

Describing resources II: Dublin Core

CERN-UNESCO School on Digital Libraries
Rabat, Nov 22-26, 2010

Annette Holtkamp
CERN

Dublin Core

- basic set of metadata elements
- to describe physical and digital resources
- 1995 workshop in Dublin, Ohio
- Dublin Core Metadata Initiative (DCMI)
 - open forum (W3C, librarians, techies)
 - committed to open standards, interoperability

<http://dublincore.org>

Simple Dublin Core: Elements I

- **title**
 - a name given to the resource
- **creator**
 - entity primarily responsible for making the resource
- **subject**
 - topic of the resource
- **description**
 - abstract, table of contents...
- **publisher**
- **contributor**
- **date**
 - point or period of time associated with an event in the lifecycle of the resource (creation date, last modified...)
- **type**
 - nature or genre of the resource (text, dataset, event...)

Simple Dublin Core: Elements II

- **format**
 - file format, physical medium, or dimensions
- **identifier**
 - unambiguous reference to the resource
- **source**
 - related resource from which the described resource is derived
- **language**
- **relation**
 - related resource
- **coverage**
 - e.g. spatial or temporal topic
- **rights**

Simple DC: Example

title: New Light on Dark Matter from the LHC

creator: Ellis, John

subject: Astrophysics

identifier: <http://inspirebeta.net/record/875145>

description: The prospects for detecting a candidate supersymmetric dark matter particle at the LHC are reviewed...

date: 2010-11-02

type: text

Simple DC: Example in XML

<dc:title>New Light on Dark Matter from the LHC</dc:title>

<dc:creator>Ellis, John</dc:creator>

<dc:subject> Astrophysics</dc:subject>

<dc:identifier><http://inspirebeta.net/record/875145>
</dc:identifier>

<dc:description>The prospects for detecting a candidate supersymmetric dark matter particle at the LHC are reviewed...</dc:description>

<dc:date>2010-11-02</dc:date>

<dc:type>text</dc:type>

Dublin Core elements

- each element optional and repeatable
- no defined order
- searchability enhanced by controlled vocabularies
- defined in a DCMI recommendation
- identified by URI in DCMI namespace

DCMI element recommendation

- Term Name: type
- URI: <http://purl.org/dc/elements/1.1/type>
- Label: Type
- Definition: The nature or genre of the resource.
- Comment: Recommended best practice is to use a controlled vocabulary such as the DCMI Type Vocabulary [DCMITYPE]. To describe the file format, physical medium, or dimensions of the resource, use the Format element.
- References: [DCMITYPE]
<http://dublincore.org/documents/dcmi-type-vocabulary/>

Application: OAI Harvesting

- Open Archives Initiative (OAI):
 - low-barrier interoperability framework for digital archives
 - to build services using metadata from many sources
- Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)
 - technical infrastructure
 - mechanism for data providers to expose their metadata
 - representation in Dublin Core (XML) mandatory
 - additional representations optional
 - widely adopted by libraries, archives, repositories

OAI-PMH: example

http://export.arxiv.org/oai2?verb=GetRecord&identifier=oai:arXiv.org:1011.0077&metadataPrefix=oai_dc

<header>

<identifier>oai:arXiv.org:1010.0077</identifier>

<timestamp>2010-11-102</timestamp> </header>

<metadata>

<oai:dc_dc

xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/...>

<dc:title>New Light on Dark Matter from the LHC</dc:title>

<dc:creator>Ellis, John</dc:creator>

<dc:subject>: Astrophysics</dc:subject>

<dc:description>The prospects for detecting a ...</dc:description>

<dc:date>2010-10-30</dc:date>

<dc:identifier>http://inspirebeta.net/record/875145</dc:identifier>

</oai_dc:dc>

</metadata>

Qualified Dublin Core

- Simple Dublin Core
- plus
- 3 additional elements
 - qualifiers for each element

Qualified DC: Additional elements

- Audience
 - for whom the resource is intended or useful
- Provenance
 - e.g. changes in ownership
- RightsHolder

Qualified DC: Qualifiers

- Element refinement
 - narrows meaning, restricts scope
description: abstract, table of contents
- Encoding scheme
 - to interpret element values
 - e.g. controlled vocabulary, formal notation
subject: LCSH, MeSH, LCC, UDC, DDC

Principles

- Simplicity of creation and maintenance
- universally understood semantics
 - creator: author, artist...
- internationalization
 - versions in many languages
- extensibility
 - domain specific extensions via Application Profiles

Dublin Core Application Profile (DCAP)

framework for defining metadata records beyond DC

- custom-tailored for specific application needs
- may combine terms from multiple namespaces
- based on globally defined vocabularies and models
- semantic interoperability
- integrable into linked data environments

<http://dublincore.org/documents/profile-guidelines/>

DCAP components

- **Functional requirements**
 - describes what a community wants to accomplish with its application
- **Domain model**
 - characterizes the types of things described by the metadata and their relationships
- **Description Set Profile + usage guidelines**
 - metadata terms to be used and rules for their use
- **Syntax guidelines and data formats**
 - machine syntax used to encode the data

MODS

Metadata Object Description Schema

- MARC21 derivative
- subset of key MARC21 elements
- a few elements without MARC21 equivalent
- no field/subfield structure
- verbal instead of numerical tags
- uses XML

MODS example

```
<mods ID="xxx">  
  <titleInfo>  
    <title>New light... </title>  
  </titleInfo>  
  <name type="personal">  
    <namePart type="given">John</namePart>  
    <namePart type="family">Ellis</namePart>  
    <role>  
      <roleTerm type="text">author</roleTerm>  
    </role>  
  </name>  
</mods>
```