



DCC Disciplinary Metadata Catalogue Demo

Alex Ball

DCC/UKOLN Informatics, University of Bath

27 March 2014



Except where otherwise stated, this work is licensed under the Creative Commons Attribution 4.0 International licence:
<http://creativecommons.org/licenses/by/4.0/>

The Research Data Alliance is supported by the European Commission, the US Government and the Australian Government.



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)

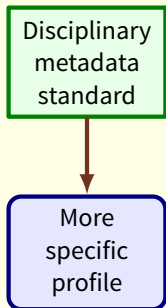


Catalogue elements

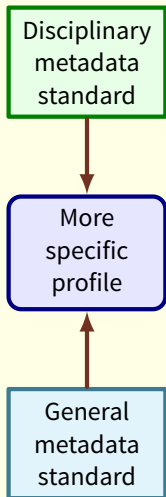
Disciplinary
metadata
standard



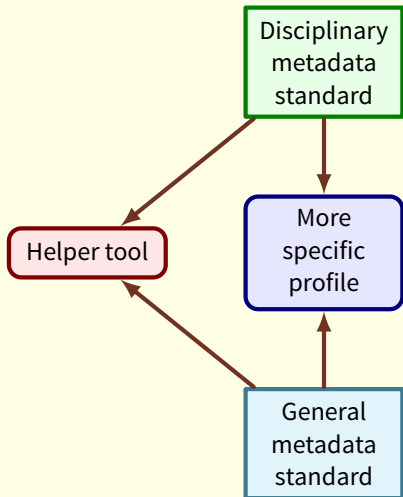
Catalogue elements



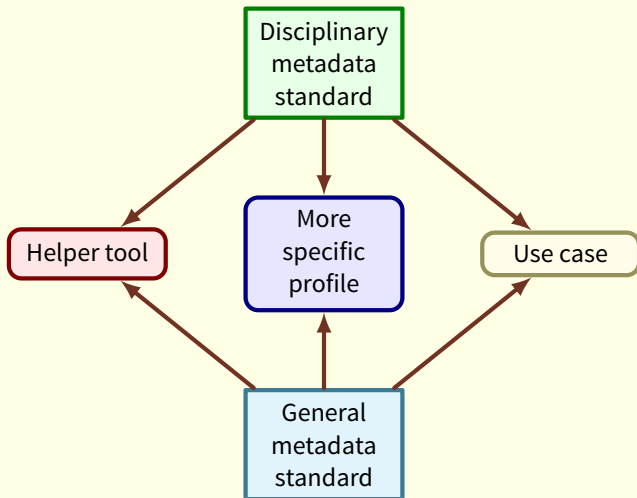
Catalogue elements



Catalogue elements



Catalogue elements



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

▶ Onward!



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 re-used.

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.

List of Metadata Standards

ABCD - Access to Biological Collection Data

The [Access to Biological Collections Data \(ABCD\) Schema](#) is an evolving comprehensive standard for the access to and exchange of data about specimens and observations (a.k.a. primary biodiversity data). The ABCD Schema attempts to be comprehensive and highly structured, supporting data from a wide variety of databases. It is compatible with several existing data standards. Parallel structures exist so that either (or both) atomised data and free-text can be accommodated.

Sponsored by Biodiversity Information Standards TDWG - the Taxonomic Databases Working Group, the current specification was last modified in 2007.

AgMES - Agricultural Metadata Element Set

A semantic standard developed by the Food and Agriculture Organization (FAO) of the United Nations, AgMES enables description, resource discovery, interoperability and data exchange of different types of information resources in all areas relevant to food production, nutrition and rural development.

Sponsored by the UN AIMS - Agricultural Information Management Standards, the current standard was issued in November 2010.

AVM - Astronomy Visualization Metadata

The [AVM](#) scheme supports the cross-searching of collections of print-ready and screen-ready astronomical imagery rendered from telescopic observations (also known as 'pretty pictures'). The scheme is compatible with the [Adobe XMP](#) specification, so the metadata can be embedded within common image formats such as JPEG, TIFF and PNG.

Such images can combine data acquired at different wavebands and from different observatories. While the primary intent is to cover data-derived astronomical images, there are broader uses as well. Specifically, the most general subset of this schema is also appropriate for describing artwork and illustrations of astronomical subject matter.

AVM is a proposed recommendation of the International Virtual Observatory Alliance and was last updated in 2011.

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 re-used.

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.

List of Metadata Profiles and Extensions

ABCDDNA

An extension of the [ABCD](#) standard for DNA data.

ABCDEFGF - Access to Biological Collection Databases Extended for Geosciences

An extension of the [ABCD](#) standard for Geosciences data.

AGLS Metadata Profile

An application of [Dublin Core](#) designed to improve visibility and availability of online resources, originally adapted from the Australian Government Locator Service metadata standard for use in government agencies.

AGRIS Application Profile

A metadata standard drawing on [Dublin Core](#) and [AgMES](#) created specifically to enhance the description, exchange and subsequent retrieval of agricultural Document-Like Information Objects.

ANZLIC Metadata Profile

A profile of [ISO 19115](#), also mapping to the AGLS profile of [Dublin Core](#), designed to facilitate efficient access to descriptions of information resources, particularly geographic or spatial data.

Apple Core

[Darwin Core](#) documentation and recommendations for herbaria.

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 re-used.

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.

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.

Mappings	OAI
Related Vocabularies	SPASE Dictionary
Specification	http://www.spase-group.org/docs/schema/
Standard's website	http://www.spase-group.org/data/

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

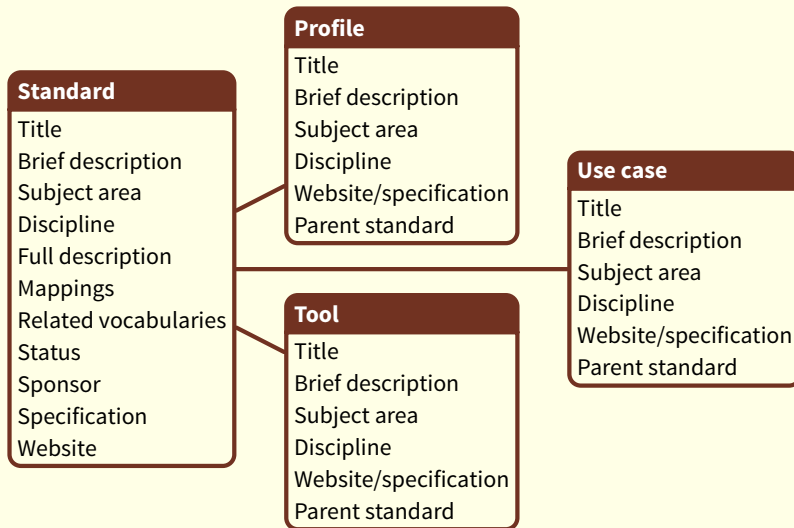
NSSDC SPASE Registry

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.

Data model



Why not just use the DCC Catalogue?

- ▶ Not obvious how to contribute updates or recommend new entries
- ▶ Part of a larger site, so limited functionality and write access
- ▶ Hosted by single organization
- ▶ Scope not quite broad enough
 - ▶ Designed to support UK Higher Education
 - ▶ Excludes contextual/administrative metadata standards
 - ▶ Excludes preservation metadata standards
 - ▶ Excludes structural metadata standards
 - ▶ Excludes standards relating to non-tabular data: audio, video, interview transcripts, etc.



Why not just use the DCC Catalogue?

- ▶ Not obvious how to contribute updates or recommend new entries
- ▶ Part of a larger site, so limited functionality and write access
- ▶ Hosted by single organization
- ▶ Scope not quite broad enough
 - ▶ Designed to support UK Higher Education
 - ▶ Excludes contextual/administrative metadata standards
 - ▶ Excludes preservation metadata standards
 - ▶ Excludes structural metadata standards
 - ▶ Excludes standards relating to non-tabular data: audio, video, interview transcripts, etc.

Good starting point, but more work needed.



Phase 1: survey

Sent to lists:

- ▶ RDA-All
- ▶ RDA MSDWG
- ▶ RDA Metadata Interest Group
- ▶ EuroCRIS
- ▶ EPOS
- ▶ DC-SAM
- ▶ ASIS&T RDAP summit series
- ▶ DataONE
- ▶ ESIP
- ▶ STFC
- ▶ ...

Metadata Directory Collection

In cooperation with the UNC SILS Metadata Research Center and the Research Data Alliance students at UNC SILS are gathering information about metadata standards that apply to scientific data. The aim of this is achieve a short term goal of the Research Data Alliance's Metadata Standards Directory Working Group to:

Develop a prototype wiki-based directory (RDA Metadata Directory) listing metadata standards applicable to scientific data. The initial emphasis will be on widely-used and domain community- endorsed metadata standards and schemas with significant interoperability / re-use capability.

Information submitted to this project will be integrated with similar information from the Digital Curation Center to prototype and build a sustainable platform for sharing and exposing information about metadata standards.

If you have previously submitted information to our collection: check the Digital Curation Center's directory of standards for science data at: <http://www.dcc.ac.uk/resources/metadata-standards>

Please submit one standard per form. At the end of the questionnaire a link to another form submission is available.

* Required

Standard Name *

Description of the Metadata Standard

May be a URL to about page or free text summary.

Metadata Standard URL *

URL for the standard's home page

Domain of the Metadata Standard *

Domains where the standard is in use

- Biology
- Earth Science
- Physical Science

18 contributions (14 new standards, 4 extensions) 13 new tools for generating metadata

DCC @: <http://www.dcc.ac.uk/resources/metadata-standards>

STANDARDS ADDED TO DCC DIRECTORY

1. ClinicalTrials.gov Protocol Data Element Definitions (DRAFT)
2. FITS (Flexible Image Transport System)
3. GenDMS
4. Genome Metadata
5. GEOSS Standards Registry
6. Journeau
7. Observ-OM
8. Observations and Measurements
9. OAI-ORE
10. PROV
11. QuDEX
12. Resource Metadata for the Virtual Observatory
13. WCS - World Coordinate System
14. The Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC)

CONTRIBUTIONS FROM:

Syracuse University
Purdue University Libraries
National Center for Ecological Analysis and Synthesis
DICE Center, UNC-CH
National Snow and Ice Data Center, Univ. of Colorado
ResearXis-Discinnet
University Medical Center Groningen
Commonwealth Scientific Industrial Research Organisation
Open Microscopy Environment
Rensselaer Polytechnic Institute
Newcastle University
Uk Data Archive
US Virtual Astronomical Observatory / Space Telescope Science Institute
EDINA, The University of Edinburgh
U. S. Geoscience Information Network
Met Office, UK
IMOS/AODN Integrated Marine Observing System/Australian Ocean Data Network
NRCan/GeoConnections

Phase 2: exploring new platforms

- ▶ Pages are generated from simple easy-to-edit text files.
- ▶ Maintainers can edit the pages directly.
- ▶ Others can fork the whole directory, make changes on their copy, and submit a pull request.

RDA Metadata Directory Standards Extensions Tools Use Cases Add • Edit View on GitHub

ISO-19115

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. This standard is internationally adopted.

Summary

Standard Website
http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020

Sponsor
Sponsored by the International Standards Organisation.

Standard Update Date
2009-02

Mappings
[FGDC/CSDGM](#)
[DIF](#)
[UK AGMAP](#)
[NatCDF](#)

Subjects
[Earth science, general, biology](#)
Subjects
[cartography, climatology, geography, geoscience, glaciology, hydrogeology, hydrography, marine science, meteorology, oceanography, topography](#)

Extensions [Add](#)

[ANZLIC Metadata Profile](#) [Add](#) [Edit](#)
A profile of ISO 19115, also mapping to the AGLS profile of Dublin Core, designed to facilitate efficient access to descriptions of information resources, particularly geographic or spatial data.

[Cruise Summary Reports](#) [Add](#) [Edit](#)
The National Oceanographic Data Centre's required format for reporting on cruises or field experiments at sea, formulated using tags from the ISO19115 metadata standard.

Tools [Add](#)

[ANZ-MEST Metadata Entry and Search Tool](#) [Add](#) [Edit](#)
A GeoNetwork web application for metadata management and searching, with profiles available for two extensions of ISO 19115: ANZLIC and the MarineCommunity Profile.

[Geodoc Metadata Editor](#) [Add](#) [Edit](#)
The Geodoc metadata editor tool allows users to create, validate, edit and export geospatial metadata records. It also supports the creation and export of metadata records as XML output files compliant with a number of standards, including UK AGMAP 2.1, ISO 19115, FGDC, DDI, and Dublin Core.

Use Cases [Add](#)

Metadata Directory is maintained by [Sean Chen](#)
This page was generated by [GitHub Pages](#).

Use cases

