

Environmental Data Management Workshop

**Presented to the
San Francisco Bay Area Conservation Community**

**Aug. 20, 2008
9:00 am – 5:00 pm**

GIS Teaching Lab, 3059 Stevenson Hall, SSU

Environmental Data Management Workshop

Presenters:

**Deanne DiPietro, Director, Research &
Information Management Program
and
Pat Stiefer, Digital Librarian**

Sonoma Ecology Center

**Sponsored by the San Francisco Bay Area Conservation Commons,
with funding from the Federal Geospatial Data Committee**

Part 1 – Metadata Basics

9:00 – 12:00

Overview

- Logistics
- Overall objectives and preview of the day
- Introductions

Logistics

- Restrooms & drinking fountains
- Parking Okay?
- Cell phones off or silent ring
- Monitors off for speakers

Workshop Objectives

- Define formal metadata
- Describe metadata's value to the organization
- Discuss why we use metadata standards
- Use FGDC standard to successfully create metadata using ArcCatalog and CEIC editors
- Create online metadata catalog at CEIC and publish metadata for harvest

Introduce Ourselves

Notes on Data Management

- Metadata is a huge topic
- It takes time to make quality spatial metadata
- It goes faster the more experience you have
- You learn a lot “on the job”, actually making the metadata happen

Metadata: What and Why

- Metadata is a description of your data
- A step in the data management process
- Immediate and long-term benefits

What is Metadata?

- A component of data management
 - Organizing and documenting information resources for retrieval, use, and sharing
- Part of project management
 - Documents development of data through a project life cycle
 - Preserves organization's investment in data for future use

What is Metadata?

- Description of an entity's important characteristics (what, who, where, when, etc.)
- Surrogate for the data
- You see metadata every day – you are making it

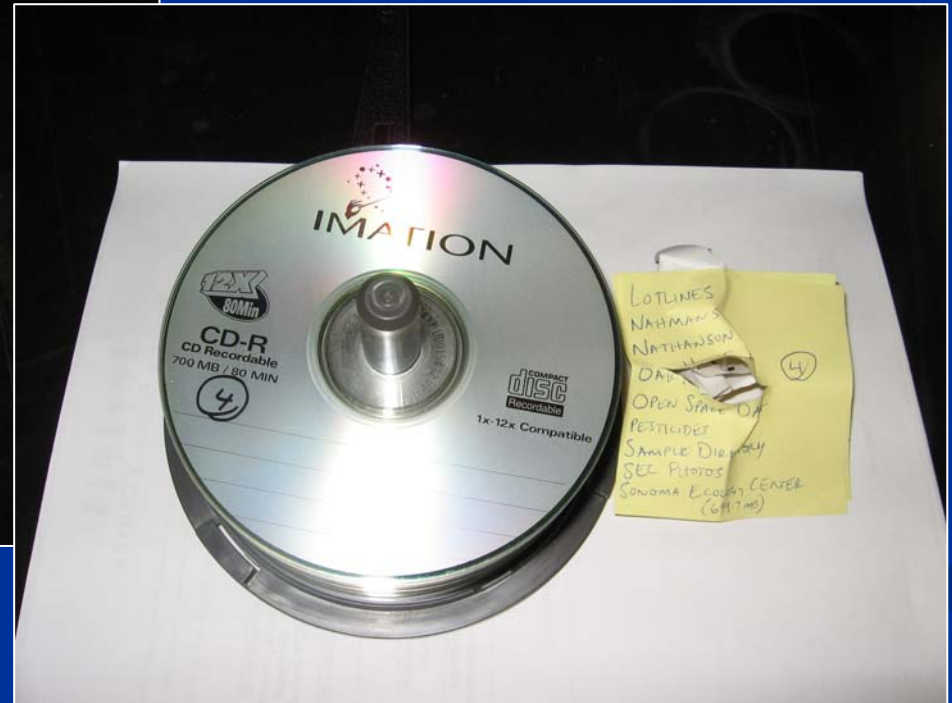
Electronic note files, emails

Formal data dictionaries

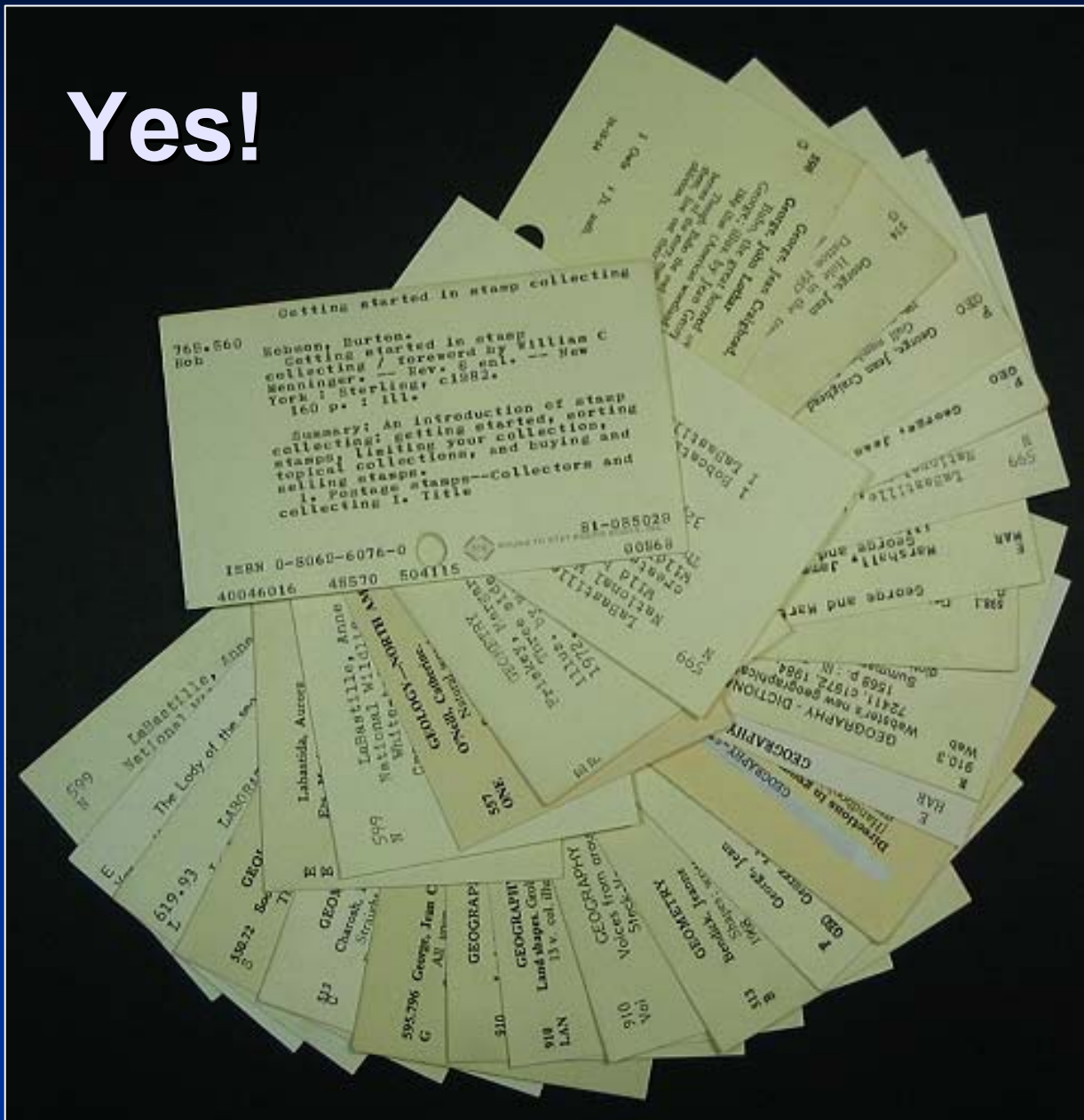
Hand-written notations

Is it Metadata?

Yes!



Yes!

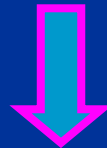


What's the Difference?

- Cards more consistency in content and presentation.
- Cards conform to a set of rules for content

What is Metadata?

- Formal metadata is specific relevant information about data, recorded in a structured way.
Metadata is considered part of the data.
 - Which aspects are relevant? (content)
 - What kind of format should we use? (structure)



Metadata Standards

Metadata Standards

- A metadata standard is a set of rules or specifications for the content and sometimes format of metadata.
- Standards enable us to consistently describe a data resource, facilitating creation and sharing of metadata
 - Metadata editing software
 - Metadata catalogs

Metadata Standards

- Some formal metadata standards in use today: different standards for different data types
 - FGDC CSDGM and ISO 19115 – most widely used US and International standards for spatial metadata, but can easily be used for other data resources
 - CEIC – focuses on subset of FGDC sections that are most useful for general purpose resource discovery tool for both spatial and nonspatial data
- Dublin Core Metadata Element Set – vocabulary of 15 properties for resource description, most useful for nonspatial resources
- Ecological Metadata Language – a metadata standard for the ecology discipline, for spatial and non-spatial information resources, based in part on Dublin Core

FGDC Standard

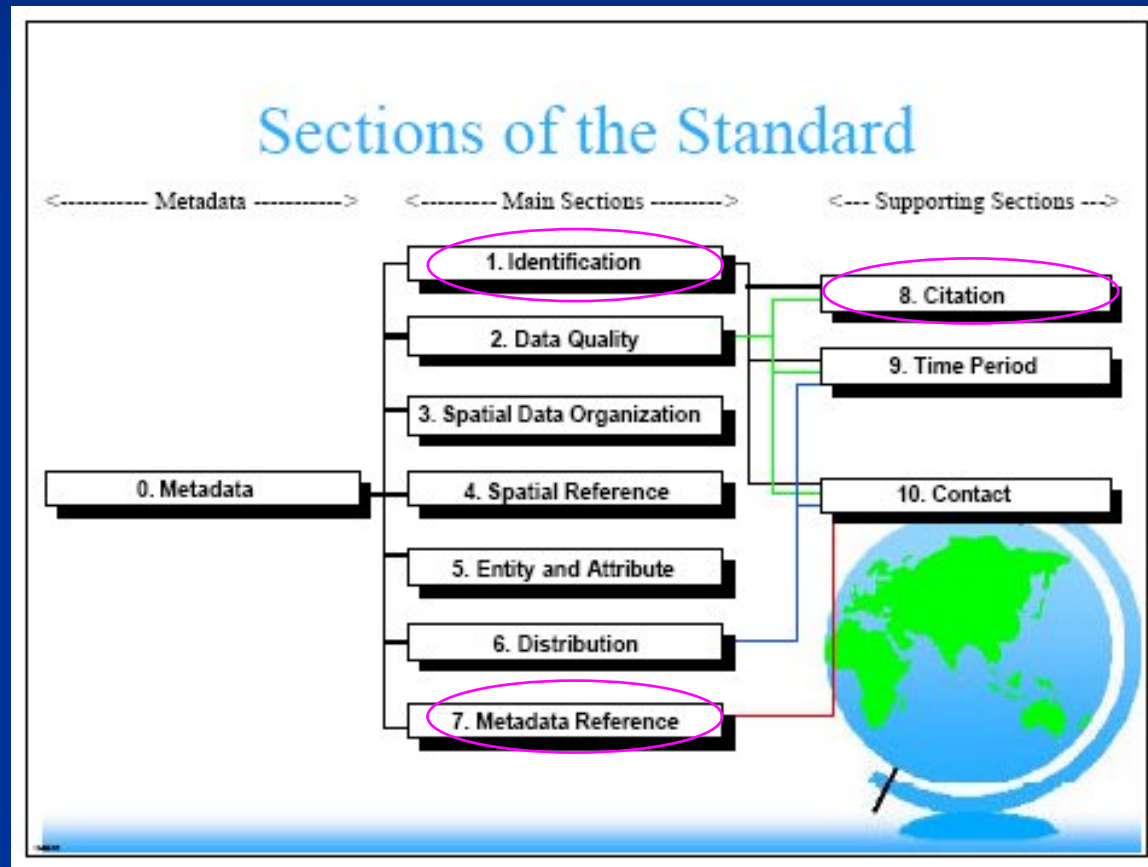
- FGDC Content Standard for Digital Geospatial Metadata (aka “FGDC Standard”)
 - currently the standard required by the US government in metadata created for its own geospatial data and for data created by its contractors
 - non-Federal agencies are encouraged to use FGDC standard to promote data sharing, support interoperability
 - specifies information content; text and xml are exchange formats

FGDC Standard

- FGDC is extensive -- it must support a wide range of geospatial data types
- Ten sections of required and optional simple and compound elements support both metadata discovery and use.
- Fear of Metadata can result!

FGDC Standard Sections

- Identification (1)
- Metadata Ref. (7)
- Citation (8)
- Data Quality (2)
- Spatial Data Org. (3)
- Spatial Reference (4)
- Entity & Attribute (5)
- Distribution (6)
- Time Period (9)
- Contact (10)

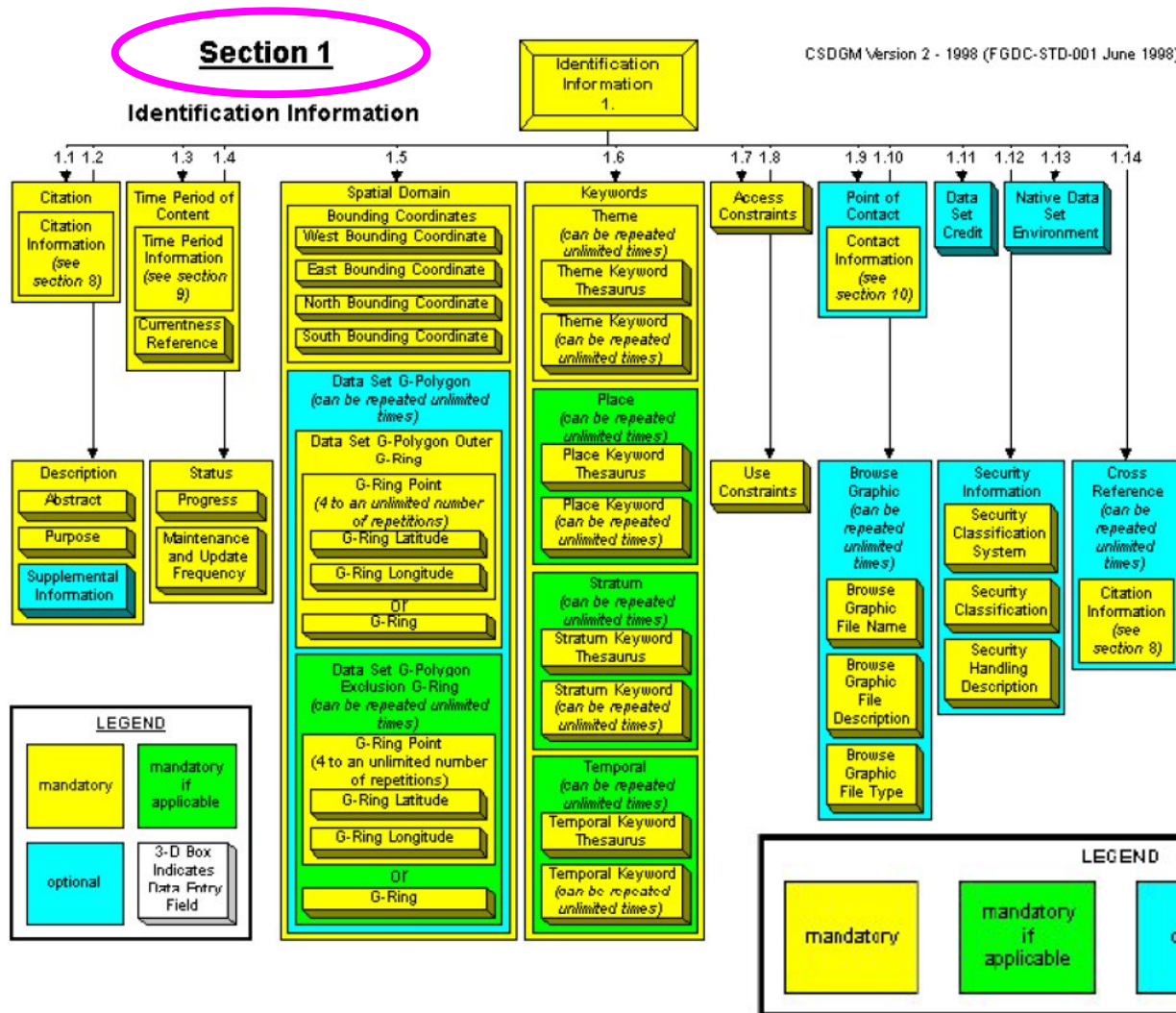


Implementation Assistance

- FGDC supports metadata/standards implementation with training sessions, educational resource materials, grants, etc.
 - Training materials:
<http://www.fgdc.gov/metadata/online-metadata-resources>
 - Cooperative Agreements Program Grants:
<http://www.fgdc.gov/grants>
 - Helpful publications, incl. “the workbook”:
<http://www.fgdc.gov/metadata/metadata-publications-list>

Implementation Support

■ CSDGM Workbook Diagram



Implementation Assistance

Additional short summary documents from FGDC:

- Geospatial Metadata Quick Guide, 2005, at <http://www.fgdc.gov/metadata/documents/MetadataQuickGuide.pdf>
- Top Ten Metadata Errors, May 2006, at <http://www.fgdc.gov/metadata/documents/top10metadataerrors.pdf>

More Implementation Assistance

■ Preview of in-context help in ArcCatalog

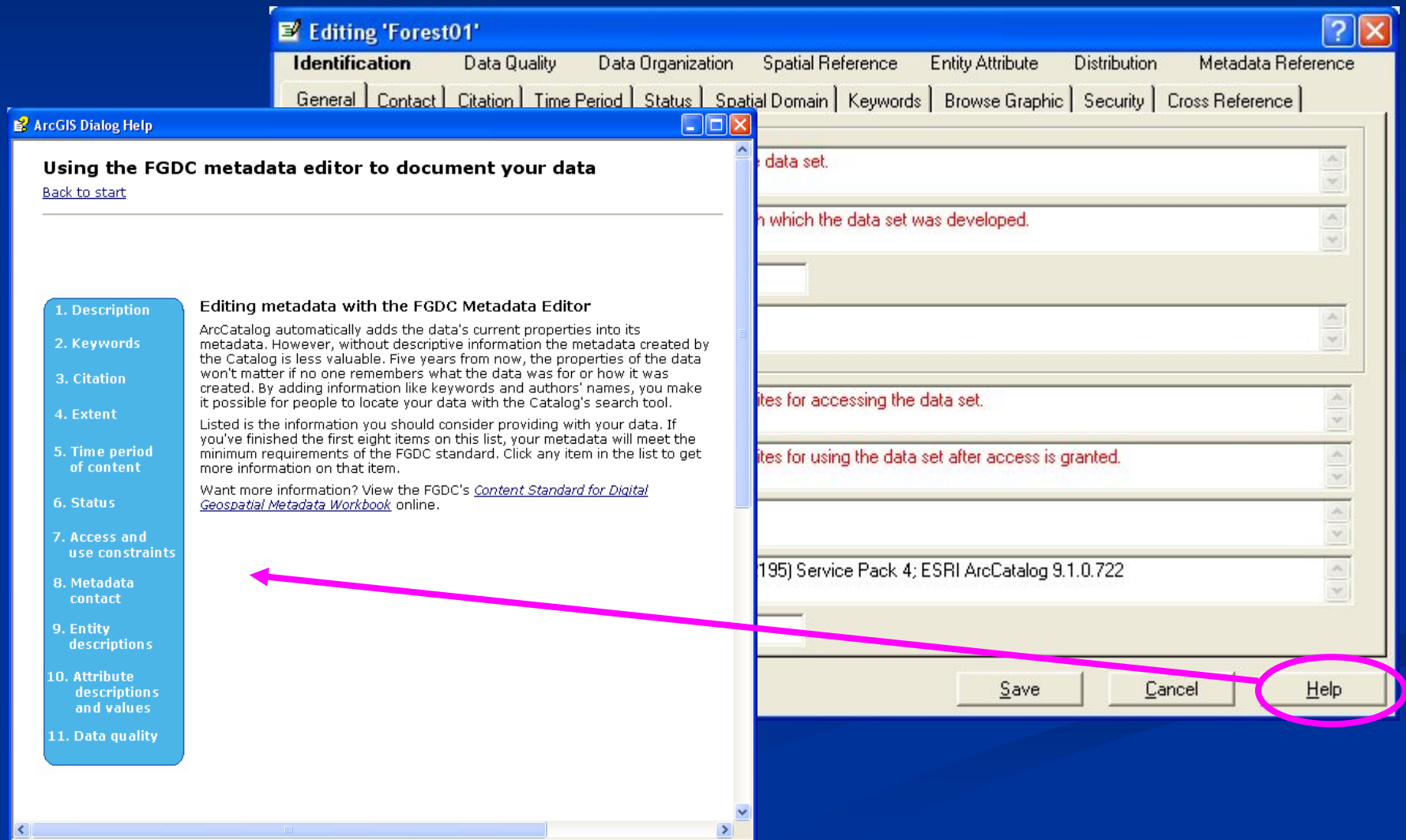
The screenshot shows the 'Editing Forest01' dialog box in ArcCatalog. The 'Identification' tab is active, showing fields for Abstract, Purpose, Language, Supplemental Information, Access Constraints, Use Constraints, Data Set Credit, Native Data Set Environment, and Native Data Set Format. A dashed pink arrow points from a question mark icon in the top right corner of the dialog box to a question mark icon in the 'Data Set Credit' field. Another dashed pink arrow points from the 'Data Set Credit' field to a help popup window.

Data Set Credit
Section: 1.11
Name: Data Set Credit
Definition: recognition of those who contributed to the data set.
Status: optional
Type: text
Domain: free text
Short Name: datacred
FAQ: / Notes:
Q: What is the purpose of the "Data Set Credit" data element?
A: The "Data Set Credit" data element was provided to allow the originator to recognize other person(s), organization(s), or events that made the data set possible. These might include sponsors, sources of funds, reviewers, dedications, etc.

Return to [Identification Information](#) **Or** **Contract or!**

More Implementation Assistance

■ Preview of help files in ArcCatalog



Why make metadata

- Immediate, practical benefits to organization
 - Data longevity - replaces human memory and allows reuse of information/data
 - Efficiency - reduces questions about how to access and use the data
 - Limit data liability - explicitly designating the limits of use of the data
- Intellectual honesty: give credit to sources you used
- Document geo-processing history

Uses for Metadata

- Supports in-house data organization
- Sharing/discovery
- Supporting data use

Break!

10:15-10:30



Creating Metadata

- ArcCatalog Tour and Practice
- CEIC Tour and Practice

ArcCatalog Tour and Hands-On Practice

ArcCatalog

- Software, part of ESRI ArcGIS suite
- Makes metadata that becomes part of the GIS dataset
- Seven forms, corresponding to FGDC standard sections
- Some elements are auto-populated

ArcCatalog

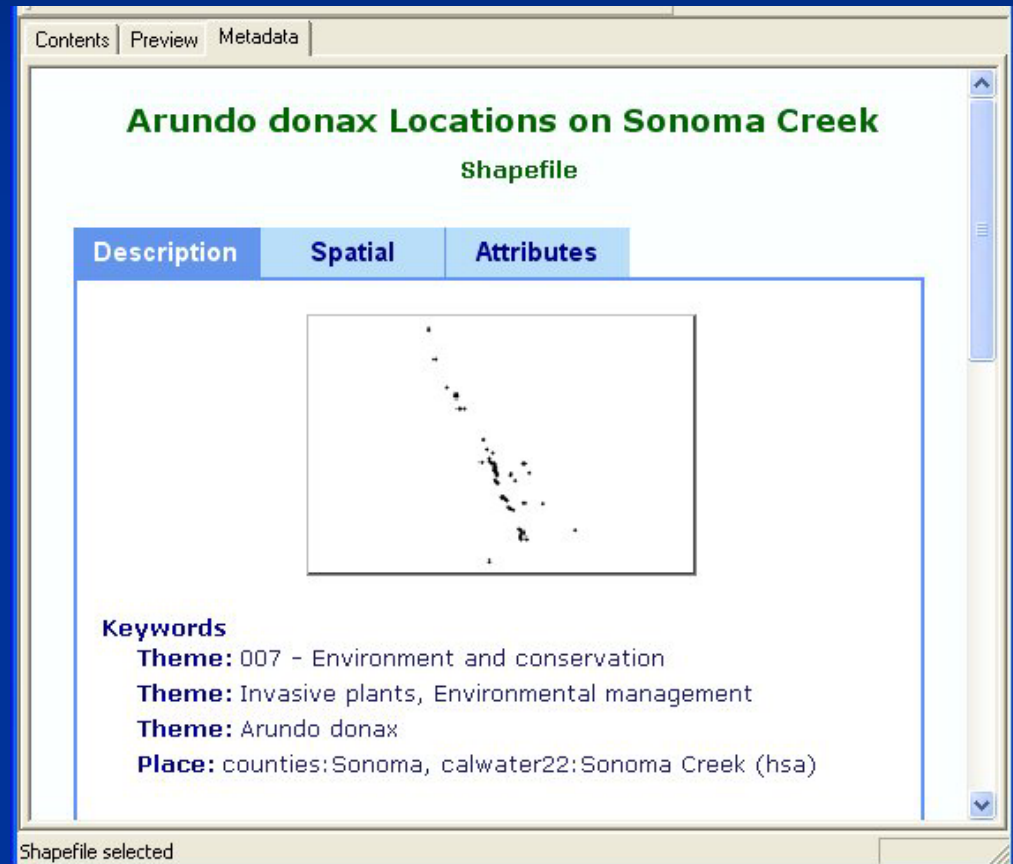
- Go Live

Key Metadata Concepts for ArcCatalog

- Documentation
- Properties
- Metadata style sheets and editors

Documentation

- Information manually added in FGDC Editor
- Title, abstract, keywords, attribute descriptions, geoprocessing steps, sources used, etc.
- Shows up in all three tabs



Properties

- Inherent in the data
 - Coordinate system information
 - Bounding coordinates
 - Attribute names
 - etc.
- Read by A/C directly from the data file
- Updated automatically as file changes
- Many properties are ESRI Profile elements (i.e., local coordinates, file or table name, etc.)

Stylesheet: FGDC ESRI

Contents Preview Metadata

Sonoma Valley Fish Passage Barri
Shapefile

Description **Spatial** Attributes

Horizontal coordinate system
Projected coordinate system name:
NAD_1983_StatePlane_California_II_FIPS_0402_Feet
Geographic coordinate system name: GCS_North_Ar

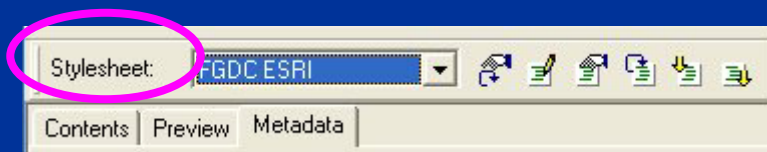
Details

Bounding coordinates
Horizontal
In decimal degrees
West: -122.576927
East: -122.421838
North: 38.442527
South: 38.265143
In projected or local coordinates
Left: 6396424.601666
Right: 6440546.712696
Top: 1923262.240910
Bottom: 1858903.804975

Stylesheets and Editors for FGDC and ISO Standards

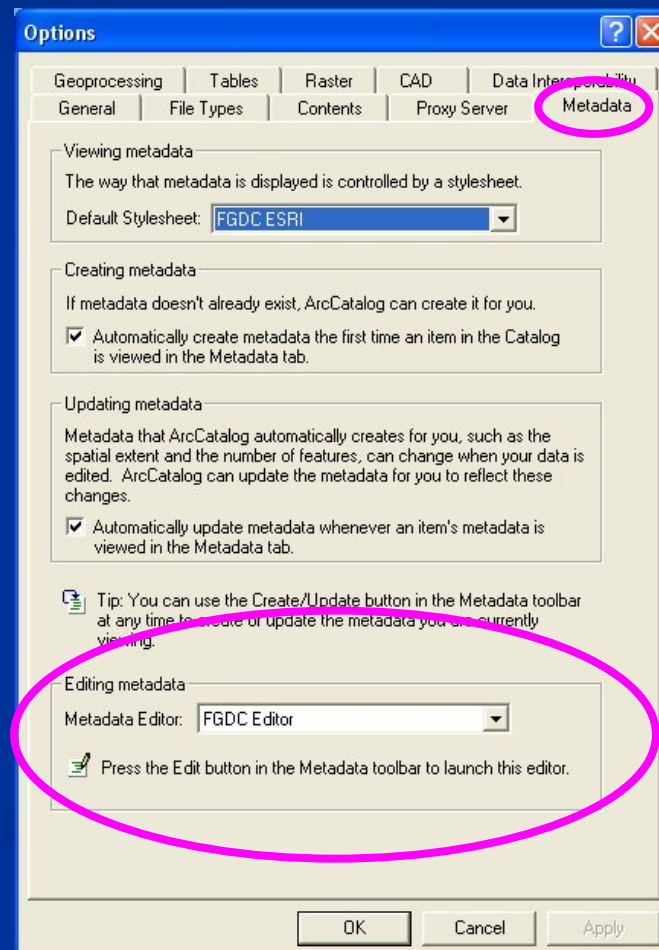
- Tools|Options|Metadata to choose editor

- Stylesheets display metadata elements for standard chosen



- Editor shows only those elements which can be edited

- A/C metadata is xml file



How Much is Enough?

- Something is ALWAYS better than nothing! (Auto-populated properties only = Nothing!!!)
- Make metadata incrementally, start with quick notes early-on, fill in more as you go
- Depends on purpose: metadata for discovery or for use?

How Much is Enough?

- Depends on audience and purpose for your metadata
 - Yourself - internal project documentation (trail of bread crumbs)
 - Organization - resource data to be used by anyone in the organization
 - Client – project deliverable
 - Publication - shared in a public catalog, clearinghouse, published to ArcIMS Metadata Service...
 - BIOS (<http://bios.dfg.ca.gov/>) requires the full standard
 - CEIC focuses on 3 sections of FGDC standard

Mistakes Remedied, Mysteries Solved

■ Troublesome Tabs

- Data Set Credit vs. Originator
- Online Linkage (Unwanted references to location of data, replace with “none”)
- Publication Information (on 2nd tab!)
- Temporal Keyword (not meant to be a date!)
- Information Source/Process Step (not used often enough)

■ Top Ten Metadata Errors:

<http://www.fgdc.gov/metadata/documents/top10metadataerrors.pdf>

Thesaurus Name/Theme Keywords

- ISO 19115 Topic Categories (thesaurus)
 - Environment and Conservation (keyword)
- CERES Theme (thesaurus)
 - Natural environment (keyword)
- NONE (thesaurus)
 - Environment (keyword)

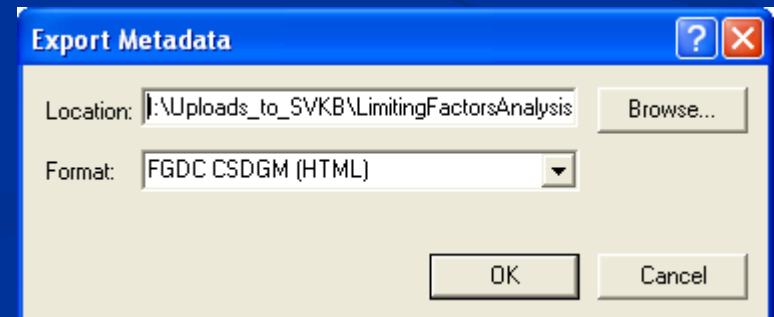
Thesaurus Name/Place

Keywords

- CalWater 2.2.1
 - Sonoma Creek hsa
- GNIS
 - Sonoma Creek
- NONE
 - Downtown Sonoma

Make Some Metadata

1. Load your data, if you have it
2. Select Railroads shapefile, or your file, in Catalog Tree
3. Click Metadata Tab on right
4. Open metadata editor and fill in as many fields as you can.
5. Export completed metadata to “Spatial Data” folder.



CEIC Tour and Hands-on Practice

- What is CEIC and what's in it
- Make your CEIC catalog
- Use interactive form to make metadata
- Upload your FGDC xml record
- Make a link to your catalog or individual records from your website or in a document

California Environmental Information Clearinghouse (CEIC)

CERES – CEIC
Live Demo



CEIC

Try it!

Lunch

12:00-1:30