Metadata services

9-1-2013

Sai Deng
University of Central Florida, sai.deng@ucf.edu

Find similar works at: https://stars.library.ucf.edu/ucfscholar

University of Central Florida Libraries http://library.ucf.edu

Part of the Cataloging and Metadata Commons, and the Scholarly Communication Commons

Original Citation


This Other Presentation is brought to you for free and open access by STARS. It has been accepted for inclusion in Faculty Scholarship and Creative Works by an authorized administrator of STARS. For more information, please contact lee.dotson@ucf.edu.
Metadata Services

Sai Deng
University of Central Florida Libraries
What will be covered?

- Metadata Services at UCF
- Understanding metadata (Definitions, types, examples)
- Metadata creation and application
- Metadata standards, related tools and data repositories
- Controlled vocabularies and thesauri
- Linked data: Metadata structure standards
- More metadata and data curation tools
Metadata
- Metadata Services
- Metadata Standards
- Descriptive Metadata
- Structural Metadata
- Administrative Metadata
- Technical Metadata
- Rights Metadata
- Preservation Metadata
- Digital Repositories
- CMS
- Metadata Creation and Application
- METS
- PREMIS
- MODS
- DC
- TEI
- DDI
- DwC
- Dryad
- GenBank
- DIF
- FGDC/CSDGM
- ISO 19115
- Metadata Mapping
- Metadata Harvesting
- Controlled Vocabulary
- Thesauri
- Linked data
- Metadata Tools

*Word cloud generated using Tagxedo.*
Metadata Services

- Metadata Services at UCF Libraries provide consultation on metadata related issues and metadata production assistance for UCF researchers in their research lifecycle.

- **Metadata Creation and Application.** It assists in:
  - Implementing metadata standards, documenting research data and creating and enhancing project metadata
  - Applying and developing controlled vocabularies and thesauri
  - Preparing metadata for deposit and preservation

- **Metadata Consultation on:**
  - Metadata standards implementation
  - Metadata schemas customization
  - Persistent identifier management
  - Metadata work to improve search and retrieval of data
  - Data curation tools and practices
  - Data repositories information
  - Metadata literature and best practices
Metadata Services

- Metadata Guide (http://guides.ucf.edu/metadata)
  - Metadata Introduction
  - Technical, Rights and Preservation Metadata
  - Research Data & Dataset
  - Data Documentation & Data Formats
  - Data Management & Metadata
  - Common Metadata Standards
  - Domain Metadata Standards
  - Basic Metadata Fields
  - Metadata Harvesting
  - Metadata Generation
  - Metadata Management
  - Data Curation
  - Digital Repositories
  - Data Repositories

- Metadata Services page: http://library.ucf.edu/ScholarlyCommunication/Metadata.php
Understanding Metadata

- **Metadata definition:**
  - Meta: Greek prefix. Means "after, behind or beyond."
  - Data: Latin word. Factual information used for calculating, reasoning or measuring.
  - Metadata means something behind or beyond data itself, and it includes data about its content, containers and contextual information.

- A formal definition: Metadata is data about data, data associated with an object, a document, or a dataset for purposes of description, administration, technical functionality and preservation.

Understanding Metadata: Types

- **Descriptive metadata**
  - Serves the purposes of discovery, identification, and selection

- **Structural metadata**
  - Describes the physical and/or logical structure of resources

- **Administrative metadata** (the use, management, and encoding processes of digital objects)
  - **Technical Metadata** (technical attributes of digital objects or datasets)
  - **Rights Metadata** (intellectual property rights, creative commons...)
  - **Preservation Metadata** (PREMIS Data Dictionary for Preservation Metadata)
Looking at Different Types of Metadata by Examining a METS File

- Metadata Encoding and Transmission Standard (METS)
  - Wrapper

- What’s wrapped in:
  - **METS Header**: `<metsHdr>`

- **Descriptive Metadata Section**: `<dmdSec>`
  - Descriptive Metadata Section `<mdWrap>`
    - Metadata Object Description Schema (MODS)

- **Administrative Metadata Section**: `< amdSec>`
  - Technical Metadata
  - Rights Metadata
  - Preservation Metadata

- **File Section**: `<fileSec>`
- **Structural Map**: `<structMap>`
Looking at Different Types of Metadata by Examining a METS file: The Record Display

<table>
<thead>
<tr>
<th>Permanent Link:</th>
<th><a href="http://ufdc.ufl.edu/CA01200034/00001">http://ufdc.ufl.edu/CA01200034/00001</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Material Information</td>
<td></td>
</tr>
<tr>
<td>Title:</td>
<td>The history, civil and commercial, of the British colonies in the West Indies.</td>
</tr>
<tr>
<td>Series Title:</td>
<td>West Indies collection</td>
</tr>
<tr>
<td>Physical Description:</td>
<td>2 v. fold. map, tables. 22cm.</td>
</tr>
<tr>
<td>Language:</td>
<td>English</td>
</tr>
<tr>
<td>Donor:</td>
<td>The William L. Bryant Foundation. (endowment)</td>
</tr>
<tr>
<td>Publisher:</td>
<td>Luke While.</td>
</tr>
<tr>
<td>Place of Publication:</td>
<td>Dublin</td>
</tr>
<tr>
<td>Publication Date:</td>
<td>1793</td>
</tr>
<tr>
<td>Subjects</td>
<td></td>
</tr>
<tr>
<td>Subjects / Keywords:</td>
<td>Indians of the West Indies West Indies, British--Description and travel West Indies, British--History Slave trade West Indies--Economic conditions Agriculture--West Indies, British. (inclusive) Botany--Jamaica Slavery--West Indies Description and travel--Jamaica History--Jamaica</td>
</tr>
<tr>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>Summary:</td>
<td>&quot;Hortus castaneus, or a catalogue of exotic plants Esquire, in the mountains of Liguanea, in the isla de cease, by Arthur Broughton.&quot;—Cf. v. 1, p. 455-4</td>
</tr>
<tr>
<td>Donation:</td>
<td>The William L. Bryant Foundation</td>
</tr>
</tbody>
</table>

Description/metadata view:
http://www.dloc.com/CA01200034/00001/citation?search=university+%3dcentral+%3dflorida

*The object/image/page view is at:
http://ufdc.ufl.edu/CA01200034/00001
Looking at Different Types of Metadata by Examining a METS File:
The “Uncollapsed” METS File

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<!---- The history, civil and commercial, of the British colonies in the West Indies. ( Book )
-->
  +<?METS:metsHdr>
  +<?METS:dmdSec ID="DMD1"/>
  +<?METS:dmdSec ID="DMD2"/>
  +<?METS:amdSec/>
  +<?METS:fileSec/>
  +<?METS:structMap ID="STRUCT1" TYPE="physical"/>
</METS:mets>
```

http://ufdcimages.uflib.ufl.edu/CA/01/20/00/34/00001/CA01200034_00001.mets.xml
Looking at Different Types of Metadata by Examining a METS File: METS Header: `<metsHdr>`

```xml
< Mets:metsHdr CREATEDATE="2009-03-02T20:18:55Z" ID="CA01200034_00001" LASTMODDATE="2012-05-02T12:43:00Z" RECORDSTATUS="PARTIAL">
  <Mets:agent ROLE="CREATOR" TYPE="ORGANIZATION">
    <Mets:name>UCFDLOC,University of Central Florida Libraries</Mets:name>
  </Mets:agent>
  <Mets:agent OTHERTYPE="SOFTWARE" ROLE="CREATOR" TYPE="OTHER">
    <Mets:name>dLOC Toolkit</Mets:name>
  </Mets:agent>
  <Mets:agent ROLE="CREATOR" TYPE="INDIVIDUAL">
    <Mets:name>LailaMiletic</Mets:name>
    <Mets:note>Online edit by Laila Miletic-Veizovic (7/26/2010)</Mets:note>
    <Mets:note>Online edit by Laila Miletic-Vejzovic (5/2/2012)</Mets:note>
  </Mets:agent>
</Mets:metsHdr>
```

What’s covered by METS Header: Agent information (e.g. Creator, software, editor)
Looking at Different Types of Metadata by Examining a METS file: Descriptive Metadata Section <dmdSec>

<dmdSec> (include either <mdWrap> or <mdRef>)

- <METS:dmdSec ID="DMD1">
  + <METS:mdWrap MDTYPE="MODS" MIME TYPE="text/xml" LABEL="MODS Metadata"></METS:mdWrap>
</METS:dmdSec>
- <METS:dmdSec ID="DMD2">
  + <METS:mdWrap MDTYPE="OTHER" OTHER MDTYPE="SobekCM" MIME TYPE="text/xml" LABEL="SobekCM Custom Metadata"></METS:mdWrap>
</METS:dmdSec>

What’s covered by Descriptive Metadata Section: Embedded descriptive metadata ("mdWrap"); (can point to external metadata too, "mdRef")
Looking at Different Types of Metadata by Examining a METS File:
Descriptive Metadata Section <dmdSec> - <mdWrap>

What’s covered by embedded MODS Descriptive Metadata:
- Summary
- Rights
- Identifier
- Language
- Physical Location
- Primary Author
- Note
- Publisher
- Date Issued
- Related Item
- Subject
- Title
- Type of resource
Looking at Different Types of Metadata by Examining a METS File: 
Administrative Metadata Section: `<amdSec>` - `<techMD>`

- `<METS:amdSec>`
  - `<METS:techMD` ID="TECHMD1">`
    - `<METS:mdWrap` MTYPE="OTHER" OTHERMDTYPE="SobekCM" MIME TYPE="text/xml" LABEL="SobekCM File Technical Details">`
      - `<METS:xmlData>`
        - `<sobekcm:FileInfo>`
          - `<sobekcm:File` fileid="JPEG1" width="630" height="971"/>
          - `<sobekcm:File` fileid="JPEG2" width="582" height="1000"/>
          - `<sobekcm:File` fileid="JPEG3" width="574" height="1000"/>
          - `<sobekcm:File` fileid="JPEG4" width="630" height="385"/>
          - `<sobekcm:File` fileid="JPEG5" width="613" height="1000"/>
          ...
        </sobekcm:FileInfo>
      </METS:xmlData>
    </METS:mdWrap>
  </METS:techMD`
</METS:amdSec>`

What’s covered by 
Administrative Metadata Section: 
Technical metadata

- `<amdSec>` include `<techMD>`, `<rightsMD>`, `<sourceMD>` and `<digiprovMD>`;
- In this example, rights, source and preservation sections do not exist in `<amdSec>`; however, some of these information are included in the descriptive metadata;
- `<techMD>` may include `<mdWrap>` or `<mdRef>`.
Looking at Different Types of Metadata by Examining a METS File:  
File Section: `<fileSec>`

```
- `<METS:fileSec>`
  - `<METS:fileGrp USE="reference">`
    - `<METS:file GROUPID="G1" ID="JPEG1" MIMETYPE="image/jpeg" SIZE="89199">`
      `<METS:FLocat LOCTYPE="OTHER" OTHERLOCTYPE="SYSTEM" xlink:href="001_FrontCover.jpg"/>`
    `</METS:file>`
  + `<METS:file GROUPID="G2" ID="JPEG2" MIMETYPE="image/jpeg" SIZE="88823">`</METS:file>
  + `<METS:file GROUPID="G3" ID="JPEG3" MIMETYPE="image/jpeg" SIZE="92908">`</METS:file>
...
- `<METS:file GROUPID="G526" ID="JP2526" MIMETYPE="image/jp2" SIZE="841173">`
  `<METS:FLocat LOCTYPE="OTHER" OTHERLOCTYPE="SYSTEM" xlink:href="526_BackCover.jp2"/>`
`</METS:file>`
`</METS:fileGrp>
`</METS:fileSec>`
```

*File section lists all files of the digital object*
Looking at Different Types of Metadata by Examining a METS File:
Structural Map: <structMap>

```xml
<METS:structMap ID="STRUCT1" TYPE="physical">
  <METS:div DMDID="DMD1 DMD2" LABEL="The history, civil and commercial, of the British colonies in the West Indies." ORDER="0" TYPE="main">
    <METS:div ID="PDIV1" ORDER="1" TYPE="Front Cover">
      <METS:div ID="PAGE1" LABEL="Front Cover 1" ORDER="1" TYPE="Page">
        <METS:fptr FILEID="JPEG1"/>
        <METS:fptr FILEID="JP21"/>
      </METS:div>
    </METS:div>
    + <METS:div ID="PAGE2" LABEL="Front Cover 2" ORDER="2" TYPE="Page"/>
  </METS:div>
  ...
  <METS:div ID="PAGE526" LABEL="Page 498" ORDER="2" TYPE="Page">
    <METS:fptr FILEID="JPEG526"/>
    <METS:fptr FILEID="JP2526"/>
  </METS:div>
</METS:structMap>
```

Structural map outlines the hierarchical structure for the digital library object.
The library has been creating cataloging and metadata records to serve the university faculty and students for teaching and learning (e.g. traditionally, using MARC for ILS, DC, MODS and METS for digital collections and DACS, DC and EAD for archives);

Will continue and expand the metadata services by:
- Working with Digital Initiatives, Special Collections and other library units to provide access to the scholarly output of the university;
- Helping in enriching, editing and enhancing metadata and documentation for researchers’ projects and their datasets if needed;
- Assisting in controlled vocabulary and authority work and recommend authorized terms, subject headings, descriptors and/or taxonomies for research project documentations, datasets and publications archiving;
- Helping with data preparation for deposit, metadata transformation and transfer between different repositories and systems if needed.
CMS Available @ UCF and Their Supported Standards

- **The UCF Libraries’ CONTENTdm**
  - Currently stores images, manuscripts, university archival materials and other digital collections.
  - Supports qualified Dublin Core, Visual Resources Association (VRA) Core element and Encoded Archival Description (EAD) standards, and also accepts locally defined elements.

- **The State University Libraries of Florida’s DigiTool**
  - Hosts local history and culture heritage materials at the library archives and special collections.
  - Supports Dublin Core, Metadata Object Description Schema (MODS), MarcXML and Metadata Encoding and Transmission Standard (METS).

- **SobekCM**
  - Stores Special Collection materials such as the Digital Library of Caribbean.
  - Supports Dublin Core, MODS, MarcXML and METS.

- **In Common**: All support Dublin Core, which allows greater interoperability. Subject specific and local information fields can be mapped to certain fields in a schema that the system supports, if applicable.

- If prefer domain repository, the library can assist in identifying data repositories to store the research data.
Selecting Metadata Standards

- In selecting a metadata standard, considering:
  - UKOLN criteria: Purpose of metadata, attributes of resource, design of standard, granularity, interoperability, support, growth, extensibility, reputation, ease of use, existing experience; ([http://www.ukoln.ac.uk/qa-focus/documents/briefings/briefing-63/html/](http://www.ukoln.ac.uk/qa-focus/documents/briefings/briefing-63/html/))
  - Your unique project needs;
  - Funding agency requirements*;
  - Standard discipline practices.

- *Funding agency requirements are covered by the “Data Management Plan” and the “Dataset Metadata” sessions.
Common Metadata Standards

- **Dublin Core (DC):** A general metadata standard for describing a wide range of digital resources.
  - 15 Elements: Title, Creator, Subject or keyword, Description, Publisher, Type, Format, Identifier, Source, Language, Relation, Coverage, Rights

- **Tools**
    (for Conversion, Crosswalks, Metadata Creation, Extraction, Generation and Harvesting and Metadata Quality)
<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Ballou's pictorial</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Earlier Title</strong></td>
<td>Gleason's pictorial</td>
</tr>
</tbody>
</table>
| **Subject** | American literature -- Periodicals  
American literature -- 19th century -- Periodicals  
Louisiana -- History -- 18th century  
Louisiana -- History -- 19th century  
Novel -- 19th century. |
| **Description** | Ballou's Pictorial, Boston, Saturday, December 22, 1855... |
| **Author** | Ballou, Maturin Murray, 1820-1895. |
| **Publisher** | M. M. Ballou |
| **Contributors** | Austin C. Burdick. |
| **Date** | 1855 |
| **Type** | Newspapers |
| **Identifier** | AP2.A389 No. 233 |
| **Language** | English |
| **Relation** | http://www.american-south.org |
| **Rights** | All rights to images are held by the respective holding institution... |
| **Tag** | DP0012649 |
| **Physical description** | Loose leaf from Vol. IX., No. 25 (Whole No. 233). Illustration included. |
| **Number of pages** | 2 pages |
| **Size** | 40 cm. |
| **Repository** | University of Central Florida Libraries, Special Collections & University Archives |
| **Repository Collection** | African American Legacy, the Carol Mundy Collection  
Carol Mundy Library |
| **Place** | Boston, Mass. |
| **Coverage-Temporal** | 1691-1850 |
| **Digital Publisher** | University of Central Florida Libraries |
| **Digital Reproduction Specifications** | Jpeg and Jpeg2000 images were derived from 400 dpi tiffs scanned on an Avision FB 6080E Book-edge Scanner. |
| **Alternative Title** | Ballou's pictorial drawing-room companion |
| **Uniform Title** | Civil War Collection  
Carol Mundy Library |
| **Provenance** | Acquired by Special Collections & University Archives, University of Central Florida Libraries from Carol Mundy, the collector of African-Americana. |
| **Later Title** | Welcome guest |

<table>
<thead>
<tr>
<th>Field name</th>
<th>DC map</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Title</td>
<td>Title</td>
</tr>
<tr>
<td>2 Earlier Title</td>
<td>Title-Alternative</td>
</tr>
<tr>
<td>3 Subject</td>
<td>Subject</td>
</tr>
<tr>
<td>4 Description</td>
<td>Description</td>
</tr>
<tr>
<td>5 Table of contents</td>
<td>Description-Table Of Contents</td>
</tr>
<tr>
<td>6 Author</td>
<td>Creator</td>
</tr>
<tr>
<td>7 Photographer</td>
<td>Creator</td>
</tr>
<tr>
<td>8 Publisher</td>
<td>Publisher</td>
</tr>
<tr>
<td>9 Studio</td>
<td>Publisher</td>
</tr>
<tr>
<td>10 Contributors</td>
<td>Contributors</td>
</tr>
<tr>
<td>11 Date</td>
<td>Date-Issued</td>
</tr>
<tr>
<td>12 Type</td>
<td>Type</td>
</tr>
<tr>
<td>13 Identifier</td>
<td>Identifier</td>
</tr>
<tr>
<td>14 Language</td>
<td>Language</td>
</tr>
<tr>
<td>15 Relation</td>
<td>Relation</td>
</tr>
<tr>
<td>16 Coverage</td>
<td>Coverage</td>
</tr>
<tr>
<td>17 Rights</td>
<td>Rights</td>
</tr>
<tr>
<td>18 Audience</td>
<td>Audience</td>
</tr>
<tr>
<td>19 Tag</td>
<td>Identifier</td>
</tr>
<tr>
<td>20 Physical description</td>
<td>Description</td>
</tr>
<tr>
<td>21 Number of pages</td>
<td>Format-Extent</td>
</tr>
<tr>
<td>22 Size</td>
<td>Format-Extent</td>
</tr>
<tr>
<td>23 Keywords</td>
<td>Subject</td>
</tr>
<tr>
<td>24 Repository</td>
<td>Source</td>
</tr>
<tr>
<td>25 Repository Collection</td>
<td>Source</td>
</tr>
<tr>
<td>26 Place</td>
<td>Coverage-Spatial</td>
</tr>
<tr>
<td>27 Coverage-Temporal</td>
<td>Coverage-Temporal</td>
</tr>
<tr>
<td>28 Digital Publisher</td>
<td>None</td>
</tr>
<tr>
<td>29 Funding Source</td>
<td>None</td>
</tr>
<tr>
<td>30 Digital Reproduction Specifications</td>
<td>Format</td>
</tr>
<tr>
<td>31 Transcript</td>
<td>Description</td>
</tr>
<tr>
<td>32 Digital Date</td>
<td>Date-Available</td>
</tr>
<tr>
<td>33 Alternative Title</td>
<td>Title-Alternative</td>
</tr>
<tr>
<td>34 Uniform Title</td>
<td>Title-Alternative</td>
</tr>
<tr>
<td>35 Provenance</td>
<td>Description</td>
</tr>
<tr>
<td>36 Later Title</td>
<td>Title-Alternative</td>
</tr>
<tr>
<td>37 Filename</td>
<td>None</td>
</tr>
</tbody>
</table>
Ballou's pictorial Gleason's pictorial Ballou's pictorial drawing-room companion Civil War Collection; Carol Mundy Library Welcome guest
Ballou, Maturin Murray, 1820-1895. <dc:creator>
American literature -- Periodicals ; American literature -- 19th century -- Periodicals ; Louisiana -- History -- 18th century ; Louisiana -- History
<dc:subject>
Ballou's Pictorial, Boston, Saturday, December 22, 1855. Loose leaf from Vol. IX., No. 25 (Whole No. 233). The back of the leaf shows page m emblematical picture of the great State of Louisiana and a novel written by Austin C. Burdick "The Visconti, or, Barbarigo the Stranger, a tale of a leaf from Vol. IX., No. 25 (Whole No. 233). Illustration included. Acquired by Special Collections & University Archives, University of Central collector of African-American.
<dc:description>
M. M. Ballou <dc:publisher>
Austin C. Burdick. <dc:contributor>
1855 <dc:date>
Newspapers <dc:type>
2 pages 40 cm. Jpeg and Jpeg2000 images were derived from 400 dpi tiffs scanned on an Avision FB 6080E Book-edge Scanner.
<dc:format>
AP2. A389 No. 233 DP0012649 <dc:identifier>
University of Central Florida Libraries, Special Collections & University Archives African American Legacy, the Carol Mundy Collection; Carol Mundy Library Welcome guest
<dc:source>
English <dc:language>
http://www.american-south.org <dc:relation>
Boston, Mass. 1691-1850 <dc:coverage>
All rights to images are held by the respective holding institution. This image is posted publicly for non-profit educational uses, excluding print images and/or for copyright information contact Special Collections & University Archives, University of Central Florida Libraries, Orlando, Florida. speccoll@mail.ucf.edu <dc:rights>
Common Metadata Standards (Cont.)

- **Encoded Archival Description (EAD)**
  - A standard for encoding archival finding aids with XML.

- **Government Information Locator Service (GILS)**
  - The Global Information Locator Service defines a core element set for government information so that it can be more searchable and discoverable by the general public.

- **ONIX for Books (ONline Information eXchange)**
  - An international standard for representing and communicating book industry product information in XML format.
Categories for the Description of Works of Art (CDWA)
A conceptual framework and guidelines for the description of art objects and images.

visual Resources Association Core Categories (VRA Core)
A data standard for the description of works of visual culture as well as the images that document them.

NISO Metadata for Digital Images
This technical metadata standard defines a set of metadata elements for raster digital images to enable users to develop, exchange, and interpret digital image files. The dictionary has been designed to facilitate interoperability between systems, services, and software as well as to support the long-term management of and continuing access to digital image collections.

PBCore
The metadata standard for audiovisual media developed by the public broadcasting community.

Technical Metadata for Multimedia: MPEG-7
The Multimedia Content Description Interface MPEG-7 is an ISO/IEC standard and specifies a set of descriptors to describe various types of multimedia information and is developed by the Moving Picture Experts Group.
Minuet of the Bride


Michel Erhart (German, ca. 1440 - after 1522)

88.5 x 40 cm (34 7/8 x 15 3/4 inches)

Oil and gold leaf on panel

2nd state 7/50

Baroque

Netherlandish

Examples from CDWA Lite: Specification 1.1
Disciplinary Metadata Standards, Related Tools and Repositories: Social Sciences & Humanities

- **DDI - Data Documentation Initiative**
  - A metadata specification for the social and behavioral sciences. Expressed in XML, the DDI metadata specification supports the entire research data life cycle.

- **DDI Tools**: The DDI site's list of tools to implement this standard.

- **Projects using DDI** ([http://www.ddialliance.org/ddi-at-work/projects](http://www.ddialliance.org/ddi-at-work/projects))

- **Social science repositories and archives**
  - **ICPSR - Inter-university Consortium for Political and Social Research**
    - A data archive providing leadership and training in data access, curation, and methods of analysis for the social science research community; all metadata conforms to the DDI standard.
  - **The Institution for Social and Policy Studies (ISPS) Data Archive**
  - **UKDA - UK Data Archive**
Social Science Dataset in DDI Format

http://www.icpsr.umich.edu/icpsrweb/ICPSR/ddi2/studies/20363
The Study Description consists of information about the data collection, study, or compilation that the DDI-compliant documentation file describes. This section includes information about how the study should be cited, who collected or compiled the data, who distributes the data, keywords about the content of the data, summary (abstract) of the content of the data, data collection methods and processing, etc.
Social Science Dataset in DDI Format
(Cont.)

Included Fields:
fileDscr
fileTxt
fileName

- `<fileDscr ID="F1">`
- `<fileTxt ID="Part1">`
  `<fileName>Female Interviews</fileName>`
  `<fileTxt>`
  `<fileDscr>`
+ `<fileDscr ID="F2"></fileDscr>`

fileDscr
(Data Files Description)

Information about the data file(s) that comprises a collection. This section can be repeated for collections with multiple files.

http://www.icpsr.umich.edu/icpsrweb/ICPSR/ddi2/studies/20363
Web Display of the ICPSR Dataset

ICPSR: Interuniversity Consortium for Political and Social Research.

Field Labels:
- Title
- Principal investigator(s)
- Summary
- Access notes
- Dataset(s)
- Study description

Citation

Funding

Scope of study

- Subject terms
- Smallest geographic unit
- Geographic coverage
- Time period
- Date of collection
- Unit of observation
- Universe
- Data types
- Data collection notes

Methodology

- Study purpose
- Study design
Disciplinary Metadata Standards, Related Tools and Repositories: Social Sciences & Humanities

○ **Text Encoding Initiative (TEI):** A standard for the representation of texts in digital form, chiefly in the humanities, social sciences and linguistics.

○ **TEI Tools**
  - Generation of schemas and documentation: Roma;
  - Stylesheets for converting TEI documents to various formats;
  - Conversion to and from TEI: OxGarage;
  - Internationalization of the TEI.

○ **Humanities repositories and Projects:**
  - [Projects Using the TEI](http://www.tei-c.org) (from the official TEI website)
  - [History Data Service (HDS)](http://www.historydata.org) (information from dataBib)
  - [Michigan Corpus of Academic Spoken English (MICASE)](http://www.micase.org)
  - [Association of Religion Data Archives (ARDA)](http://www.arda.com)

```xml
<TEI>
  <teiHeader>
    <!-- ... metadata describing the manuscript -->
  </teiHeader>
  <facsimile>
    <!-- ... metadata describing the digital images -->
  </facsimile>
  <text>
    <!-- (optional) transcription of the manuscript -->
  </text>
</TEI>
```

The minimal required structure for `teiHeader`:

```xml
<TEI>
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>[Title of manuscript]</title>
      </titleStmt>
      <publicationStmt>
        <distributor>[name of data provider]</distributor>
        <idno>[project-specific identifier]</idno>
      </publicationStmt>
      <sourceDesc>
        <msDesc xml:id="ex5" xml:lang="en">
          <!-- [full manuscript description] -->
        </msDesc>
      </sourceDesc>
      <revisionDesc>
        <change when="2008-01-01">
          <!-- [revision information] -->
        </change>
      </revisionDesc>
    </fileDesc>
  </teiHeader>
</TEI>
```
msDesc provides detailed information about a single manuscript or other text-bearing object.

Fields:
- msDesc
  - msIdentifier
    - Settlement
    - repository
    - Idno
    - altIdentifier
  - msContents
    - quote
    - title
  - physDesc
    - material
  - History
    - origPlace
    - origDate
    - quote

For more element descriptions and examples, see [http://projects.oucs.ox.ac.uk/ENRICH/Deliverables/referenceManual_en.html](http://projects.oucs.ox.ac.uk/ENRICH/Deliverables/referenceManual_en.html)
Disciplinary Metadata Standards, Related Tools and Repositories: Biological Sciences - Standards

**ABCD - Access to Biological Collection Data**
A standard for the access to and exchange of data about specimens and observations (a.k.a. primary biodiversity data), including specimens and observations.

**Darwin Core**
A metadata specification for information about the geographic occurrence of species and the existence of specimens in collections.

**EML: Ecological Metadata Language**
A metadata specification developed by the ecology discipline and for the ecology discipline. EML is implemented as a series of XML document types that can be used in a modular and extensible manner to document ecological data.
Disciplinary Metadata Standards, Related Tools and Repositories: Biological Sciences- Tools

- **Darwin Core Archive Assistant**
  - A web application that offers data publishers wishing to serve to the **Global Biodiversity Information Facility (GBIF) network** an easy interface for describing data elements as text (CSV) files, with a XML descriptor file to accompany them.

- **Morpho**
  - An application for accessing and manipulating metadata and data (both locally and on the network), with wizards creating metadata files using a subset of **Ecological Metadata Language (EML)**.
Disciplinary Metadata Standards, Related Tools and Repositories: Biological Sciences - Repositories

**Dryad**
(https://datadryad.org/)
A digital repository for data underlying the international scientific publications, with an initial focus on evolutionary biology and related fields.

**GBIF - Global Biodiversity Information Facility**
GBIF is a free and open access global web portal promoting and facilitating the mobilization, access, discovery and use of biodiversity data.
The preferred format for publishing data to the GBIF network is the Darwin Core Archive (DwC-A), and its Integrated Publishing Toolkit uses EML as its data standard.

**GenBank**
GenBank ® is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences.
<table>
<thead>
<tr>
<th>dc.contributor.author</th>
<th>Crawford, Nicholas G.</th>
</tr>
</thead>
<tbody>
<tr>
<td>dc.contributor.author</td>
<td>Faircloth, Brant C.</td>
</tr>
<tr>
<td>dc.contributor.author</td>
<td>McCormack, John E.</td>
</tr>
<tr>
<td>dc.contributor.author</td>
<td>Brumfield, Robb T.</td>
</tr>
<tr>
<td>dc.contributor.author</td>
<td>Winker, Kevin</td>
</tr>
<tr>
<td>dc.contributor.author</td>
<td>Glenn, Travis C.</td>
</tr>
<tr>
<td>dc.date.accessioned</td>
<td>2012-05-18T15:48:08Z</td>
</tr>
<tr>
<td>dc.date.available</td>
<td>2012-05-18T15:48:08Z</td>
</tr>
<tr>
<td>dc.date.issued</td>
<td>2012-05-16</td>
</tr>
<tr>
<td>dc.identifier</td>
<td>doi:10.5061/dryad.75nv22qj</td>
</tr>
<tr>
<td>dc.identifier.uri</td>
<td><a href="http://hdl.handle.net/10255/dryad.38214">http://hdl.handle.net/10255/dryad.38214</a></td>
</tr>
<tr>
<td>dc.description</td>
<td>We present the first genomic-scale analysis addressing the phylogenetic position of turtles, using over 1,000 loci from representatives of all major reptile lineages including tuatara…</td>
</tr>
</tbody>
</table>

Dryad
(https://datadryad.org/)

This is an example of full metadata view.

Title: Data from: More than 1000 ultraconserved elements provide evidence that turtles are the sister group of archosaurs

http://www.datadryad.org/handle/10255/dryad.38214?show=full
Title: Data from: More than 1000 ultraconserved elements provide evidence that turtles are the sister group of archosaurs
### DSpace: A sample herbarium record from “Virtual Herbarium Collection” in Shocker Open Access Repository (SOAR)

**Files in this item**

Name: HERBARIUM_58.jpg  
Size: 1010.Kb  
Format: image/jpeg  
[View/Open](http://soar.wichita.edu/handle/10057/3031?show=full)

This herbarium record in DSpace uses a combination of DC and DwC for metadata description.

**Achillea lanulosa Nutt. [Taxon]**  
[http://soar.wichita.edu/handle/10057/3031?show=full](http://soar.wichita.edu/handle/10057/3031?show=full)
Biotechnology Dataset in GenBank

Saccharomyces cerevisiae TCP1-beta gene, partial cds; and Axl2p (AXL2) and Rev7p (REV7) genes, complete cds


<table>
<thead>
<tr>
<th>LOCUS</th>
<th>DEFINITION</th>
<th>ACCESSION</th>
<th>REFERENCES</th>
<th>SOURCE</th>
<th>KEYWORD</th>
<th>FEATURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCP1</td>
<td>DNA linear PLM</td>
<td>U49595.1</td>
<td>1(2004)</td>
<td>Saccharomyces cerevisiae (baker's yeast)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Disciplinary Metadata Standards, Related Tools and Repositories:
Earth Science - Standards

**AgMES**
Agricultural Metadata Element Set
AgMES is designed to include agriculture specific extensions for terms and refinements from established metadata standard such as Dublin Core and AGLS to facilitate resource discovery, interoperability and data exchange in the agriculture domain.

**ISO 19115:2003**
An internationally-adopted schema for describing geographic information and services. It provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data.

**DIF**
Directory Interchange Format
An early metadata initiative from the Earth sciences community, intended for the description of scientific data sets. It includes elements focusing on instruments that capture data, temporal and spatial characteristics of the data, and projects with which the dataset is associated.

**FGDC/CSDGM**
Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata
Content standard for digital geospatial metadata maintained by the Federal Geographic Data Committee (FGDC). Often referred to as the “FGDC Metadata Standard.”

**CF**
(Climate and Forecast) Metadata Conventions
A standard for climate and forecast “use metadata” that aims both to distinguish quantities (such as physical description, units, or prior processing) and to locate the data in space-time.
Geospatial Dataset in FGDC/CSDGM Format

Top level elements:
- **idinfo**: Identification Information;
- **dataqual**: Data Quality Information;
- **spdoinfo**: Spatial Data Organization Information;
- **spref**: Spatial Reference Information;
- **eainfo**: Entity and Attribute Information;
- **distinfo**: Distribution Information;
- **metainfo**: Metadata Reference Information;

CSDGM Fields (under idinfo):
- **Idinfo**
- **Citation**
  - **citeinfo**
  - **Origin**
  - **Pubdate**
  - **Title**
  - **Pubinfo**
  - **Onlink**
  - **Descript**
  - **Abstract**
  - **Purpose**
  - **Supplinfo**
  - **Timeperd**
  - **Status**
  - **Spdom**
  - **Keywords**
  - **Accconst**
  - **Useconst**
  - **Ptcontac**
  - **Native**
  - **Crossref**

XML view

FGDC/CSDGM:
Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata

Web view:

http://catalog.data.gov/harvest/object/df1e4a33f-0fd2-4135-ae33-a1c526bd7a73/original

Biological data of field activity 08CRD01 (B-1-08-VI) in U.S. Virgin Islands from 05/30/2008 to 06/13/2008

http://www.geoplatform.gov/node/243/bf5a5c64-085e-4c68-a489-93e8608d3ad1
Measurements Of Pollution In The Troposphere (MOPITT) Level 3 Gridded Daily CO Retrievals (MOP03) V004

http://gcmd.gsfc.nasa.gov/KeywordSearch/Metadata.do?Portal=langley&KeywordPath=Parameters%7CATMOSPHERE%7CAIR+QUALITY%7CCARBON+MONOXIDE&OrigMetadataNode=GCMD&EntryId=MOP034&MetadataView=Full&MetadataType=0&lbnode=mdlb1

Labels:
Summary
Related URL
Geographic Coverage
Spatial coordinates
Temporal Coverage

Measurements Of Pollution In The Troposphere (MOPITT) Level 3 Gridded Daily CO Retrievals (MOP03) V004

http://gcmd.gsfc.nasa.gov/KeywordSearch/Metadata.do?Portal=langley&KeywordPath=Parameters%7CATMOSPHERE%7CAIR+QUALITY%7CCARBON+MONOXIDE&OrigMetadataNode=GCMD&EntryId=MOP034&MetadataView=Full&MetadataType=0&lbnode=mdlb1
Earth Science Dataset in DIF Format (web display)(Cont.)

Directory Interchange Format (DIF)


Origin: DIF was the product of an Earth Science and Applications Data Systems Workshop (ESADS) held February 24-26, 1987 on catalog interoperability (CI). (http://gcmd.gsfc.nasa.gov/add/difguide/whatisadif.html)

Labels:
Location Keywords
Science Keywords
ISO Topic category
Platform
Instrument
Project
Ancillary Keywords
Data Set Progress
Data Center
Personnel
Extended Metadata Properties
Creation and Review Dates
...

Measurements Of Pollution In The Troposphere (MOPITT) Level 3 Gridded Daily CO Retrievals (MOP03) V004
http://gcmd.gsfc.nasa.gov/KeywordSearch/Metadata.do?Portal=langley&KeywordPath=Parameters%7CATMOSPHERE%7CAIR+QUALITY%7CCARBON+MONOXIDE&OrigMetadataNode=GCMD&EntryId=MOP034&MetadataView=Full&MetadataType=0&lbnode=mdlb1
Earth Science Dataset in ISO 19115 Format

ISO 19115:2003
An internationally-adopted schema for describing geographic information and services.

It provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data.

NASA: Metadata for dataset 'Measurements Of Pollution In The Troposphere (MOPITT) Level 3 Gridded Daily CO Retrievals (MOP03) V004'
http://gcmd.gsfc.nasa.gov/KeywordSearch/MetadataTransformation.do?Portal=langley&KeywordPath=Parameters|ATMOSPHERE|AIR+QUALITY|CARBON+MONOXIDE&EntryId=MOP034&MetadataType=0&to=iso-csw&lbnode=mdlb1
ISO 19115:2003 provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data.

(\url{http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020})

NASA: Metadata for dataset 'Measurements Of Pollution In The Troposphere (MOPITT) Level 3 Gridded Daily CO Retrievals (MOP03) V004'

\url{http://gcmd.gsfc.nasa.gov/KeywordSearch/MetadataTransformation.do?Portal=langley&KeywordPath=Parameters|ATMOSPHERE|AIR+QUALITY|CARBON+MONOXIDE&EntryId=MOP034&MetadataType=0&to=iso-csw&lbnode=mdlb1}
Disciplinary Metadata Standards, Related Tools and Repositories:
Earth Science- Tools

- **Geospatial Metadata Tools** (from FGDC website)

- **AgriMetamaker**
  - A service to facilitate the publication of metadata in the AGRIS (International System for Agricultural Science and Technology) Repository; it conforms to the AGRIS Application Profile, which draws from the Dublin Core and AgMES standards.

- **WiscLINC Metadata Tools Review**
  - The Wisconsin Land INformation Clearinghouse review of metadata tools used for documenting geospatial data and serving geospatial metadata. It includes tools for CSDGM, ISO 19115, Dublin Core, and DIF.
Disciplinary Metadata Standards, Related Tools and Repositories: Earth Science- Repositories and Data Centers

**AGRIS - International System for Agricultural Science and Technology**

A global public domain database using the AgMES standard to describe structured bibliographical records on agricultural science and technology.

**NCDC - National Climatic Data Center**

The world's largest climate data archive, providing climatological services and data worldwide. It currently promotes the FGDC/CSDGM metadata standard for its datasets.

**CEOS International Directory Network**

An international effort to assist users in locating Earth science data sets, data services, and visualizations using DIF metadata. It provides free, online access to metadata on scientific data in the Earth sciences: geoscience, hydrospheric, biospheric, satellite remote sensing, and atmospheric sciences.
Disciplinary Metadata Standards, Related Tools and Repositories:

Physical Science- Standards and Tools

- **CIF - Crystallographic Information Framework**
  - An extensible standard file format and set of protocols for the exchange of crystallographic and related structured data.

- **Tools:**
  - **Software for CIF**
    - The International Union of Crystallography's list of programs and libraries available for use with CIF files.
  - **A Guide to CIF for Authors**
Disciplinary Metadata Standards, Related Tools and Projects: Physical Science- CIF Example

data_global
_chemical_name_mineral 'Abernathyite'
loop_
_publ_author_name
'Ross M'
'Evans H T'
_journal_name_full 'American Mineralogist'
_journal_volume 49
_journal_year 1964
_journal_page_first 1578
_journal_page_last 1602
_publ_section_title
;
Studies of the torbernite minerals (I): The crystal structure of abernathyite and the structurally related compounds NH4(UO2AsO4).3H2O and K(H3O)(UO2AsO4)2.6H2O;
_database_code_amcsd 0000130
_chemical_compound_source 'Fuemrol No. 2 mine, Temple Mountain, Emery County, Utah, USA'
_chemical_formula_sum '(K U) As O9 H6'
_cell_length_a 7.176
_cell_length_b 7.176
_cell_length_c 18.126
_cell_angle_alpha 90
_cell_angle_beta 90
_cell_angle_gamma 90
_cell_volume 933.398
_exptl_crystal_density_diffrn 3.573
_symmetry_space_group_name_H-M 'P 4/n c c'
One of the symmetry operations of the space group in algebraic form. Symmetry operations are used to generate the atoms in the unit cell defined by the setting used.

Fractional atomic coordinates of atoms in a cell

6522 matching records for this search.

Abernathyite
Ross M. Evans H T
American Mineralogist 49 (1964) 1578-1602
Studies of the torbernite minerals (I): The crystal structure of abernathyite and the structurally related compounds NH4(UO2AsO4).3H2O and K(H3O)(UO2AsO4)2.6H2O
Locality: Puerrol No. 2 mine, Temple Mountain, Emery County, Utah, USA

Record web display
American Mineralogist Crystal Structure Database

A CIF crystal structure database that includes every structure published in the American Mineralogist, The Canadian Mineralogist, European Journal of Mineralogy and Physics and Chemistry of Minerals, as well as selected datasets from other journals.

Crystallography Open Database

An open-access collection of crystal structures of organic, inorganic, metal-organic compounds and minerals, many of which are in CIF form.
Disciplinary Metadata Standards, Related Tools and Repositories

For more information on disciplinary metadata standards, tools, and use cases, please refer to UK Digital Curation Centre (DCC)’s Disciplinary Metadata page.

DataBib: Databib is a community-driven, annotated bibliography of research data repositories. http://databib.org/


Open Access Directory: Data Repositories A list of repositories and databases for open data. It is part of the Open Access Directory maintained by Simmons College. http://oad.simmons.edu/oadwiki/Data_repositories

For more information on data repositories and digital repositories, please refer to Databib, OpenDOAR and OAD.
<table>
<thead>
<tr>
<th>Standard DIF</th>
<th>Dublin Core Metadata</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entry_Title</td>
<td>Title</td>
</tr>
<tr>
<td>Data_Set_Citation: Dataset_Creator</td>
<td>Creator</td>
</tr>
<tr>
<td>Personnel: Role: Investigator: Last_Name</td>
<td>Subject and Keywords</td>
</tr>
<tr>
<td>Personnel: Role: Investigator: First_Name</td>
<td></td>
</tr>
<tr>
<td>Personnel: Role: Investigator: Middle_Name</td>
<td></td>
</tr>
<tr>
<td>Keyword</td>
<td></td>
</tr>
<tr>
<td>Parameters: Category</td>
<td></td>
</tr>
<tr>
<td>Parameters: Topic</td>
<td></td>
</tr>
<tr>
<td>Parameters: Term</td>
<td></td>
</tr>
<tr>
<td>Parameters: Variable</td>
<td></td>
</tr>
<tr>
<td>Parameters: Detailed_Variable</td>
<td></td>
</tr>
<tr>
<td>Source_Name</td>
<td></td>
</tr>
<tr>
<td>Sensor_Name</td>
<td></td>
</tr>
<tr>
<td>Project</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td>Description</td>
</tr>
<tr>
<td>Data_Set_Citation: Dataset_Publisher</td>
<td>Publisher</td>
</tr>
<tr>
<td>Data_Center: Data_Center_Name</td>
<td></td>
</tr>
<tr>
<td>Data_Center: Data_Center_URL</td>
<td></td>
</tr>
<tr>
<td>Data_Center: Data Center Contact, Last_Name</td>
<td></td>
</tr>
<tr>
<td>Data_Center: Data Center Contact, First_Name</td>
<td></td>
</tr>
<tr>
<td>Data_Center: Data Center Contact, Middle_Name</td>
<td></td>
</tr>
<tr>
<td>Personnel: Role:</td>
<td>Contributor:</td>
</tr>
<tr>
<td>Personnel: Last_Name</td>
<td></td>
</tr>
<tr>
<td>Personnel: First_Name</td>
<td></td>
</tr>
<tr>
<td>Personnel: Middle_Name</td>
<td></td>
</tr>
<tr>
<td>Data_Set_Citation: Dataset_Release_Date</td>
<td>Date</td>
</tr>
<tr>
<td>Data_Set_Citation: Data_Presentation_Form</td>
<td>Resource Type</td>
</tr>
</tbody>
</table>

**Metadata Mapping**

A **crosswalk** provides a mapping of metadata elements from one metadata standard to another for the purpose of:

- Lowering the barrier for the management and aggregation of digital resources;
- Increasing data interoperability and data sharing;
- Supporting cross-domain searching and discovery.

**Mapping example: DIF to Dublin Core**
### Metadata Mapping (Cont.)

- Compared to generic Dublin Core, discipline metadata standards provide more granularity;
- Dublin Core (DC), however, provides greater interoperability for data sharing.
- OAI_DC is used within the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).

http://www.icpsr.umich.edu/icpsrweb/ICPSR/dc/studies/20363

**Fields:**
- oai_dc:dc
  - dc:title
  - dc:creator
  - dc:subject
  - dc:description
  - dc:date
  - dc:type
  - dc:identifier
  - dc:coverage
  - dc:rights

**Metadata Harvesting:** an automated, regular process of collecting metadata descriptions from different sources to create useful aggregations of metadata and related services.

**OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting)** is a protocol developed by the Open Archives Initiative.

OAI-PMH data harvesting requires that data providers support the oai_dc metadata format.
Controlled Vocabulary

- **Controlled vocabulary** is a standardized set of terms used to organize knowledge for subsequent retrieval. It can facilitate search and browsing. It can be universally agreed on or locally created.

- What to consider in applying or designing a thesauri for your project:
  - **Scope** of the material (core and surrounding topics, your purpose, existing thesauri and your resource);
  - Your **project needs** and intended **audience**;
  - **Funder** requirements and **institutional** expectation;
  - What **types** of controlled vocabularies you may need: subject, genre, physical format, personal names, organization names, events...
  - When choosing particular terms over others, consider **three warrants**: **literary warrant** (discipline and field literature), **user warrant** and **organizational warrant**. (Gazan, CONTROLLED VOCABULARY & THESAURUS DESIGN, http://www.loc.gov/catworkshop/courses/thesaurus/pdf/cont-vocab-thes-trnee-manual.pdf)
Controlled Vocabulary (Cont.)

- For traditional library catalog:

- For digital and online resources:
  - Internet Media Types [www.iana.org/assignments/media-types/index.html](http://www.iana.org/assignments/media-types/index.html)
  - MODS Note Types [http://www.loc.gov/standards/mods/mods-notes.html](http://www.loc.gov/standards/mods/mods-notes.html)
Controlled Vocabulary: Subject Thesauri

- **Subject Thesauri**
  - AGROVOC (Agricultural Organization of the United Nations Vocabulary)
  - Astronomy Thesaurus
  - CAB Thesaurus (for life sciences, technology and social sciences)
  - CIF dictionaries (for Physics)
  - Eurovoc (European Union Thesaurus)
  - Ethnographic Thesaurus
  - GeoNames
  - Getty Institute Art and Architecture Thesaurus Online
  - Getty Institute Thesaurus of Geographic Names
  - Library of Congress Authorities for subject headings
  - Library of Congress Thesaurus for Graphic Materials
  - MESH (Medical Subject Headings)
  - Public Health Language
  - Rare Books and Manuscripts Section (RBMS) Controlled Vocabularies
  - STW Thesaurus for Economics
  - UNBIS Thesaurus
  - UNESCO Thesaurus
  - USDA National Agricultural Library Agriculture Thesaurus
**Controlled Vocabulary: Name Authorities**

**Library of Congress Name Authority File (LC/NAF)**

The LC/NAF provides authoritative data for names of persons, organizations, events, places, and titles.

**Virtual International Authority File (VIAF)**

The VIAF™ (Virtual International Authority File) combines multiple name authority files into a single OCLC-hosted name authority service. The goal of the service is to lower the cost and increase the utility of library authority files by matching and linking widely-used authority files and making that information available on the Web.

**Getty Union List of Artist Names (ULAN)**

The ULAN includes proper names and associated information about artists. Artists may be either individuals (persons) or groups of individuals working together (corporate bodies). Artists in the ULAN generally represent creators involved in the conception or production of visual arts and architecture.
Web Ontology Language (OWL)
The OWL Web Ontology Language is designed for use by applications that need to process the content of information instead of just presenting information to humans. OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF, and RDF Schema (RDF-S) by providing additional RDF vocabulary along with a formal semantics.

Resource Description Framework (RDF)
RDF is a standard model for data interchange on the Web. RDF extends the linking structure of the Web to use URIs to name the relationship between things as well as the two ends of the link (this is usually referred to as a “triple”). Using this simple model, it allows structured and semi-structured data to be mixed, exposed, and shared across different applications.

SKOS: Simple Knowledge Organization for the Web
SKOS is a W3C recommendation designed for representation of thesauri, classification schemes, taxonomies, subject-heading systems, or any other type of structured controlled vocabulary.

Library linked data examples:
- Virtual International Authority File (VIAF);
- FAST: Faceted Application of Subject Terminology;
- Dewey Decimal Classification;
- Open Metadata Registry (RDA vocabularies);

MADS/RDF
The Metadata Authority Description Schema (MADS) is an XML schema for an element set that may be used to provide metadata about authorized forms of agents (people, organizations), events, and terms (topics, geographics, genres, etc.). MADS/RDF builds on MADS/XML as a knowledge organization system.
More Metadata and Data Curation Tools

DataUp
http://dataup.cdlib.org/

Colectica for Microsoft Excel
http://www.colectica.com/software/colecticaforexcel

OpenRefine (ex-Google Refine)
http://openrefine.org/

Nesstar Publisher
http://www.nesstar.com/software/publisher.html

DataCite
http://www.datacite.org/

FITS (to validate and extract technical metadata)
http://code.google.com/p/fits/

XMetaL
http://xmetal.com/content-xmetal-for-sharepoint/

Schematron
http://xml.ascc.net/resource/schematron/schematron.html

Altova XMLSpy
http://www.altova.com/xmlspy.html

LabTrove
http://www.labtrove.org/

Kepler
https://kepler-project.org/

MIXED (Migration to Intermediate XML for Electronic Data)
https://sites.google.com/a/datanetworkservice.nl/mixed/

TemaTres (Controlled vocabulary server)
http://sourceforge.net/projects/tematres/

DMPTool
https://dmp.cdlib.org/

FITS
http://dataup.cdlib.org/

Colectica for Microsoft Excel
http://www.colectica.com/software/colecticaforexcel

OpenRefine (ex-Google Refine)
http://openrefine.org/

Nesstar Publisher
http://www.nesstar.com/software/publisher.html

DataCite
http://www.datacite.org/

FITS (to validate and extract technical metadata)
http://code.google.com/p/fits/

XMetaL
http://xmetal.com/content-xmetal-for-sharepoint/

Schematron
http://xml.ascc.net/resource/schematron/schematron.html

Altova XMLSpy
http://www.altova.com/xmlspy.html

LabTrove
http://www.labtrove.org/

Kepler
https://kepler-project.org/

MIXED (Migration to Intermediate XML for Electronic Data)
https://sites.google.com/a/datanetworkservice.nl/mixed/

TemaTres (Controlled vocabulary server)
http://sourceforge.net/projects/tematres/

DMPTool
https://dmp.cdlib.org/

DataCite
http://www.datacite.org/
Related Library Resources

○ Scholarly Communication:  (http://library.ucf.edu/ScholarlyCommunication/)
  ○ SC Contact Information (http://library.ucf.edu/ScholarlyCommunication/Contact.php)
  ○ Metadata Services (http://library.ucf.edu/ScholarlyCommunication/Metadata.php)

○ UCF Library Research Guides (http://guides.ucf.edu)
  ○ Metadata guide (http://guides.ucf.edu/metadata)

○ UCF Digital Collections (http://library.ucf.edu/Systems/DigitalCollections/)

○ Research and Information Services (http://library.ucf.edu/Reference/)
  ○ Subject Librarians (http://library.ucf.edu/SubjectLibrarians/)

○ Please refer to “Dataset Metadata” session for information on: data documentation, metadata requirements for data management, researcher identifiers and digital object identifiers.
Contact:
Sai Deng
Metadata Librarian
sai.deng@ucf.edu
407-823-4312 (Office)

Thank you!