Documentation Resources on the ESIP Wiki

Ted Habermann, John Kozimor, Sean Gordon
The HDF Group
thabermann@hdfgroup.org

This work was supported by NASA/GSFC under Raytheon Co. contract number NNG15HZ39C
Terminology

**Concept**: General term for describing a documentation entity.

**Dialect**: A particular form of the documentation language that is specific to a community.

**Recommendation**: A set of concepts that a group believes is required for achieving a documentation goal.

**Spiral**: A set of concepts required to support a particular documentation need or use case.

**Collection**: A group of metadata records, commonly organized by data center, organization or project and often stored in a database or web accessible folder.
Documentation Connections

Category:Documentation Connections

Introduction

The ESIP Community supports a vast array of systems that are accessed and utilized by a diverse group of users. Historically, groups within the community have approached metadata differently in order to effectively describe their data. As a result, similar dialects have emerged to address specific user requirements. The multi-dialect approach described above hinders interoperability—as it results in different terminology being used to describe the same concepts. By clearly depicting fundamental documentation needs and concepts and mapping them in the different dialects, confusion is minimized and interoperability is facilitated. Thus, demonstrating connections between dialects increases discoverability, accessibility, and reusability of data via consistent, compatible metadata.

This document describes the connections between fundamental concepts in dialects used throughout the ESIP Community—such that effective communication is maintained even when different metadata models are employed.

Preface

The following pages are intended to help understand connections between metadata dialects in order to implement complete and consistent metadata. For convenience, this reference is divided into 6 major sections, consisting of: introductory material, documentation recommendations, metadata implementation, documentation selection scenarios, and the Appendices. Each of these sections is further divided into subsections. For your reference, a full table of contents is provided below.

Table of Contents

Documentation Terminology
Concepts Glossary
Documentation Recommendations
Metadata Dialects
Metadata Implementation
Documentation Selection Scenarios

Subcategories

This category has only the following subcategory.

M
* Metadata Implementation

http://wiki.esipfed.org/index.php/Category:Documentation_Connections
## Concepts Glossary

<table>
<thead>
<tr>
<th>Concept</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>A paragraph describing the resource.</td>
</tr>
<tr>
<td>Acknowledgement</td>
<td>Provides a place to acknowledge various types of support for the project that produced the data.</td>
</tr>
<tr>
<td>Additional Attributes</td>
<td>Attributes used to describe unique characteristics of the collection which extend beyond those defined in the model.</td>
</tr>
<tr>
<td>Additional Attributes - Content Information</td>
<td>This concept stores the additional attributes that are related to granule content.</td>
</tr>
<tr>
<td>Additional Attributes - Descriptive Keywords</td>
<td>This concept stores the additional attributes values for descriptive keywords.</td>
</tr>
<tr>
<td>Additional Attributes - Geographic Identifier</td>
<td>This concept stores the additional attributes values for geographic identifiers.</td>
</tr>
<tr>
<td>Additional Attributes - Instrument</td>
<td>This concept stores the additional attributes values for instrument information.</td>
</tr>
<tr>
<td>Additional Attributes - Platform</td>
<td>This concept stores the additional attributes values for platform information.</td>
</tr>
<tr>
<td>Additional Attributes - Quality Information</td>
<td>This concept stores the additional attributes values for Quality Information.</td>
</tr>
<tr>
<td>Additional Lineage Documentation</td>
<td>Additional descriptive information, such as a publication that describes the process for generating the resource.</td>
</tr>
<tr>
<td>Additional Resource Documentation</td>
<td>Other documentation associated with the resource.</td>
</tr>
<tr>
<td>Additional Usage Documentation</td>
<td>Publications that describe usage of data.</td>
</tr>
<tr>
<td>Address</td>
<td>Address line for the location</td>
</tr>
<tr>
<td>Algorithm Citation</td>
<td>Information identifying the algorithm and version or date.</td>
</tr>
</tbody>
</table>
Recommendations

Documentation Recommendations

While the procedure for a design effort is standardized — define the requirements, design and develop, and implement — the specific requirements and subsequent designs will vary depending on the needs of the user(s). Metadata requirements are no exception, as different communities have different requirements. Thus historically, metadata content has been approached in a variety of ways depending on the needs of specific user communities and resulted in the development of multiple, diverse metadata “dialects.” Consequently, sharing data outside specific user communities is hindered by interoperability challenges.

In order to facilitate discoverability, accessibility, usability, understandability, and interoperability of data across disciplines with differing requirements, the gap between the needs/requirements of the provider community and the needs/requirements of other communities must be bridged.

Recall that recommendations are the metadata concepts (elements) that are required, recommended, or suggested for a particular documentation need. This section provides documentation recommendations for a variety of metadata purposes. Included in each documentation recommendation on the following pages is: 1) Recommendation concepts with a brief description thereof, and 2) a list of XPaths for each concept that best bridges the gap to other dialects. Thus, the recommendations specify both the “what” (concepts) as well as the “how” (XPaths) – which enables the user to get from the needs of a provider community (dialect X) to the needs of other communities (dialect Y, Z, etc.). Therefore, through use of these recommendations, data can become available to a significantly larger pool of users.

The recommendations are subdivided into four categories: Discovery, Accessibility/Usability, Understanding.

### Contents

1. Discovery
2. Data Accessibility/Usability
3. Data Understanding
4. Documentation Recommendation Comparison Tool

### Discovery

Discoverability is the ability of information to be found. Discoverability of data is particularly important in the sciences, as the benefits of scientific investigation can’t be used if people can’t find the data. The discoverability of a dataset depends on the completeness and compatibility of the metadata. The more complete and compatible the metadata, the more likely it will be that a user will be able to discover the data they are seeking.

This sub-section aims to clearly depict discoverability documentation needs and concepts and map to them in the different dialects — thereby increasing compatibility and by default, discoverability.

- Data Discovery (ACDD)
- Data Discovery (CSW)
- Data Discovery (QudatCid)
Dialects

Metadata Dialects

Metadata content can be approached in a variety of "dialects," depending on the needs of specific user communities. Though different, these languages also significantly overlap — as the "who, what, when, why, and how" must always be addressed, regardless of the community approach. Thus, in reality, these differences in approach are more akin to dialects of a universal documentation language more than disparate languages. As such, for the purposes of this work, the term "metadata dialect" will refer to standardized metadata documentation approaches, in order to promote emphasis on universal documentation concepts as opposed to implementation of individual standards. The following are some of the most common dialects used throughout the ESIP community.

Note: While they are discussed independently, a dialect can use aspects of other dialects within its own — if the two dialects have the same/similar structure or the same life format.

- ADTag (Alaska Data Integration Working Group)
- CDGM (FGDC Content Standard for Digital Geospatial Metadata)
- DCAT (Data Catalog Vocabulary)
- Ddf (DataCite 3.1)
- DIF (Directory Interchange Format)
- Dryad
- ECHO (EARTH OBSERVING SYSTEM (EOS) CLEARINGHOUSE)
- ECS (EOSDIS Core System)
- EML (Ecological Metadata Language)
- HCLB (Dataset Descriptions: HCLB Community Profile)
- HDF ECDS (Hierarchical Data Format Earth Observing System 5)
- ISO
- ISO-1
- netCDF (Network Common Data Format) Conventions
- SRF (Service Entry Resource Format)
- SOS (Sensor Observation Service)
- THREDDS (Thematic Remote Environmental Distributed Data Services)
- WSDL (Web Service Description Language)

Let's Start at the Beginning
Guidance Pages

Citation Documentation

I need to provide references for datasets and associated resources.

Contents [hide]

1 Overview
2 Content Standard for Digital Geospatial Metadata (CSDGM)
   2.1 Structure
   2.2 Usage
3 Directory Interchange Format (DIF)
   3.1 Structure
4 EOS Clearinghouse (ECHO)
   4.1 Structure
5 International Standards Organization (ISO)
   5.1 Structure
   5.2 Usage
   5.3 CI_Citation++
   5.4 Revisions
6 Crosswalks
7 Notes
   7.1 CodeLists as Types
ISO Metadata Explorer

http://wiki.esipfed.org/index.php/MI_Metadata

http://wiki.esipfed.org/index.php/CI_Responsibility

http://wiki.esipfed.org/index.php/ISO_19115
ISO Metadata Explorer
Acknowledgements

This work was partially supported by contract number NNG15HZ39C from NASA.

Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author and do not necessarily reflect the views of NASA or The HDF Group.