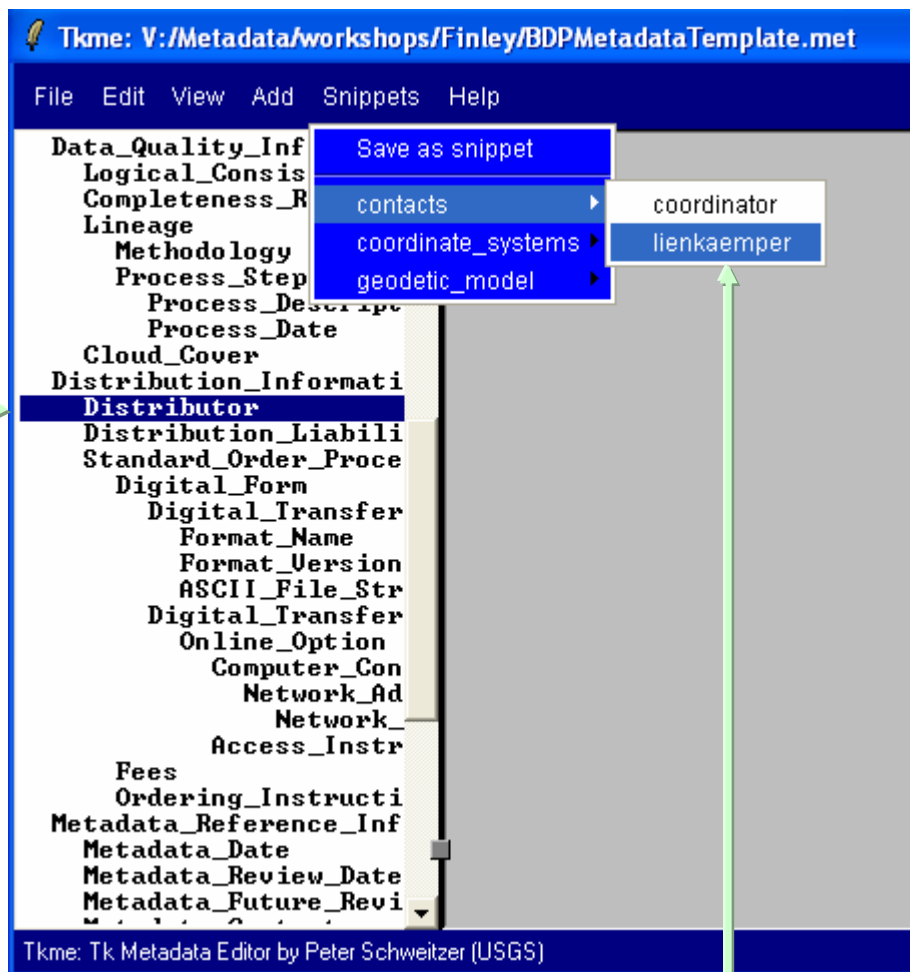
Storing Snippets



Snippets are stored as text files in a \snippets directory in the USGS folder that contains the TKME program. Folders may be created in \snippets to help organize snippet files. Name snippets without extensions.

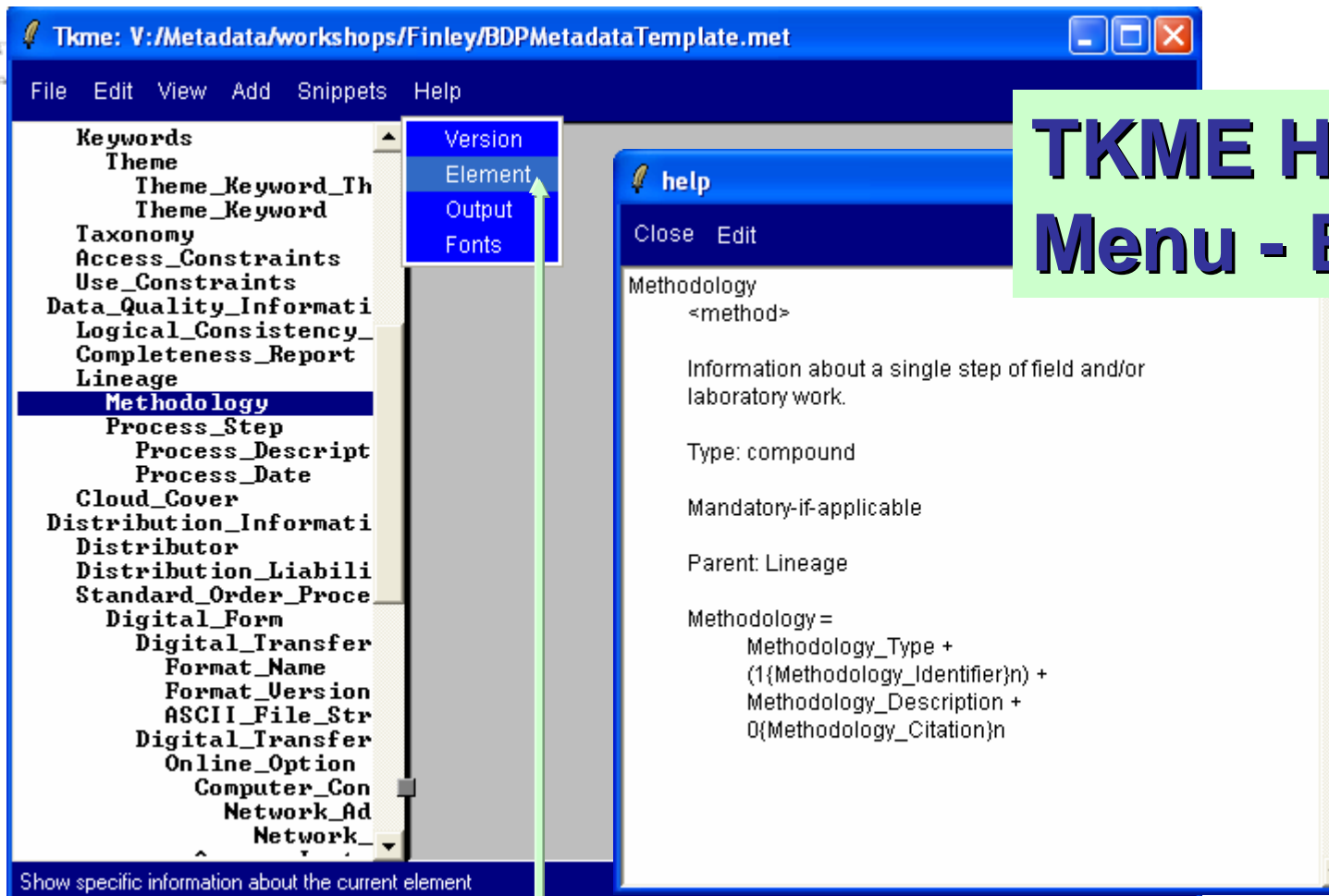
Placing Snippets

Select a parent element where the snippet will be inserted



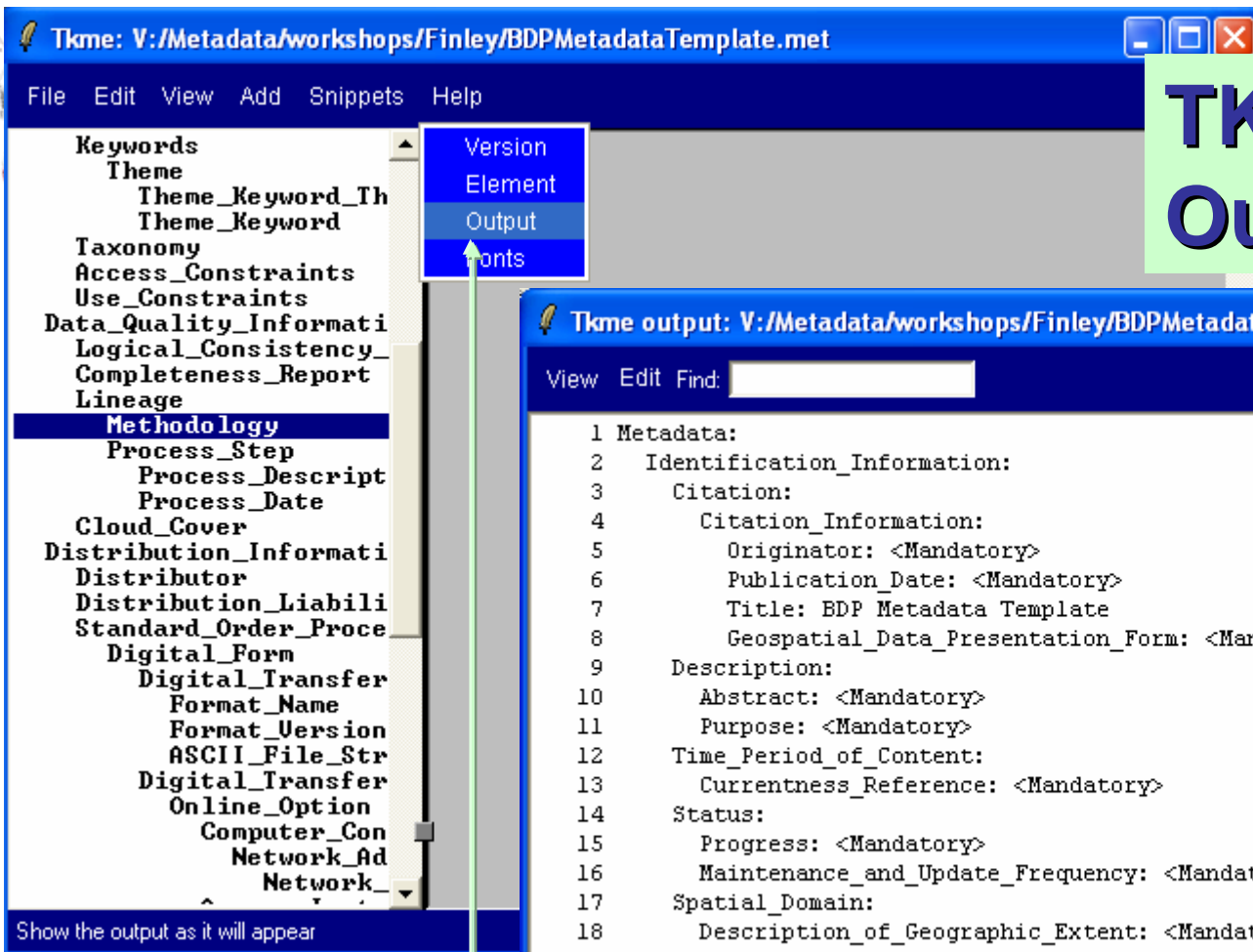
Choose a snippet from the Snippets menu and the elements will be placed in the metadata record



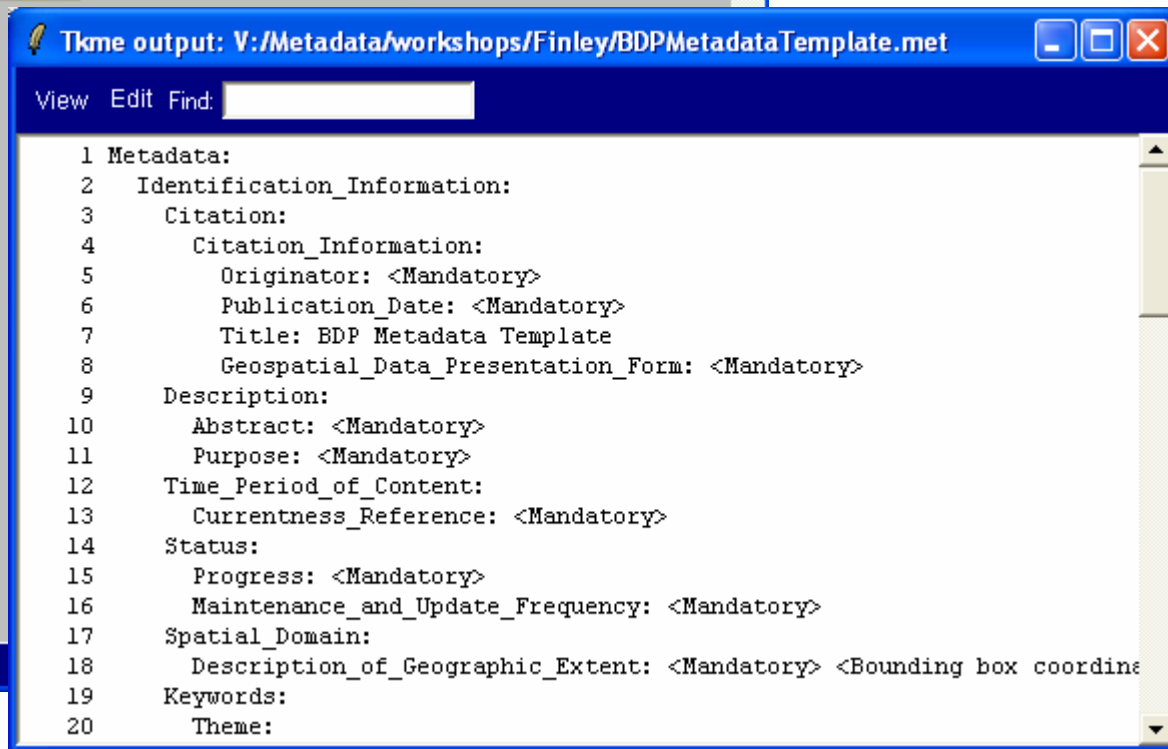


TKME Help Menu - Element

Highlight a metadata element and choose Element from the Help menu. An entry from the FGDC workbook will be displayed

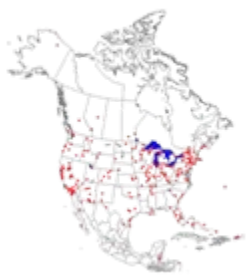


TKME Menu-Output



View the current status of your metadata record by selecting Output from the Help menu





Exercise - Metadata Verification

- Select your data set in ArcCatalog
 - Select the metadata tab
- Run Spellcheck from the NPS Metadata menu
- Run MP from the NPS Metadata menu
 - Choose MP options
- Examine the Error File
- Revise and run MP again
- Save your metadata output files to C:\temp or to a floppy disk

