Metadata*

Bill Michener – University of New Mexico

*Special Credit to Viv Hutchison (USGS)
What Is Metadata?

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?
Why metadata?: Data entropy

(Michener et al. 1997)
Metadata in the Real World

Metadata is all around...

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Boullosa, Carmen.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title(s)</td>
<td>They're cows, we're pigs / by Carmen Boullosa</td>
</tr>
<tr>
<td>Physical Descr</td>
<td>viii, 180 p ; 22 cm.</td>
</tr>
<tr>
<td>Subject(s)</td>
<td>Pirates Caribbean Area Fiction.</td>
</tr>
<tr>
<td>Format</td>
<td>Fiction</td>
</tr>
</tbody>
</table>
What does a metadata record look like?

North American Breeding Bird Survey (BBS)

Identification Information:
Citation:

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/nbi_metadata/brdpwr0004.txt (text format) and ftp://cameron.cr.usgs.gov/pub/nbi_metadata/brdpwr0004.html (HTML format) and ftp://cameron.cr.usgs.gov/pub/nbi_metadata/brdpwr0004.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 bird populations and trends for more than 400 species of birds. These data are widely used by research 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 in the U.S. and were sprayed on not only killed insects but also killed birds, raising serious concerns over its impact on birds. Unfortunately, no long-term regional or continental population data were available for wildlife biologists to determine trends in bird populations. The Bird Breeding Survey provides information on bird population trends. Robbins and others (1986) provided the first continent-wide bird population survey for songbirds based on BBS data. When viewed at continental or regional scales, these data provide a record of the relative abundance of species that are well sampled by the BBS. In addition, the data provide information on temporal patterns in trends. Populations of permanent resident and short-distance migratory birds are sampled, and long-term data from the BBS are used to study population trends in North America. USA 1993-1994.

Taxonomy:

Family: Pissardiidae
Genus: Pissardiia
Species: Pissardiia australis
Getting Support: The Value of Metadata

Metadata saves time, money, & frustration
Metadata allows data to be understood and reused
Preserves institutional memory and investment in data; written documentation rather than in one person’s brain
Promotes partnerships and “advertises” data collections; easier to share reliable information
Creates efficiency – identify and use existing datasets, avoid duplication of effort
Gives the dataset creator(s) credit
Metadata Critical to Data Sharing

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?

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
Multiple metadata standards exist

Content Standard for Digital Geospatial Metadata (CSDGM)
- Federal Geographic Data Committee (FGDC)
- Emphasis on geospatial data; environmental sciences
- Includes: **Biological Data Profile (BDP)**
  *Emphasis on biological data (and geospatial)*

**Darwin Core**
- Emphasis on museum specimens

**Dublin Core Element Set**
- Emphasis on web resources, publications

**Ecological Metadata Language (EML)**
- Emphasis on ecology

**ISO 19115  Geographic information: Metadata**
- Emphasis on geospatial data and services
Metadata can be accessed through Clearinghouses:

USGS Clearinghouse:
- over 100,000 records
- http://metadata.ornl.gov/clearinghouse

Metacat:
- metadata registries including KNB, UCNRS, OBFS, NCEAS, PISCO, ESA
- http://knb.ecoinformatics.org/software/index.jsp
DataONE: Quickly discover content
“Working Smart” is the Key to Creating Metadata

Create and use metadata standards and templates
Capitalize on capabilities of various software packages to automatically ingest information
Use existing documentation so as to not recreate information
  • e.g., Project proposals and notes
Consider a team approach to create metadata: experts of specific areas contribute that information
Titles are critical: make yours informative at a glance
Be specific when you can: create a ‘data dictionary’
Tools: Metavist (FGDC Standard Metadata)
NOAA Mermaid: An online tool

NDCDC 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, NCDCC 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.

http://www.ncddc.noaa.gov/Metadata/Tools
Morpho: An Ecological Metadata Language editor
Thank you!

Questions? Comments?

"Professor LaVonne had many enemies in the entomological world, detective, but if you examine that data label, you'll find exactly when and where he was—shall we say—'collected.'"