DCC Disciplinary Metadata

Alex Ball

DCC/UKOLN Informatics, University of Bath

2013-09-06
Motivation

“...disciplinary metadata standards... indicate the domain-specific information that will allow data to be interpreted correctly by others in the field. Since data curators cannot become experts in all of the subjects under research within their institutions, a particular need exists for guidance regarding disciplinary metadata standards.”

— Liz Bedford (emphasis added)
Catalogue elements

Disciplinary metadata standard
Catalogue elements

Disciplinary metadata standard

More specific profile
Catalogue elements

Disciplinary metadata standard

More specific profile

General metadata standard
Catalogue elements

- Disciplinary metadata standard
  - Helper tool
  - More specific profile
  - General metadata standard
Scope

- Metadata standards, but not
  - languages/protocols
  - taxonomies/vocabularies
- Descriptive metadata, but not
  - administrative metadata
  - preservation metadata
  - structural metadata
- Active research data, but not
  - publications
  - learning objects
- Tabular data, but not
  - audio
  - video
  - narrative text
Process

1. Literature review
   ▶ Ball, *Scientific Data Application Profile Scoping Study*
   ▶ Riley & Becker, *Seeing Standards: A Visualization of the Metadata Universe*
   ▶ DCC Diffuse
   ▶ UKOLN Application Profiles Support Project

2. Repository review
   ▶ Databib

3. Early thoughts about implementation

4. Choose taxonomy of disciplines (HESA JACS 3.0)

5. Assemble information about standards, profiles/extensions, tools and use cases

6. Finalize implementation and enter information on website

7. Testing and release
Disciplinary Metadata catalogue

Search by Discipline

- Biology
- Earth Science
- General Research Data
- Physical Science
- Social Science & Humanities

Search by Resource Type

**Metadata Standards**
- Specifications for the minimum information that should be collected about research data in order for it to be reused.

**Profiles and Extensions**
- Standards that have been adapted for use in particular types of repositories, or for particular types of data.

**Use cases**
- Institutional repositories and data portals using standards to determine which metadata should be collected upon data deposit.

**Tools**
- Software that has been developed to capture or store metadata conforming to a specific standard.
Physical Science

Nuclear and Particle Physics  Chemistry  Physics  Crystallography  Materials  Science  Solar physics  Space science  Astronomy  Multi-disciplinary  Biochemistry

Metadata Standards

**AVM - Astronomy Visualization Metadata**
A standard defining discovery metadata for fully rendered astronomical imagery.

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

**CSMD-CCLRC Core Scientific Metadata Model**
A study-data oriented model that captures high-level information about scientific studies and the data that they produce, primarily tailored for the physical sciences.

**International Virtual Observatory Alliance Technical Specifications**
A set of specifications, including metadata standards, that enable the integration of many astronomical archives into an international virtual observatory.

**SPASE Data Model**
An information model for describing the elements of the heliophysics data environment.

Extensions

**eBank UK Metadata Application Profile**
A Dublin Core Metadata Application Profile created for the eBank UK project, which provides access to the detailed results of scientific experiments in crystallography.
SPASE Data Model

An information model for describing the elements of the heliophysics data environment, and a set of resource types which can be used to describe data along with its scientific context, source, provenance, content and location. It is designed to support a federated data system where data may reside at different locations and may be separated from the metadata which describes it. The preferred expression form is XML.

The Space Physics Archive Search and Extract (SPASE) effort is implemented by the SPASE Consortium which is composed of representatives of the international Heliophysics data community. The Current Release of the data model (2.2.2) was updated in October 2012.

<table>
<thead>
<tr>
<th>Mappings</th>
<th>OAI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Related Vocabularies</td>
<td>SPASE Dictionary</td>
</tr>
<tr>
<td>Specification</td>
<td><a href="http://www.spase-group.org/docs/schema/">http://www.spase-group.org/docs/schema/</a></td>
</tr>
<tr>
<td>Standard's website</td>
<td><a href="http://www.spase-group.org/data/">http://www.spase-group.org/data/</a></td>
</tr>
</tbody>
</table>

Extensions

**IMPEX Data Model**
A simulation extension to the SPASE data model.

Tools

**SPASE Metadata Editor**
A web-based editor for generating SPASE descriptions.

**SPASE Tools**
The SPASE website's list of tools for working with SPASE metadata and the SPASE framework.

Use Cases

**NSSDC SPASE Registry**
The National Space Science Data Center's registry of SPASE-described space science mission data.

**SPASE Inside**
The SPASE website's list of systems that use SPASE compliant metadata to enable search services.
Next steps

- Review periodically for currency
- Add entries in response to suggestions
- Work with RDA Metadata Standards Directory Working Group

not conducive to collective maintenance. In the UK, JISC and Digital Curation Centre recently launched the Disciplinary Metadata resource covering a variety of disciplines, and the RDA MSDIG has been evaluating this resource. The assessment to-date reveals that this is an important accomplishment to build upon. It is apparent that collaborative metadata registration systems have...
Thank you for your attention

DCC Website: http://www.dcc.ac.uk/
Alex Ball: http://alexball.me.uk/

DCC Disciplinary Metadata:
http://www.dcc.ac.uk/resources/metadata-standards