

Making Metadata Work

Using metadata to document your science

August 1st, 2010



Presenter

Best Practices

- Viv Hutchison
 - US Geological Survey – NBII program
 - Metadata Coordinator
 - Location: USGS Western Fisheries Research Center - Seattle, WA USA
 - Contact: vhutchison@usgs.gov
or 206.526.6282 x329



A screenshot of the National Biological Information Infrastructure (NBII) website. The header includes the NBII logo and navigation links: NBII Home, About NBII, Plants, Animals & Other Organisms, Habitats, Ecological Topics, Geographic Perspectives, and Toolkit. A search bar is visible. The main content area features a video player titled "Early Detection, Rapid Response (EDRR)" with a description: "A catalog of online invasive species resources, for identification, reporting, expertise, occurrences, assessment, planning, and response". Below the video are "PROGRAM AREAS" numbered 1 to 5. To the right is a "NBII: RSS Sources" section listing various RSS feeds. The footer contains "National Biological Information Infrastructure at a Glance" and "News from the NBII Blog".

Overall Topics

- Metadata:
 - What it is
 - Why it is valuable
- Clearinghouses
- How to Write Metadata
- Tools and Resources
- Metadata Step by Step
- Understanding the fields in a standard

If you don't know
where you are going...

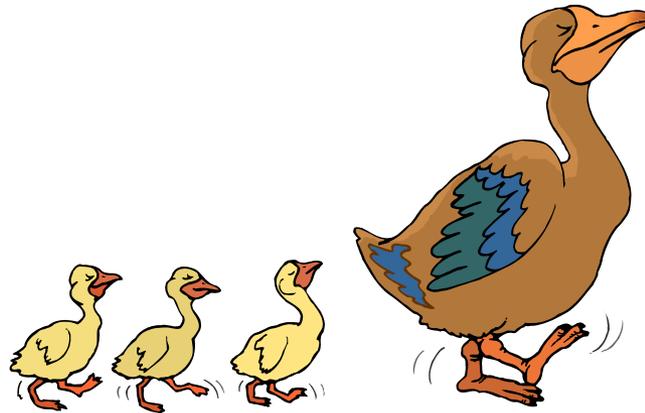
you'll wind up
somewhere
else.

—YOGI BERRA





Getting Started with Metadata...



Topics

Best Practices



- Define metadata
- Look at examples of metadata
- Determine the types of information included in a metadata record

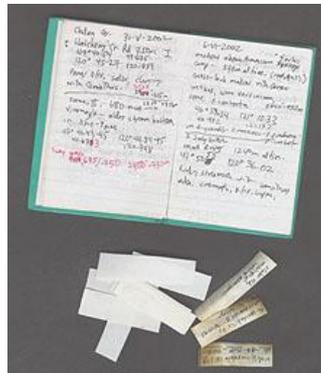
Data Collection



Best Practices

A diagram showing a page from a field notebook with red arrows pointing to various sections, labeled with 'Best Practices' categories:

- AUTHOR AND YEAR:** Points to 'J.M. Bowles 1999'.
- DATE:** Points to '12 May'.
- WEATHER:** Points to 'Overnight low 18°C, clouds 1/2, wind 3, gusting 5. Arizona Upland, Palo Verde, Saguaro, Organop and Centillo near the top of a bajada. A dawn birdwalk along the arroyo was fairly unproductive except for sighting a pair of Coyotes hollering off in the dawn light. One looked very poor compared to birds I encountered during a good year of a party foraging in Palo Verde along the edge of the mesquite near Laguna Colorado singing. Gambel's Quail, Cassin's Sparrow, Curve-billed Thrasher, Black-headed Grosbeak, White-winged Dove were common. I have never seen so many of one on a single bush at this time of year.' (Note: 'white-winged dove' is circled in red).
- LOCATION INFORMATION:** Points to 'Alamo Canyon Damsground, Organop, Cochise, Maricopa, Arizona'.
- PAGE NUMBER:** Points to '10'.
- HABITAT:** Points to 'Arizona Upland, Palo Verde, Saguaro, Organop and Centillo near the top of a bajada'.
- COMMON NAMES:** Points to 'Gambel's Quail, Cassin's Sparrow, Curve-billed Thrasher, Black-headed Grosbeak, White-winged Dove'.
- SCIENTIFIC NAME:** Points to 'white-winged dove'.
- NARRATIVE:** Points to 'During breakfast we watched White-winged Doves and Gambel's Thrashers, taking nectar from flowers at the top of a Saguaro. They plunged their faces right into the flowers and came out covered with pale yellow pollen. About 4 flowers were'.



Birds found near Carlisle Pa		Birds found at Carlisle Pa		First Met	Years
<u>Lane Pigeon</u>		26 th Dec 1839	always	5 th May 1839	13 yrs
XX	" Mourning Dove	3 rd Sept 1838	13 yrs	25 th July 1839	13 yrs
XX	" Pigeon	2 nd Sept 1838	13 yrs	24 th July 1839	13 yrs
XX	" White-throated Sparrow	August 1831	10 yrs	July 1830	13 yrs
XX	" Blue Jay	October 1835	13 yrs	Aug 1832	13 yrs
XX	" Great Horned Owl	"	14 yrs	July 1830	13 yrs
XX	" Nuthatch	11 March 1837	13 yrs	22 April 1839	13 yrs
XX	" Mottled Owl	25 th Oct 1834	2 yrs	18 May 1839	13 yrs
		24 th Dec 1839	13 yrs	4 July 1839	13 yrs
				27 th April 1839	13 yrs
				6 th May 1839	13 yrs
				1 st July 1837	12 yrs
				2 nd Dec 1860	6 yrs

From Field Notes to Data Sets

Average Temperature of Observation for Each Species

Species	Average Temperature	Temperature Standard Deviation	Number of Observations	Minimum Temperature	Maximum Temperature
Northern Red-legged Frog	4.4	---	1	4.4	4.4
Tailed Frog	7.0	3.0	3	4	10
Arizona Toad	10.0	---	1	10	10
Strecker's Chorus Frog	10.5	2.0	11	9	16
Oregon Spotted Frog	11.0	15.5	2	0	22
New Jersey Chorus Frog	11.5	4.5	17	3	22
Wood Frog	12.5	5.5	897	0	28.8
Spring Peeper	13.2	5.6	569	-1	32
Red-legged Frog	13.3	5.9	16	4	27

Best Practices



From Data Sets to Published Papers

Best Practices



What is a Data Set?

Best Practices

- A collection of data that is raw or statistically analyzed
- **Generally data sets are defined as:**
 - Spatial – a collection of logically related features arranged in a prescribed manner such as GIS map layers, water features, etc
 - Tabular – a file, spreadsheet, data in a table
- **Elements in a data set can include:**
 - Values, measures, points, coordinates, conditions, qualities, frequencies, or attributes that are a result of an observational study



What Is Metadata?

Best Practices

- **Metadata is: Data ‘reporting’**
 - **WHO** created the data?
 - **WHAT** is the content of the data?
 - **WHEN** was it created?
 - **WHERE** is it geographically?
 - **HOW** was the data developed?
 - **WHY** was the data developed?



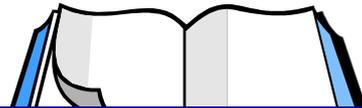
Metadata in the Real World

- *Metadata is all around...*

Best Practices



Nutrition Facts	
Serving Size ½ cup (114g)	
Servings Per Container 4	
Amount Per Serving	
Calories 90	Calories from Fat 30
% Daily Value*	
Total Fat 3g	5%
Saturated Fat 0g	0%
Cholesterol 0mg	0%
Sodium 300mg	13%
Total Carbohydrate 13g	4%
Dietary Fiber 3g	12%
Sugars 3g	
Protein 3g	
Vitamin A 80%	• Vitamin C 60%
Calcium 4%	• Iron 4%
* Percent Daily Values are based on a 2,000 calorie diet. Your daily values may be higher or lower depending on your calorie needs:	
	Calories: 2,000 2,500
Total Fat	Less than 65g 80g
Sat Fat	Less than 20g 25g
Cholesterol	Less than 300mg 300mg
Sodium	Less than 2,400mg 2,400mg
Total Carbohydrate	300g 375g
Dietary Fiber	25g 30g
Calories per gram:	
Fat 9 • Carbohydrate 4 • Protein 4	



- **Author(s)** Boullosa, Carmen.
- **Title(s)** They're cows, we're pigs / by Carmen Boullosa
- **Place** New York : Grove Press, 1997.
- **Physical Descr** viii, 180 p ; 22 cm.
- **Subject(s)** Pirates Caribbean Area Fiction.
- **Format** Fiction

Working With Data

Best Practices

- When you ***provide*** data to someone else, what types of information would you want to include with the data?



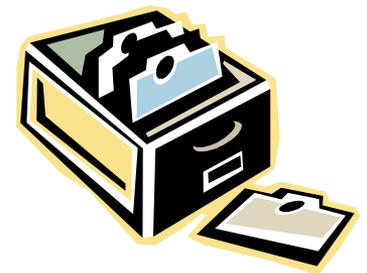
- When you ***receive*** a dataset from an external source, what types of information do you want to know about the data?

Working With Data

Best Practices

- **Providing data:**
 - Why was the data created?
 - What limitations, if any, do the data have?
 - What does the data mean?
 - Who should be cited if someone publishes something that utilized your data?

- **Receiving data:**
 - What are the data gaps?
 - What processes were used for creating the current data?
 - Are there any fees associated with the data?
 - In what scale were the data created?
 - What do the values in the tables mean?
 - What software do I need in order to read the data?
 - What projection is the data in?
 - Can I give this data to someone else?



What is a metadata standard?

Best Practices

A Standard provides a structure to describe data with:

- Common terms to allow consistency between records
- Common definitions for easier interpretation
- Common language for ease of communication
- Common structure to quickly locate information

In search and retrieval, standards provide:

- documentation structure in a reliable and predictable format for computer interpretation
- a uniform summary description of the data set



What does a metadata record look like?

Best Practices

North American Breeding Bird Survey (BBS)

Identification Information:

Citation:

Citation Information:

Originator: Patuxent Wildlife Research Center, Biological Resources Division, U.S. Geological Survey (USGS)

Publication Date: 1997

Title:

North American Breeding Bird Survey (BBS)

Publication Information:

Publication Place: Laurel, MD

Publisher:

Patuxent Wildlife Research Center, Biological Resources Division, U.S. Geological Survey (USGS)

Other Citation Details:

This metadata file can be found at: ftp://cameron.cr.usgs.gov/pub/nbii_metadata/brdpwrc0004.txt (text format) and ftp://cameron.cr.usgs.gov/pub/nbii_metadata/brdpwrc0004.html (HTML format) and ftp://cameron.cr.usgs.gov/pub/nbii_metadata/brdpwrc0004.sgml (SGML format).

Description:

Abstract:

The North American Breeding Bird Survey (BBS), which is coordinated by the Biological Resources Division and Canadian Wildlife Service, is a primary source of population trend and distribution information for most species of North American birds. The BBS was initiated during 1966 by Chan Robbins and his associates at the Patuxent Wildlife Research Center to monitor the populations of all breeding bird species across the continental U.S., Canada, and Alaska. Approximately 2200 skilled observers participate in the survey each year. The BBS has accumulated 30 years of data on the abundance, distribution, and trends for more than 400 species of birds. These data are widely used by researchers, various federal and state agencies, non-governmental organizations, and the general public. Analyses of BBS data have been instrumental in the development of innovative approaches for analyzing trends of wildlife populations.

Purpose:

In the 1960's, chlorinated hydrocarbon pesticides and similar poisons were widely used for spraying not only killed insects but also killed birds, raising serious concerns over its effects on bird populations. Unfortunately, no long-term regional or continental population data were available for birders to demonstrate declines in bird populations. The Bird Breeding Survey has provided information on bird population trends. Robbins et al. (1986) provided the first continental trends of songbirds based on BBS data. When viewed at continental or regional scales, these trends show the relative abundance of species that are well sampled by the BBS. In addition, the survey has provided temporal patterns in trends. Populations of permanent resident and short-distance migrants are generally stable or increasing, while populations of long-distance migrants are generally declining.

Place Keyword: Mexico

Taxonomy:

Keywords/Taxon:

Taxonomic Keyword Thesaurus: none

Taxonomic Keyword: birds

Taxonomic Keyword: breeding birds

Taxonomic Classification:

Taxon Rank Name: Kingdom

Taxon Rank Name: Animalia

Taxonomic Classification:

Taxon Rank Name: Phylum

Taxon Rank Name: Chordata

Taxonomic Classification:

Taxon Rank Name: Class

Taxon Rank Name: Aves

Taxonomic Classification:

Taxon Rank Name: Order

Taxonomic Classification:

Taxon Rank Name: Family

Taxonomic Classification:

Taxon Rank Name: Genus

Taxonomic Classification:

Taxon Rank Name: Species



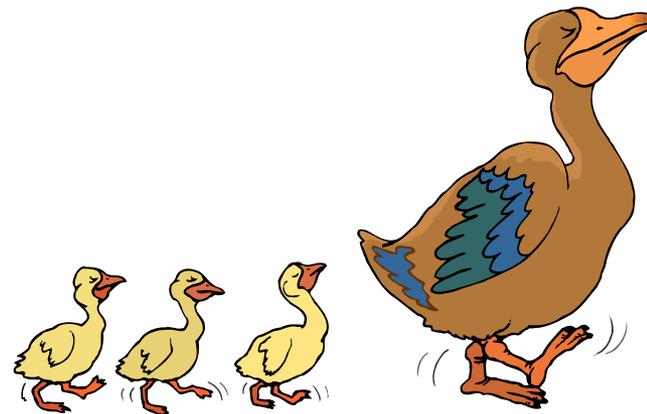
Multiple metadata standards exist

Best Practices

- **Ecological Metadata Language (EML)**
 - Emphasis on ecology
- **Content Standard for Digital Geospatial Metadata (CSDGM)**
 - Federal Geographic Data Committee (FGDC)
 - Emphasis on geospatial data
- **Biological Data Profile (BDP) of the CSDGM**
 - Emphasis on biological data (and geospatial)
- **ISO 19115 Geographic information: Metadata**
 - Emphasis on geospatial data and services
- **Dublin Core Element Set**
 - Emphasis on web resources, publications
- **Darwin Core**
 - Emphasis on museum specimens



The Value of Metadata: Why Do It?



Topics

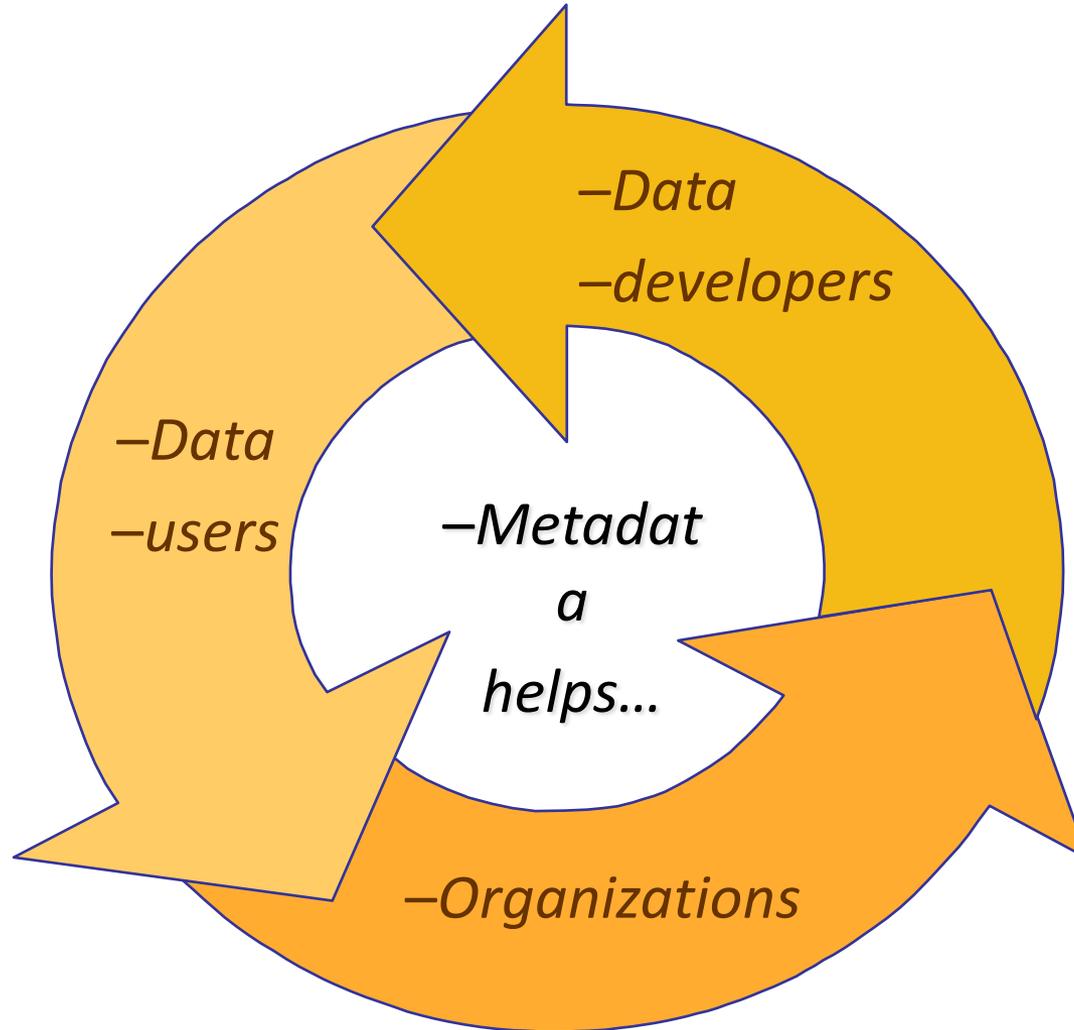
Best Practices

- Determine the value of metadata
- Define reasons for a standard



Metadata has value to all:

Best Practices



What is the value to *Data Developers*?

Best Practices

- Metadata allows data developers to:
 - Avoid data duplication
 - Share reliable information
 - Publicize efforts – promote scientist's work and contributions to a field of study
 - Reduce workload



What is the value to *Data Users*?

Best Practices

- Metadata gives a user the ability to:
 - Search, retrieve, and evaluate dataset information from both inside and outside an organization
 - Find data - determine what data exist for a geographic location and/or topic
 - Determine applicability – decide if a data set meets your needs
 - Discover how to acquire the data set you identified, process and use the data set



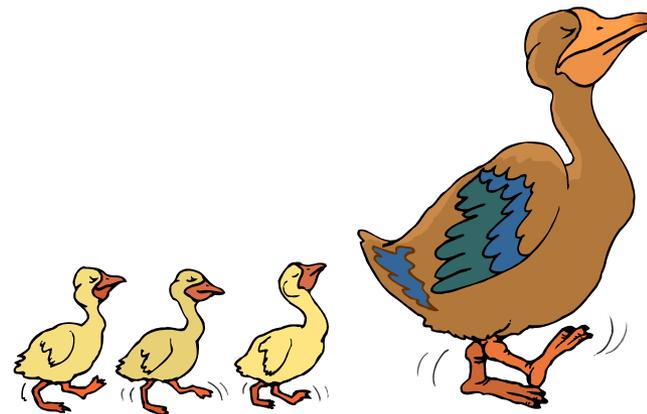
What is the value to *Organizations*?

Best Practices

- Metadata helps ensure an organization's investment in data:
 - Documentation of data processing steps, quality control, definitions, data uses, and restrictions
- Transcends people and time:
 - offers data permanence
 - creates institutional memory
- Advertises an organization's research
 - creates possible new partnerships and collaborations thru data sharing



Metadata Clearinghouses



Topics

Best Practices



- Define Clearinghouse
- Examples
- How to use the NBII Clearinghouse

Metadata Clearinghouses

Best Practices

- A metadata clearinghouse is a location — typically accessed through the Internet — to search for spatial data sets
- Clearinghouses make metadata records easy to find

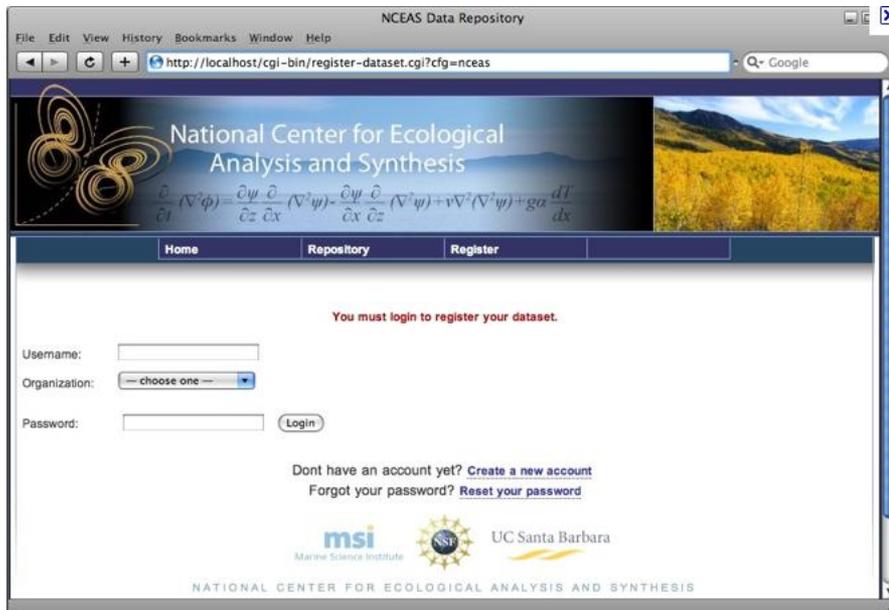


Examples of Clearinghouses:

NBII Clearinghouse
<http://metadata.nbio.gov>



The screenshot shows the NBII Metadata Clearinghouse search page. At the top, there is a banner with the NBII logo and the text "National Biological Information Infrastructure" and "NBII Metadata Clearinghouse | HELP". Below the banner are two tabs: "Simple Search" and "Advanced Search". The main heading is "Search All Records For". There is a search input field, a "SEARCH" button, and a "Results/Page" dropdown menu set to "10". A hint below the search field reads: "Hint: boolean operators, wildcards and phrases are allowed. ex: precipitation or (rain* and *moisture content)". Below the search area, it says "Query being built:" followed by a blurred query string and "Not Editable". A "CLEAR QUERY" link is at the bottom left of the query area.



The screenshot shows the NCEAS Data Repository registration page. The browser address bar shows "http://localhost/cgi-bin/register-dataset.cgi?cfg=nceas". The page header includes "National Center for Ecological Analysis and Synthesis" and a mathematical equation:
$$\frac{\partial}{\partial t} (\nabla^2 \phi) - \frac{\partial \psi}{\partial z} \frac{\partial}{\partial x} (\nabla^2 \psi) - \frac{\partial \psi}{\partial x} \frac{\partial}{\partial z} (\nabla^2 \psi) + v \nabla^2 (\nabla^2 \psi) + g \alpha \frac{dT}{dx}$$
. The navigation menu has "Home", "Repository", and "Register". A red message says "You must login to register your dataset." There are input fields for "Username:", "Organization:" (with a dropdown menu), and "Password:". A "Login" button is next to the password field. Below the login fields, there are links: "Dont have an account yet? [Create a new account](#)" and "Forgot your password? [Reset your password](#)". At the bottom, there are logos for "msi Marine Science Institute", "UC Santa Barbara", and "NATIONAL CENTER FOR ECOLOGICAL ANALYSIS AND SYNTHESIS".

 [My NBII](#) | [NBII Disclaimer and Privacy Statement](#) | [Feedback Form](#) | [Email Us](#) | [Security Warning](#)

Metacat
<http://knb.ecoinformatics.org>



Searching the NBII Clearinghouse:

The screenshot shows the top navigation bar with the NBII logo (National Biological Information Infrastructure) and a bird illustration. Below the logo is the text "NBII Metadata Clearinghouse" and a "HELP" button. The main search area has two tabs: "Simple Search" (selected) and "Advanced Search". The search prompt is "Search All Records For". There is a search input field, a "SEARCH" button, and a "Results/Page" dropdown menu set to "10". A hint text reads: "Hint: boolean operators, wildcards and phrases are allowed. ex: precipitation or (rain* and \"moisture content\")". Below the search area is a section titled "Query being built:" with a large, faint, illegible query string. A "CLEAR QUERY" link is located at the bottom left of this section.

nbii National Biological Information Infrastructure

NBII Metadata Clearinghouse | HELP

Simple Search | Advanced Search

Search All Records For

SEARCH

Hint: boolean operators, wildcards and phrases are allowed.
ex: precipitation or (rain* and "moisture content")

Results/Page
10

Query being built: _____

Not Editable

[CLEAR QUERY](#)



[My NBII](#) | [NBII Disclaimer and Privacy Statement](#) | [Feedback Form](#) | [Email Us](#) | [Security Warning](#)

Advanced Search

NBII Metadata Clearinghouse

HELP

Simple Search

Advanced Search

activities



Fielded Search

FullText OR

FullText OR

FullText OR

Add more options

Help | clear

Date Search

Data Collection Date

Publication Date

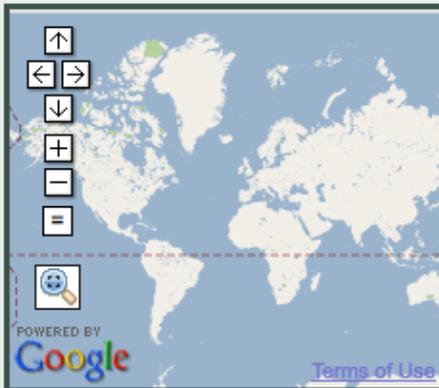
Either

during thru

mm/dd/yyyy mm/dd/yyyy

Help | clear

Geographic Search



List Areas in:

USA WORLD

Select from list

Search Area:

overlaps encloses

North

West East

South

*Click on to select an area

Place:

Help | clear

Format

All

Maps and Data

Publications

Tools and Software

Data providers

All

U.S Geological Survey (USGS)

- BRD Bibliographic Metadata Node
- Columbia Environmental Research Center Metadata Node
- Fire Research and Management Exchange System (FRAMES)
- Forest and Rangeland Ecosystem Science Center Metadata Node
- Great Basin Information Project
- Hawaii Biodiversity and Mapping Program
- National Gap Analysis Program Metadata Node
- National Wetlands Research Center

Query being built:

Not Editable

[CLEAR QUERY](#)

Results



National Biological Information Infrastructure

Metadata Summary

Email

Bookmark

RSS
Feed

Help

Your search found: 28 documents.

You searched for: fullText:cattle and ranch AND datasource:(all)

Filter by resource type	Filter by decade	Filter by originator
Maps and Data (19) others (9)	2000 (15) 1990 (13)	Gosz, James (10) Moore, Doug (8) usgs (6) Parmenter, Robert (2) Principal investigator(s) (1)

Viewing Documents 1 - 10 out of 28

[Prev](#) [1](#) [2](#) [3](#) [Next](#)

[Return to Search](#)

Sort By:

Index Rank

Pubdate

Source

Filter by data providers

[Long Term Ecological Research \(LTER\) Network \(20\)](#)

[Metadata Clearinghouse](#)

ELK AND BISON GRAZING ECOLOGY AND MANAGEMENT IN GREAT SAND DUNES NATIONAL PARK AND PRESERVE

10/01/2004 - 05/31/2008

Data provider: METADATA CLEARINGHOUSE PRINCIPAL NODE

The establishment of Great Sand Dunes National Park and Preserve and the new Baca National Wildlife Refuge in the San Luis Valley

Results

Viewing Documents 1 - 10 out of 28

[Prev](#) [1](#) [2](#) [3](#) [Next](#)

[Return to Search](#)

Sort By:

Index Rank

Pubdate

Source

Filter by data providers

[Long Term Ecological Research \(LTER\) Network \(20\)](#)

[Metadata Clearinghouse Principal Node \(6\)](#)

[Baruch Coastal Monitoring \(1\)](#)

[California Environmental Information Catalog \(1\)](#)

ELK AND BISON GRAZING ECOLOGY AND MANAGEMENT IN GREAT SAND DUNES NATIONAL PARK AND PRESERVE

10/01/2004 - 05/31/2008

Data provider: METADATA CLEARINGHOUSE PRINCIPAL NODE

The establishment of Great Sand Dunes National Park and Preserve and the new Baca National Wildlife Refuge in the San Luis Valley of Colorado was one of the most significant land conservation actions in the U.S. west in recent years. The action was a result of cooperation between the NPS, FWS, BLM, USFS, and The Nature Conservancy. The new national park will consist of 107,265 acres, the new national preserve, 41,872 acres, and the new national wildlife refuge, 92,180 acres. The area encompassed by this designation protects a number of natural wonders and features including a unique ecosystem o



[View full metadata](#)

STATE OF CALIFORNIA GRAZING ALLOTMENTS

01/01/2003 - 12/31/2003

Data provider: CALIFORNIA ENVIRONMENTAL INFORMATION CATALOG

BLM grazing allotments in California.



[View full metadata](#)

ELK AND BISON GRAZING ECOLOGY AND MANAGEMENT IN GREAT SAND DUNES NATIONAL PARK AND PRESERVE, ELK BODY CONDITION AND LOCATION DATA FROM 2005 TO 2008.

01/17/2005 - 05/08/2007

Data provider: METADATA CLEARINGHOUSE PRINCIPAL NODE

The establishment of Great Sand Dunes National Park and Preserve (GRSA) and the new Baca National Wildlife Refuge in the San Luis Valley of Colorado was one of the most significant land conservation actions in the U.S. west in recent years. The action was a result of cooperation between the NPS, FWS, BLM, USFS, and The Nature Conservancy. The new national park will consist of 107,265 acres, the new national preserve, 41,872 acres, and the new national wildlife refuge, 92,180 acres. The area encompassed by this designation protects a number of natural wonders and features including a unique ecos

Results



National Biological Information Infrastructure

Metadata Report

[Bookmark](#) [Email](#)

You searched for: **fulltext:cattle and ranch and datasource:(all)**

[Return to Search](#)

[Return to Results](#)

[View this page in FGDC](#)

Title: Elk and bison grazing ecology and management in Great Sand Dunes National Park and Preserve

Investigator(s): U.S. Geological Survey, Fort Collins Science Center (FORT)

Status: In work

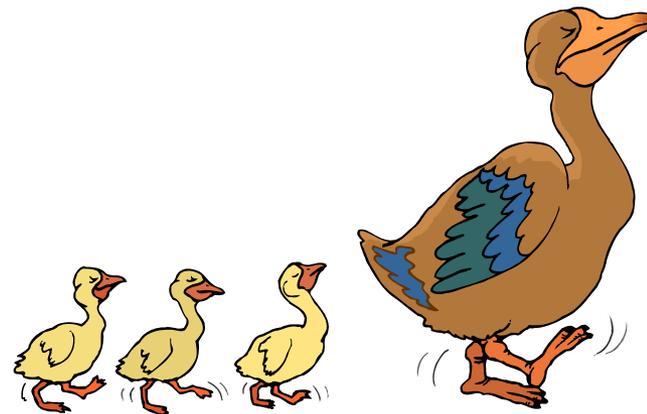
Access Restriction: Contact principal investigator after completion of project

Data Center Contact: Phone: Email:

Keyword(s):	Theme Keyword Thesaurus	Theme Keyword
	None	ecology management ELK bison

Abstract: The establishment of Great Sand Dunes National Park and Preserve and the new Baca National Wildlife Refuge in the San Luis Valley of Colorado was one of the most significant land conservation actions in the U.S. west in recent years. The action was a result of cooperation between the NPS, FWS, BLM, USFS, and The Nature Conservancy. The new national park will consist of 107,265 acres, the new national preserve, 41,872 acres, and the new national wildlife refuge, 92,180 acres. The area encompassed by this designation protects a number of natural wonders and features including a unique ecosystem of natural sand dunes, the entire

How to write good metadata



Topics

Best Practices

- Steps for preparing quality metadata
- Tips for writing good metadata



Steps to Create Quality Metadata

Best Practices

- Organize your information
- Write your metadata
- Review for accuracy and completeness
- Have someone else read your file
- Revise it, based on comments from your reviewer
- Review it once more before you publish it



Tips for Writing Good Metadata

Best Practices

Think about the long-term effects:

- Don't use jargon
- Define technical terms and acronyms:
CA, LA, GPS, GIS
- Clearly state data limitations
- Cite examples



Tips for Writing Good Metadata

Best Practices

Titles, Titles, Titles...

- Titles are *critical* in helping readers find your data
 - While individuals are searching for the most appropriate datasets on the clearinghouses, they are most likely going to use the title as the first criteria to determine if a dataset meets their needs.
 - Treat the title as the opportunity to sell your dataset.
- A *complete title* includes: *What, Where, When, Who, and Scale*
- An *informative title* includes: *topic, timeliness of the data, specific information about place and geography*



Tips for Writing Good Metadata



- Which title is better?

- *Rivers*

OR

- *Greater Yellowstone Rivers from 1:126,700 U.S. Forest Service Visitor Maps (1961-1983)*

Greater Yellowstone (where) Rivers (what) from 1:126,700 (scale) U.S. Forest Service (who) Visitor Maps (1961-1983) (when)

Tips for Writing Good Metadata

Best Practices

Be specific and quantify when you can: the goal of a metadata record is to give the user enough information to know if they can use it without contacting the dataset owner.



Vague: *We checked our work and it looks complete.*

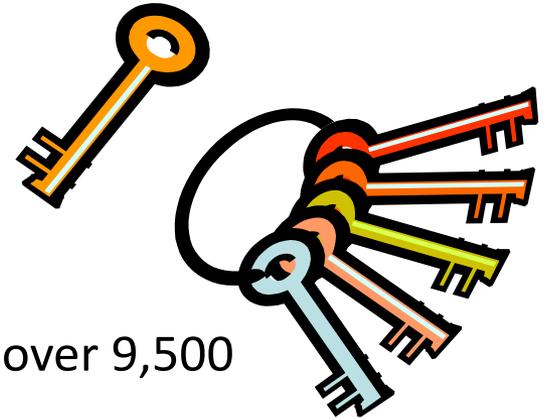
Specific: *We checked our work using 3 separate sets of check plots reviewed by 2 different people. We determined our work to be 95% complete based on these visual inspections.*

Tips for Writing Good Metadata

Best Practices

Select keywords wisely

- Use unambiguous words
- Use descriptive words
- Fully qualify geographic locations
- Use thesauri whenever possible
 - Example: NBII Biocomplexity Thesaurus (over 9,500 terms (<http://thesaurus.nbii.gov>))



Tips for Writing Good Metadata

Best Practices

Remember: a computer will read your metadata

- Don't use symbols that might be misinterpreted
 - ! @ # % { } | / \ < > ~
- Don't use characters with dual interpretations
 - ie: < > signs are used in html coding
- Don't use tabs or indents
- When copying and pasting from other sources, use a text editor to eliminate hidden characters

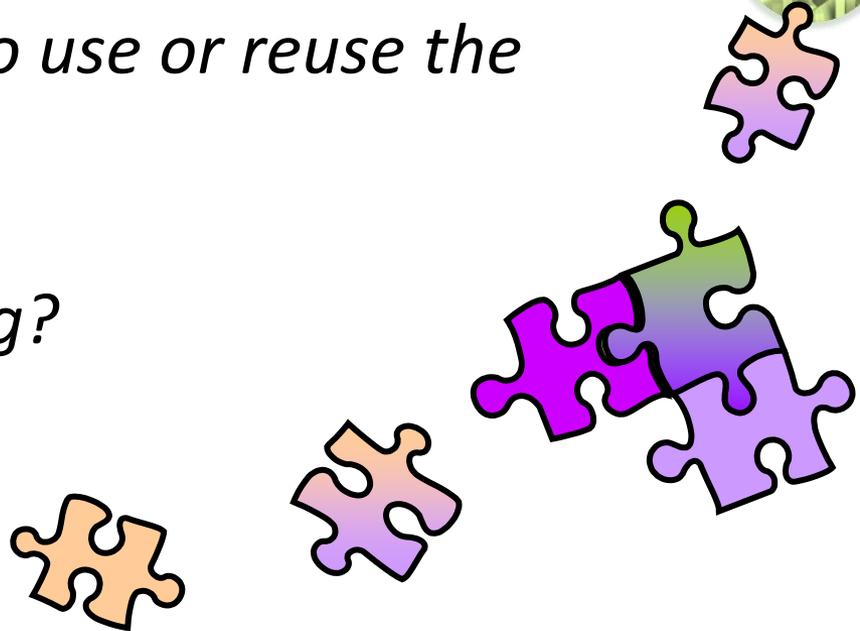


Tips for Writing Good Metadata

Best Practices

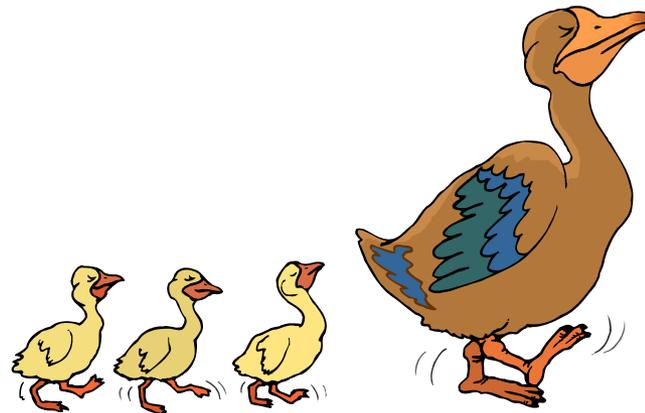
*****Review your final product*****

- *Does the documentation present all the information needed to use or reuse the data?*
- *Are any pieces missing?*





Metadata Tools and Resources



ArcCatalog (ArcGIS)

The image displays two overlapping dialog boxes in ArcCatalog. The background dialog, titled "Editing 'NatureServe Element Occurrence Data for Rare and Endangered Species in the United States and Canada'", has tabs for Identification, Data Quality, Data Organization, Spatial Reference, Entity Attribute, Distribution, and Metadata Reference. The "Identification" tab is active, showing sub-tabs: General, Contact, Citation, Time Period, Status, Spatial Domain, Keywords, Browse Graphic, Security, and Cross Reference. The "Description" section includes fields for Abstract, Purpose, Language, and Supplemental Information. The "Access Constraints" section includes fields for Use Constraints, Data Set Credit, Native Data Set Environment, and Native Data Set Format. Buttons for Save, Cancel, and Help are at the bottom.

The foreground dialog, also titled "Editing 'NatureServe Element Occurrence Data for Rare and Endangered Species in the United States and Canada'", has tabs for Identification, Data Quality, Data Organization, and Spatial Reference. The "Identification" tab is active, showing sub-tabs: Entity Type, Attribute, General, Dates, and Attribute Domain Values. The "General" sub-tab is active, showing fields for Label (GRANK), Type, Width, Precision, Indexed, Definition (Global Rank - The NatureServe Conserv), and Definition Source (NatureServe - Heritage methodology). Buttons for Save, Cancel, and Help are at the bottom.

Metavist

Metavist - natsrv_d_eo.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) Publication Date

Lynn Kutner (ed.), NatureServe

Title

NatureServe Element Occurrence Data for Rare and Endangered Species in the

Abstract

NatureServe, in collaboration with its member Natural Heritage Programs and C maintains a database of rare and imperiled species and plant communities across Canada. The Element Occurrence (EO) records that form the core of the Nature information on the location, status, characteristics, numbers, condition, and diversity using established Natural Heritage Methodology developed by Nature's Conservancy (TNC). An Element Occurrence (EO) is an area of land and/or water community is, or was, present. An EO should have practical conservation value by potential continued (or historical) presence and/or regular recurrence at a given

Purpose

The Element Occurrence (EO) data are collected and maintained by the Nature Conservation Data Centres using a standardized methodology to provide accurate conservation information to public and private agencies and individuals. The pr

Best Practices



Contact Information

The primary contact is a(n):

Person Organization

Contact Person

Shara Howie

Contact Organization (optional)

NatureServe

Contact Position (optional)

Director, Heritage Data Services

Contact Address

mailing and physical address: Arlington

Add Edit Delete

Contact Voice Telephone

(703) 908-1855

Contact TDD/TTY Telephone (optional)

Contact Fax Telephone (optional)

(703) 908-1917

Contact E-Mail Address (optional)

shara_howie@natureserve.org

Hours of Service (optional)

Contact Instructions (optional)

OK Cancel

NOAA Mermaid: An online tool



NOAA Satellite and Information Service
National Environmental Satellite, Data, and Information Service (NESDIS)



National Coastal Data
Development Center

home » mermaid metadata resources » tools

Home

About Us

Metadata Resources

Regional Offices

Projects

training

tools

additional information

references

Metadata Enterprise Resource Management Aid (MERMAid)



Getting Started with MERMAid

- [Request an Account](#)
- [V1.2 Getting Started Guide](#)
(PDF 5 MB)
- [V1.2 Getting Started Guide](#)

NCDDC provides coastal data resources (organizations and individuals) with a tool to develop, validate, manage and publish metadata records via secure internet access. The Metadata Enterprise Resource Management Aid (MERMAid) allows users/data providers to establish unlimited metadata databases to organize their metadata records any way they see fit (i.e. by program, project, data type, personnel). Some of the key features in MERMAid include (1) user-defined roles and permissions at the metadata management and database levels; (2) change tracking; and (3) enhanced validation. Also, your existing FGDC compliant metadata (in XML format) can be ingested into and managed through MERMAid.

In the near future, NCDDC will be shifting from its current metadata catalog to a knowledge base catalog. MERMAid will play an integral role in this transition. To better leverage these new capabilities, enhanced search and discovery tools will be made available to the public and metadata managers that will provide powerful drill-down features.

NCDDC Services

[Site Map](#)
[Metadata Search](#)
[Middleware Technology](#)
[Download](#)
[Interactive Maps](#)
[Regional Ecosystems](#)
[C-SIDE](#)

NOAA Services

[Central Library](#)
[Photo Library](#)
[Video Library](#)
[Visualization Lab](#)
[Education Resources](#)
[National Locator](#)
[NOAA In Your State](#)

SSC Visitor Services

[Regional Map](#)
[Highway Map](#)
[Additional Information](#)

<http://www.ncddc.noaa.gov/Metadata/Tools>

Morpho: Ecological Metadata

Best Practices



Morpho
Data Management for Ecologists

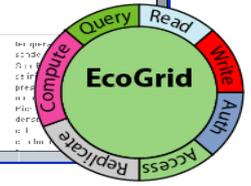
Accession	Deployment	Col Name	Col	Name	Temperature Monitored
1530-C1	08/7/2002	1530-32	1530-02	AP15	1530
1530-C1	08/7/2002	1530-32	1530-02	AP15	1530
1530-C1	08/7/2002	1530-32	1530-02	AP15	1530
1530-C1	08/7/2002	1530-32	1530-02	AP15	1530
1530-C1	08/7/2002	1530-32	1530-02	AP15	1530
1530-C1	08/7/2002	1530-32	1530-02	AP15	1530
1530-C1	08/7/2002	1530-32	1530-02	AP15	1530
1530-C1	08/7/2002	1530-32	1530-02	AP15	1530
1530-C1	08/7/2002	1530-32	1530-02	AP15	1530
1530-C1	08/7/2002	1530-32	1530-02	AP15	1530

Biocomplexity Data Search

SEARCH FOR DATA ON THE KND

456 data packages found

Title	Contacts	Organizations
Bates and Hughes	Yves F. de	
Productivity, Diversity and Soil Data from two North American Grasslands	Doe	

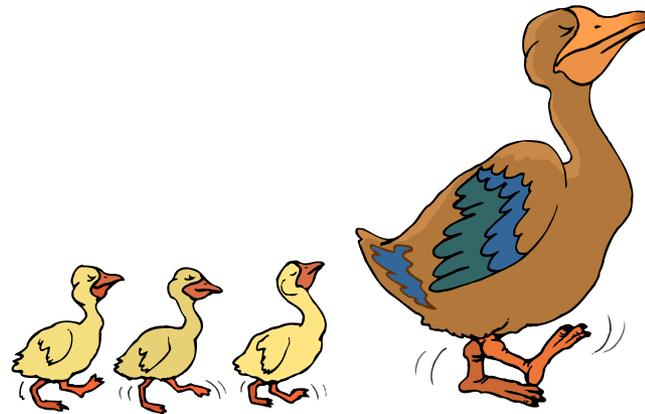


Kepler

Gene Accession Number → Gene Accession Number Sequences Display

<http://knb.ecoinformatics.org>

The Metadata Process: Step by Step



Understand the standard you are using

Best Practices

**Orient yourself
with
documentation
about your
metadata
standard of
choice**



National Spatial Data Infrastructure

Biological Data Profile Workbook (For use with FGDC-STD-001-1998)

Federal Geographic Data Committee

March 30, 2001

Federal Geographic Data Committee

Department of Agriculture • Department of Commerce • Department of Defense • Department of Energy
Department of Health & Human Services • Department of Housing and Urban Development
Department of the Interior • Department of Justice • Department of State
Department of Transportation • Environmental Protection Agency
Federal Emergency Management Agency • Library of Congress
National Aeronautics and Space Administration • National Archives and Records Administration
National Science Foundation • Tennessee Valley Authority



Use established thesauri for keywords

Browse the Biocomplexity Thesaurus

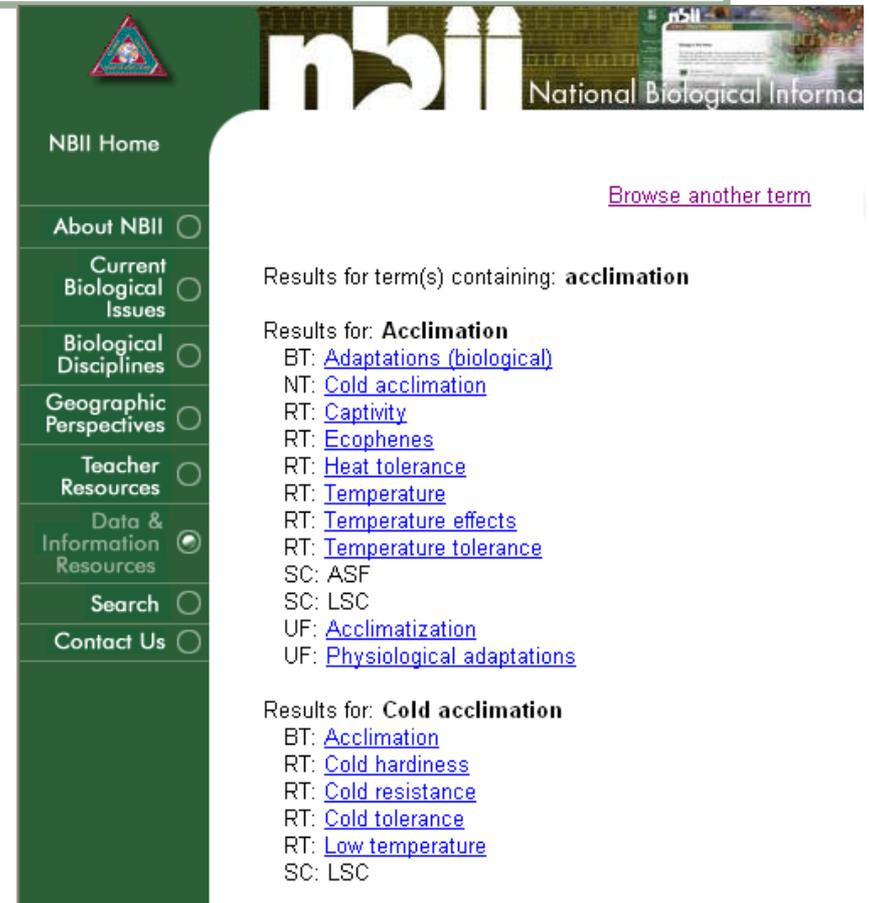
The Biocomplexity Thesaurus was developed through a partnership between the NBII and [CSA](#), a leading bibliographic database provider.

Please enter a term to browse:

Check Thesaurus

- Search on single term or a phrase. (e.g. biodiversity, invasive species)
- Search performs automatic stemming for prefixes and suffixes (e.g. entering [face](#) will yield cell surface, air-ice interface, face, and other variations; entering [bio](#) will yield abiotic factors, biometrics, biodiversity, ectosymbionts, etc.)
- [More about the Biocomplexity Thesaurus](#)

<http://thesaurus.nbii.gov>



The screenshot shows the NBII website interface. At the top right, there is a banner for the National Biological Information Infrastructure (NBII). On the left side, there is a vertical navigation menu with the following items: NBII Home, About NBII, Current Biological Issues, Biological Disciplines, Geographic Perspectives, Teacher Resources, Data & Information Resources (which is selected), Search, and Contact Us. The main content area displays the search results for the term 'acclimation'. It includes a link to 'Browse another term' and two sections of results. The first section is for 'Acclimation' and lists related terms: Adaptations (biological), Cold acclimation, Captivity, Ecophenes, Heat tolerance, Temperature, Temperature effects, and Temperature tolerance. The second section is for 'Cold acclimation' and lists related terms: Acclimation, Cold hardiness, Cold resistance, Cold tolerance, and Low temperature.

Enter information using a metadata creation tool

Best Practices

The screenshot shows the Metavist software interface. The title bar reads "Metavist" and includes standard window controls. Below the title bar is a menu bar with "File", "Options", and "Help". A toolbar contains icons for a folder and a document. The main interface is divided into several tabs: "Identification", "Data Quality", "Spatial Data Org", "Spatial Reference", "Entity & Attribute", "Distribution", and "Metadata Ref". Under the "Identification" tab, there are sub-tabs: "Basic Info", "Spatial Domain", "Keywords", "Taxonomy", "Access", "Analytical Tools", and "Miscellaneous". The "Basic Info" sub-tab is active, showing a "Citation for the data set" section with fields for "Author(s)", "Publication Date", and "Title", along with an "Edit Citation" button. Below this is an "Abstract" section with a large text area and a "Purpose" section with a smaller text area. The interface is yellow and green.



Keep and Share your metadata

Best Practices

- Keep your metadata with your dataset
- Select a Clearinghouse

A typical process:

- Upload records to a web-accessible folder
- Records should be in .xml format
- Clearinghouse harvests records



Search for records in Clearinghouses



A Pilot Catalog For Earth Observations

DataONE Metadata Clearinghouse
HELP

Simple Search
Advanced Search

Search All Records For

Hint: boolean operators, wildcards and phrases are allowed.
ex: precipitation or (rain* and "moisture content")

Results/Page

Query being built:

Not Editable

CLEAR QUERY

Best Practices



Viewing Documents 1 - 10 out of 28

Prev 1 2 3 Next

Return to Search

Sort By: Index Rank Pubdate Source

BISON GRAZING ECOLOGY AND MANAGEMENT IN GREAT SAND DUNES NATIONAL PARK AND PRESERVE 10/01/2004 - 05/31/2008

- [Long Term Ecological Research \(LTER\) Network \(20\)](#)
- [Metadata Clearinghouse Principal Node \(6\)](#)
- [Baruch Coastal Monitoring \(1\)](#)
- [California Environmental Information Catalog \(1\)](#)

Data provider: METADATA CLEARINGHOUSE PRINCIPAL NODE

The establishment of Great Sand Dunes National Park and Preserve and the new Baca National Wildlife Refuge in the San Luis Valley of Colorado was one of the most significant land conservation actions in the U.S. west in recent years. The action was a result of cooperation between the NPS, FWS, ELM, USFS, and The Nature Conservancy. The new national park will consist of 107,265 acres, the new national preserve, 41,872 acres, and the new national wildlife refuge, 92,180 acres. The area encompassed by this designation protects a number of natural wonders and features including a unique ecosystem o

★★★★★★★★

[View full metadata](#)

STATE OF CALIFORNIA GRAZING ALLOTMENTS

01/01/2003 - 12/31/2003

Data provider: CALIFORNIA ENVIRONMENTAL INFORMATION CATALOG
BLM grazing allotments in California.

★★★★★★★★

[View full metadata](#)

ELK AND BISON GRAZING ECOLOGY AND MANAGEMENT IN GREAT SAND DUNES NATIONAL PARK AND PRESERVE, ELK BODY CONDITION AND LOCATION DATA FROM 2005 TO 2008.

01/17/2005 - 05/08/2007

Data provider: METADATA CLEARINGHOUSE PRINCIPAL NODE

The establishment of Great Sand Dunes National Park and Preserve (GRNSA) and the new Baca National Wildlife Refuge in the San Luis Valley of Colorado was one of the most significant land conservation actions in the U.S. west in recent years. The action was a result of cooperation between the NPS, FWS, ELM, USFS, and The Nature Conservancy. The new national park will consist of 107,265 acres, the new national preserve, 41,872 acres, and the new national wildlife refuge, 92,180 acres. The area encompassed by this designation protects a number of natural wonders and features including a unique ecos

Who contributes to creating a metadata file?

Best Practices

Single individual or team approach?

- Team Leader / Project Manager
- GIS Specialist
- Field Personnel
- Database Manager
- Science Staff
- Data Analysis Lead



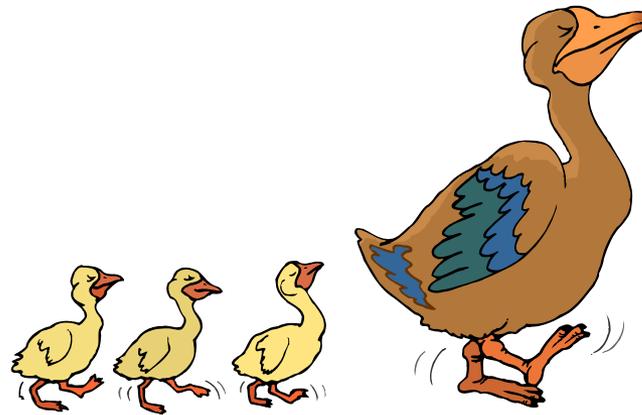
More Metadata Training & Information Online

Best Practices

- FGDC metadata trainer directory
- FGDC metadata training calendar
 - <http://www.fgdc.gov>
- NBII metadata training program
 - <http://www.nbii.gov>
- NOAA Training Materials
 - <http://www.csc.noaa.gov/metadata/>
- Email listserv
 - <http://lists.geocomm.com/mailman/listinfo/metadata>



Understanding the fields in a standard



Metavist

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)

Publication Date

Title

Abstract

Purpose

Time Period of Content

Data are from A single date Multiple dates A range of dates Unknown

The calendar is: Gregorian Geologic

Time Period(s):

Currentness Reference

Time Period of Content tells when the data were current.

Currentness Reference tells how "current" is defined.

Ex. 1: Ground condition – data capture dates as in photography or field data collection.

Ex. 2: Publication date – information was officially recorded, as in a deed.

Ex. 3: Publication date – data from a '1978 USGS Topo map'.

Ex. 4: Observed – observations made in the lab.

If one of the 3 standard options does not fit your needs, type in your own description.

Status

Progress

Updating

Best Practices



Identification Information

Citation Information

Best Practices



Citation Information

Primary Citation

Author(s)

Publication Date

Unknown

Unpublished material

Specify date

Publication Time (optional)

Local time (HHMMSS):

Title

Edition

Data Presentation Form

Series information

Series name:

Issue:

Publication information

Publication place

Publisher

Other Citation Details

Online Linkages [URLs] (optional)

Create Larger Work Citation? Yes No

OK **Cancel**

Spatial Domain Information

The screenshot shows a software interface with a menu bar (File, Options, Help) and a toolbar. Below the toolbar are several tabs: Identification, Data Quality, Spatial Data Org, Spatial Reference, Entity & Attribute, Distribution, and Metadata Ref. Under the 'Spatial Data Org' tab, there are sub-tabs: Basic Info, Spatial Domain, Keywords, Taxonomy, Access, Analytical Tools, and Miscellaneous. The 'Spatial Domain' sub-tab is active, showing a form with the following sections:

- Description of Geographic Extent:** A large text area for a description.
- Bounding Coordinates:** Four input fields for West (longitude), East (longitude), North (latitude), and South (latitude), each with a value of 0.00000.
- Bounding Altitudes (optional):** Two input fields for Altitude Minimum and Altitude Maximum, each with a value of 0.000.
- Altitude Distance Units:** A dropdown menu.
- Data Set G-Polygon(s) (optional):** An input field with three buttons: Add Polygon, Edit Polygon, and Delete Polygon.

Best Practices



Spatial Domain:

- Description of Geographic Extent is a short, textual description of the geographic domain of the dataset. (Example: “Manistee Watershed” or “Ponds and Reservoirs larger than 2 acres in Jefferson County, CO”)
- Bounding Coordinates are the limits of coverage of the dataset expressed in longitude and latitude.

Keywords

Best Practices

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

Theme

Add Edit Delete

Place

Add Edit Delete

Stratum

Add Edit Delete

Temporal

Add Edit Delete



Keywords — Theme

Thesaurus

Keywords

OK Cancel

Taxonomy

Best Practices



File Options Help

Basic Info Spatial Domain Keywords Taxonomy Access Analytical Tools Miscellaneous

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

Keywords/Taxon

Add Edit Delete

Taxonomic Classification

Add Child Edit Delete

Import Export

Taxonomic System

Classification System or Authority

Add Edit Delete

Identification Reference

Add Edit Delete

Identifier (optional)

Add Edit Delete

Taxonomic Procedures

Taxonomic Class

Data fields ITIS

Taxon Rank Name [Dropdown]

Taxon Rank Value [Text Field]

(If rank name = Empire, then value = Biovitae)

Applicable Common Names (optional)

[Text Field]

OK Cancel

Most taxonomic descriptions can be found at:
[Integrated Taxonomic Information System](#)

Access and Use

File Options Help

Basic Info Spatial Domain Keywords Taxonomy Access Analytical Tools Miscellaneous

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

Access Constraints (default is "None")

Use Constraints (default is "None")

Point of Contact (optional)

Edit

Security Information (optional)

Security Classification System

Security Classification

Security Handling Description

Best Practices



Use Constraints: Restrictions and legal prerequisites for using the dataset after it is granted. (disclaimer)

Access Constraints: Restrictions and legal prerequisites for accessing the dataset, including any access constraints applied to ensure the protection of private or intellectual property, and restrictions on obtaining the dataset.

Analytical Tools

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

Analytical Tool

Add Edit Delete

Tools, models, or statistical procedures that the data set is intrinsically bound to an
Examples include reconstructions of phylogenies, population viability analyses, cor
analyses, and inferences on the effects of climate change on forest composition a
Enough information should be included so that a potential data user can easily det
methodology to acquire it.

Best Practices



Analytical Tool

Analytical Tool Description

Tool Contact (optional)

Edit

Tool Citation (optional)

Edit

OK Cancel

Tool Access Information

Online Linkage(s) [URLs]

Tool Access Instructions

Tool Computer and Operating System (optional)

--Tools, models, or statistical procedures that the dataset is intrinsically bound to and are available for use in analyzing the dataset.

Data Quality: Attribute Accuracy

The screenshot shows a software interface with a menu bar at the top containing: Identification, Data Quality, Spatial Data Org, Spatial Reference, Entity & Attribute, Distribution, and Metadata Ref. Below the menu bar is a sub-menu with icons and labels for: Attribute Accuracy, Consistency & Completeness, Positional Accuracy, Lineage, and Cloud Cover. The main content area is titled "Attribute Accuracy Report" and features a large empty text box on the left. To the right of this box is explanatory text: "Explanation of the accuracy of the identification of entities and assignment of attribute values in the data set." and "The report can reference more extensive descriptions in other documents." Below the text box is a section titled "Quantitative Attribute Accuracy Assessment (optional)" with an empty input field and three buttons labeled "Add", "Edit", and "Delete".

Practices



Attribute Accuracy Report: This element should contain an explanation of the accuracy of the identification of the entities, assignment of values in the dataset, and description of the tests used in the development of the dataset.

Example: How sure are you that it IS a pine tree?

Data Quality

File Options Help

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

Attribute Accuracy Consistency & Completeness Positional Accuracy Lineage Cloud Cover

Logical Consistency Report

Explanation of the fidelity of relationships in the data set and tests used.

The report can reference more extensive descriptions in other documents.

If creating a Logical Consistency Report is not logical, use the default "not applicable"

Completeness Report

Information about omissions, selection criteria, generalizations, definitions used, and other rules used to derive the data set.

The report can reference more extensive descriptions in other documents.

Best Practices



Logical Consistency: Explanation of the exactness of relationships in the dataset and tests used. Are any polygons too small, or any lines too close?

Completeness: How complete is your dataset?

Best Practices for Preparing Ecological Data Sets, ESA, August 2010

Lineage

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

Attribute Accuracy | Consistency & Completeness | Positional Accuracy | Lineage | Cloud Cover

Methodology
Information about a single step of field and/or laboratory work.

Add Edit Delete

Source Information
List of sources and a short discussion of the information contributed by each.

Add Edit Delete

Process Step
Information about a single data processing event.
Can describe process applied to an acquired data set or to raw data you collected.

Add Edit Delete

Best Practices



- Lineage is information about the events, parameters, and source data which was used to construct the dataset, and information about the responsible parties. This section contains 3 subsections: Methodology, Source Information, and Process Steps.
- Tip: Think of the lineage element in the metadata as the “how-to” section. If someone was to read the lineages section of the metadata, could they recreate the work?

Source Citation

Source Information

Source Citation
:
Edit

Source Citation Abbreviation

Type of Source Media

Source Scale Denominator

Time Period of Content

Data are from A single date Multiple dates A range of dates

The calendar is: Gregorian Geologic

Time Period(s):

Add Edit Delete

Currentness Reference

OK Cancel

Best Practices



Source Citation:
Reference for a
source dataset.

Spatial Reference

Best Practices



File Options Help

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

Horizontal Coordinate System Definition Vertical Coordinate System Definition

Horizontal Coordinate System Definition: Geographic Planar Local No definition

Geographic

Latitude Resolution

Longitude Resolution

Geographic Coordinate Units

Planar

Add Edit Delete

Local

Local Description

Local Georeference Information

Geodetic Model

Horizontal Datum Name

Semi-major Axis

Ellipsoid Name

Denominator of Flattening Ratio

Coordinate Systems Definitions:

This is the reference frame or system from which linear or angular quantities are measured and assigned to the position the point occupies.

Entities and Attributes

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

⚡ Detailed Descriptions | ⚡ Overview Descriptions | ⓘ Information

Edit elements

Entity Types

Attributes

Domain Values

Select elements

Manipulate attachments

Root

Branch

View Existing Branches

Attach Detach

Clear fields

Valid linkages, root + branch:
Entity type + Attribute
Attribute + Domain Value
Enumerated Domain + Attribute
Range Domain + Attribute

Best Practices



- Think of your dataset in the form of a spreadsheet. The “Entity Type” would be the name of the table
- “Attributes” are the column headers
- “Attribute Domain Values” are the types of information that can be filled in in the column (allowable values)

Entities and Attributes: Defining Attribute Domain Values

Best Practices

- **Enumerated Domain**
 - A finite list of allowable values
 - Categorical or Coded Data
 - Specify the value or code, the definition, and the source of the definition (can be 'dataset developer')
- **Codeset Domain**
 - Published codes
 - Example: genus and species codes found in the PLANTS database (USDA)
- **Range Domain**
 - Measured values – minimum and maximum allowable values and units of measurement, if appropriate.
- **Unrepresentable Domain**
 - Use when none of the other domains apply.
 - Lists, Index Values, Site IDs
 - Explanation of why data are unrepresentable as a domain is mandatory.



Distribution

Best Practices



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

Distributors | Information

Distributors

Add Edit Delete

Standard Order Process

Format Non-digital Digital

Digital Form

Add Edit Delete

Non-digital Form

Fees *Ordering Instructions (optional)*

Turnaround (optional)

Distribution Information

Distributor Ordering Prerequisites & Availability OK/Cancel

Technical Prerequisites (optional)

Describe any technical capabilities the consumer must have to use the data set in the form(s) provided by the distributor.

Available Time Period (optional)

Data will be available A single date Multiple dates A range of dates

Time Period(s):

Add Edit Delete

Metadata Reference

Best Practices

File Options Help

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

Metadata information | Extensions | Constraints & Security

Metadata Standard's Name: FGDC Biological Data Profile of the Content Standard for Digital Geospatial Metadata

Version of Standard: FGDC-STD-001.1-1999

1999 = Biological
1998 = Plain spatial data
(type in your own if needed)

"FGDC Biological" is default
"National Biological" is alternate
"FGDC Content" is for plain spatial data

Metadata Creation/Update Date: Tuesday, October 09, 2007

Metadata Contact:

Metadata Review Date (optional):

Tuesday, October 09, 2007 (date later than Creation/Update)

Metadata Future Review Date (optional):

Tuesday, October 09, 2007 (date later than Metadata Review)



I dig
Metadata!



Thank you!

Best Practices



Questions?
Comments?