

Creating Metadata Records

Introduction

This report describes the creation of metadata related to five electronic resources. The two metadata schemes used are the Electronic Resource Citation (ERC) developed by John Kunze and the Dublin Core (DC) developed by the Dublin Core Metadata Initiative. Following the ten example records is a discussion of some of issues involved in creating metadata records.

Electronic Resource Citation Records

erc:
who: Lynch, Clifford A
what: Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age
when: 2003 - 02
where: <http://www.arl.org/newsltr/226/ir.html>
in: ARL Bimonthly Report 226
erc-about:
what: Institutional Repositories

erc:
who: Powell, Andy | Lyon, Liz
what: JISC Information Environment Architecture - Distributed Systems - UKOLN
when: 2003-07-03
where: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/>
erc-about:
what: DNER Technical Architecture | Standards Framework | Usage Scenarios | JISC Information Environment and Web services | 5 steps to becoming a content provider in the JISC Information Environment | JISC IE Service Registry | Baseline Portal Functional Specification

erc:
who/created: DCMI-Libraries Working Group
who/contributed: Clayphan, Robina | Guenther, Rebecca
what: DC-Library Application Profile (DC-Lib)
when: 2004-09-10
where: <http://dublincore.org/documents/2004/09/10/library-application-profile/>
erc-about:
what: Dublin Core Metadata Element Set | Library Application Schema

erc:
who: Brand, Amy | Daly, Frank | Meyers, Barbara
what: Metadata Demystified: a guide for publishers
when: 2003-07
where: http://www.niso.org/standards/resources/Metadata_Demystified.pdf
erc-about:
what: Metadata | Book Publishing | Journal Publishing | ONIX

erc:
who: Library of Congress
what: METS: An Overview & Tutorial
when: 2004-09-24

where: <http://www.loc.gov/standards/mets/METSOverview.v2.html>
 erc-about:
 what: METS | Metadata Encoding & Transmission Standard | Structural |
 Administrative | Structural | Metadata | XML Schema

Dublin Core Records

```
<META NAME="DC.Title" LANG="en" CONTENT="Institutional Repositories: Essential
Infrastructure for Scholarship in the Digital Age">
<META NAME="DC.Creator" LANG="en" CONTENT="Lynch, Clifford A.">
<META NAME="DC.Subject" LANG="en" CONTENT="Institutional Repositories">
<META NAME="DC.Publisher" LANG="en" CONTENT="The Association of Research
Libraries">
<META NAME="DC.Date" LANG="en" CONTENT="2003-02">
<META NAME="DC.Type" LANG="en" CONTENT="text">
<META NAME="DC.Format" LANG="en" CONTENT="text/html">
<META NAME="DC.Identifier" LANG="en" SCHEME="URI"
CONTENT="http://www.arl.org/newsltr/226/ir.html">
<META NAME="DC.Language" LANG="en" CONTENT="en">
<META NAME="DC.IsFormatOf" LANG="en" CONTENT="Electronic version from ARL, no.
226.">
<META NAME="DC.Rights" LANG="en" CONTENT="The Association of Research
Libraries">
```

```
<META NAME="DC.Title" LANG="en" CONTENT="JISC Information Environment
Architecture - Distributed Systems - UKOLN">
<META NAME="DC.Creator" LANG="en" CONTENT="Powell, Andy">
<META NAME="DC.Creator" LANG="en" CONTENT="Lyon, Liz">
<META NAME="DC.Subject" LANG="en" CONTENT="DNER Technical Architecture;
Standards Framework; Usage Scenarios; JISC Information Environment and Web
services; 5 steps to becoming a content provider in the JISC Information
Environment; JISC IE Service Registry; Baseline Portal Functional
Specification">
<META NAME="DC.Publisher" LANG="en" CONTENT="UKOLN">
<META NAME="DC.Date" LANG="en" CONTENT="2003-07-03">
<META NAME="DC.Type" LANG="en" CONTENT="text">
<META NAME="DC.Format" LANG="en" CONTENT="text/html">
<META NAME="DC.Identifier" LANG="en" SCHEME="URI"
CONTENT="http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/">
<META NAME="DC.Language" LANG="en" CONTENT="en">
```

```
<META NAME="DC.Title" LANG="en" CONTENT="Library Application Profile">
<META NAME="DC.alternative" LANG="en" CONTENT="DC-Library Application Profile
(DC-Lib)">
<META NAME="DC.Creator" LANG="en" CONTENT="DCMI-Libraries Working Group">
<META NAME="DC.Subject" LANG="en" CONTENT="Dublin Core Metadata Element Set">
<META NAME="DC.Subject" LANG="en" CONTENT="Library Application Schema">
<META NAME="DC.abstract" LANG="en" CONTENT="This document proposes a possible
application profile that clarifies the use of the Dublin Core Metadata Element
Set in libraries and library-related applications and projects. It was
originally prepared by the DCMI-Libraries Application Profile drafting
committee, a subset of the DCMI-Libraries Working Group. This revision was
prepared in August 2004 and incorporates decisions made by the DCMI Usage
Board at its meetings in 2003 and issues discussed in the WG meeting in
Seattle in September 2003. It has been reformatted in conformance with the
Dublin Core Application Profile Guidelines produced by the CEN MMI-DC
Workshop.">
<META NAME="DC.Publisher" LANG="en" CONTENT="Dublin Core Metadata Initiative">
<META NAME="DC.Contributor" LANG="en" CONTENT="Clayphan, Robina">
<META NAME="DC.Contributor" LANG="en" CONTENT="Guenther, Rebecca">
<META NAME="DC.Date" LANG="en" CONTENT="2004-09-10">
<META NAME="DC.Type" LANG="en" CONTENT="text">
```

```

<META NAME="DC.Format" LANG="en" CONTENT="text/html">
<META NAME="DC.Identifier" LANG="en"
CONTENT="http://dublincore.org/documents/2004/09/10/library-application-
profile/">
<META NAME="DC.Language" LANG="en" CONTENT="en">
<META NAME="DC.Rights" LANG="en"
CONTENT="http://dublincore.org/about/copyright/index.shtml#copyright">

<META NAME="DC.Title" LANG="en" CONTENT="Metadata Demystified: a guide for
publishers">
<META NAME="DC.Creator" LANG="en" CONTENT="Brand, Amy">
<META NAME="DC.Creator" LANG="en" CONTENT="Daly, Frank">
<META NAME="DC.Creator" LANG="en" CONTENT="Meyers, Barbara">
<META NAME="DC.Subject" LANG="en" CONTENT="Metadata; Book Publishing; Journal
Publishing; ONIX">
<META NAME="DC.Description" LANG="en" CONTENT="An overview of metadata
conventions in publishing and related initiatives designed to standardize how
metadata is structured and disseminated online.">
<META NAME="DC.Publisher" LANG="en" CONTENT="NISO Press">
<META NAME="DC.Publisher" LANG="en" CONTENT="The Sheridan Press">
<META NAME="DC.Date" LANG="en" CONTENT="2003-07">
<META NAME="DC.Type" LANG="en" CONTENT="text">
<META NAME="DC.Format" LANG="en" CONTENT="application/pdf">
<META NAME="DC.Identifier" LANG="en" SCHEME="URI"
CONTENT="http://www.niso.org/standards/resources/Metadata_Demystified.pdf">
<META NAME="DC.Language" LANG="en" CONTENT="en">
<META NAME="DC.isFormatOf" LANG="en" SCHEME="ISBN" CONTENT="1-880124-59-9">
<META NAME="DC.Rights" LANG="en" CONTENT="Copyright 2003, The Sheridan Press
and NISO Press">

<META NAME="DC.Title" LANG="en" CONTENT="METS: An Overview & Tutorial">
<META NAME="DC.Creator" LANG="en" CONTENT="Library of Congress">
<META NAME="DC.Subject" LANG="en" CONTENT="METS; Metadata Encoding &
Transmission Standard; XML Schema; Structural; Administrative; Structural;
Metadata">
<META NAME="DC.Description" LANG="en" CONTENT="METS, a Digital Library
Federation initiative, provides an XML document format for encoding metadata
necessary for both management of digital library objects within a repository
and exchange of such objects between repositories (or between repositories and
their users). ">
<META NAME="DC.Publisher" LANG="en" CONTENT="Library of Congress">
<META NAME="DC.Date" LANG="en" CONTENT="2003-09-24">
<META NAME="DC.Type" LANG="en" CONTENT="text">
<META NAME="DC.Format" LANG="en" CONTENT="text/html">
<META NAME="DC.Identifier" LANG="en"
CONTENT="http://www.loc.gov/standards/mets/METSOverview.v2.html">
<META NAME="DC.Language" LANG="en" CONTENT="en">

```

Discussion

The four required elements of the ERC anchoring story and the fifteen core elements of the simple DC are deceptively simple. Though they can both become complex very easily when the qualifiers are taken into account. One of the more difficult decisions is deciding “where to stop” in metadata creation: how many subject terms, how deep does one go with related format linking? These questions relate to the intended audience and the precision and recall requirements not covered in this introductory assignment.

Values for the ERC four required elements were fairly straight-forward to find either in the content of the resource itself or in its HTML page source. I appreciated the option of including a reason for why an element did not have a value as that reason could also be informative to the user. More information on the refinements allowed on the elements would be helpful.

Each of the sample ERC records includes the `erc-about: story` because a field that describes the intellectual content of the resource seemed appropriate. A combination of Library of Congress Subject Headings and natural language taken from the document text were used. Only the `what:` element was included because `who:`, `when:` and `where:` were unnecessary with these resources. To facilitate resource discovery perhaps an `about: element` should be required in the anchoring story since every resource has an “aboutness” to it. However, since the primary purpose of ERC is for archival and access management, an `about: required` element is probably unnecessary if full-text searching becomes more sophisticated.

Similarly, the values for the DC records were relatively straight-forward to determine. I felt a need to try to find a value for every element, even though one could not always be found. The simple DC elements allow for great flexibility because they are not precisely defined. The qualified DC elements help to narrow the elements definition, but depending on the needs of the user community those extra elements may not be necessary. At times trying to determine the values for the qualified elements was overwhelming when no firm guidelines are available. Though specific application profiles will provide those guidelines for different communities.

The general metadata editor in Notetab was very simple to use and made nicely formatted code while easily allowing for any extra editing. (Being a Mac user, I would like to see a platform-independent, Web-based solution, but Notetab worked very well with VirtualPC.)

Summary and Conclusion

The ERC readings stress the simplicity of this metadata scheme as a way of enticing resource creators to attach metadata. However, the ERC quickly becomes complex when additional stories and element refinements are considered. The Dublin Core elements, while having more elements at the onset, appear more well-defined in semantics, yet flexible enough to be used in most, if not all, situations. Possibly the Dublin Core could have a subset (a sort of DC-mini) that consists of required elements to minimally describe a resource, yet be expanded as needed with the remaining DC elements.

The Dublin Core has many examples of its use in different communities. The ERC still needs more technical documentation and guidelines and some practical examples before improvements over the Dublin Core can be realized.