

Introduction

This report contains the Electronic Resource Citation and Dublin Core records created for the Group E resources and my comments on and discussion of their creation.

To create consistency among all the records, I followed certain syntactical and semantic conventions. All personal names are listed in inverted order (e.g., Jenness, Jennifer) and the dates in both ERC and DC records are listed in YYYY-MM-DD format. ERC records were limited to the anchoring story only in all cases but one, and the DC records were as complete as could be managed with the information provided by the item. This provided an interesting contrast between the systems and served to highlight the advantages and disadvantages of the two schemata.

Electronic Resource Citation Records

Record 1:

who: Brand, Amy | Daly, Frank | Meyers, Barbara

what: Metadata Demystified: A Guide for Publishers

when: 200307

where: http://www.niso.org/standards/resources/Metadata_Demystified.pdf

Record 2:

who: DCMI-Libraries Working Group | Clayphan, Robina | Guenther, Rebecca

what: DC - Library Application Profile

when: 20040910

where: <http://dublincore.org/documents/library-application-profile/>

Record 3:

who: Powell, Andy | Lyon, Liz

what: JISC Information Environment Architecture

when: 20030703

where: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/>

Record 4:

who: Lynch, Clifford A.

what: Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age

when: 200302

where: <http://www.arl.org/newsltr/226/ir.html>

who/published: Association of Research Libraries

what/published: ARL Bimonthly Report 226

when/published: 200302

where/published: <http://www.arl.org/newsltr/226/index.html>

Record 5:

who: The Library of Congress

what: METS (Metadata Encoding and Transmission Standard): An Overview & Tutorial

when: 20040923

where: <http://www.loc.gov/standards/mets/METSOverview.v2.html>

DC Records:

Record 1:

```
<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">
<META NAME="DC.Subject" LANG="en" CONTENT="Publishing">
<META NAME="DC.Description" LANG="en" CONTENT="A guide to metadata conventions
in publishing">
<META NAME="DC.Publisher" LANG="en" CONTENT="The Sheridan Press">
<META NAME="DC.Publisher" LANG="en" CONTENT="NISO Press">
<META NAME="DC.Contributor" LANG="en" CONTENT="Harris, Pat (editor)">
<META NAME="DC.Contributor" LANG="en" CONTENT="Parente, Susan (editor)">
<META NAME="DC.Contributor" LANG="en" CONTENT="Pirkey, Kevin (editor)">
<META NAME="DC.Contributor" LANG="en" CONTENT="Suprock, Greg (editor)">
<META NAME="DC.Contributor" LANG="en" CONTENT="Witkowski, Mark (editor)">
<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"
CONTENT="http://www.niso.org/standards/resources/Metadata_Demystified.pdf">
<META NAME="DC.Language" LANG="en" CONTENT="en">
```

Record 2:

```
<META NAME="DC.Title" 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="Application profiles">
<META NAME="DC.Subject" LANG="en" CONTENT="Dublin Core Metadata Element Set">
<META NAME="DC.Description" LANG="en" CONTENT="A proposal for clarifying an
application profile for the Dublin Core Metadata Element Set ">
<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://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/">
 <META NAME="DC.Relation" LANG="en" CONTENT="IsVersionOf
 http://dublincore.org/documents/2002/09/24/library-application-profile/">

Record 3:

<META NAME="DC.Title" LANG="en" CONTENT="JISC Information Environment
 Architecture">
 <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="Networked Services">
 <META NAME="DC.Subject" LANG="en" CONTENT="Standards and Protocols">
 <META NAME="DC.Subject" LANG="en" CONTENT="Information access">
 <META NAME="DC.Description" LANG="en" CONTENT="A webpage which provides the
 standards and protocols utilized in the JISC Information Environment (JISC-IE)">
 <META NAME="DC.Publisher" LANG="en" CONTENT="UK Office for Library Networking
 (UKOLN)">
 <META NAME="DC.Publisher" LANG="en" CONTENT="The Joint Information Systems
 Committee (JISC)">
 <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" CONTENT="http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/">
 <META NAME="DC.Language" LANG="en" CONTENT="en">

Record 4:

<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="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"
 CONTENT="http://www.arl.org/newsltr/226/ir.html">
 <META NAME="DC.Language" LANG="en" CONTENT="en">
 <META NAME="DC.Relation" LANG="en" CONTENT="IsPartOf ARL Bimonthly Report
 226">

Record 5:

<META NAME="DC.Title" LANG="en" CONTENT="METS: An Overview and Tutorial">

```

<META NAME="DC.Creator" LANG="en" CONTENT="The Library of Congress">
<META NAME="DC.Subject" LANG="en" CONTENT="Metadata Encoding and Transmission
Standard (METS)">
<META NAME="DC.Description" LANG="en" CONTENT="A document describing the
purpose and format of METS intended to provide an introduction to the schema.">
<META NAME="DC.Publisher" LANG="en" CONTENT="The Library of Congress">
<META NAME="DC.Date" LANG="en" CONTENT="2004-09-23">
<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.Relation" LANG="en" CONTENT="IsPartOf
http://www.loc.gov/standards/mets/">

```

Discussion

Working on the creation of these metadata records made me realize just how thoroughly I've been indoctrinated into "proper" cataloging procedures. I continually found myself searching for rules that don't exist, clarification on syntax and semantics, and more specifics on what exactly belongs where. Although I recognize that these schemas are designed to be created and used by laypersons with little or no training, the lack of clear directions for use was frustrating for me. Several conflicts arose which highlighted this problem for me.

One of the problems I encountered was in properly identifying creators or contributors, or, in ERC, the "who," within the record. How complete a title should a person be given? For instance, which would be more correct: Lynch, Clifford A., or Lynch, Clifford A., Executive Director, Coalition for Networked Information? Just giving the creator's name is the simpler option, but providing his title also helps in establishing the authority and provenance of the item. The DC schema has no established convention for this, although *Using Dublin Core* provides a set of input rules for the schema. Another example would be in the DC.Contributor field. Which is the most correct way to list an editor; as Harris, Pat, as Harris, Pat (editor), as Harris, Pat (contributing editor), or as Harris, Pat | Contributing Editor? Because of the vagueness of the syntax and the lack of input rules, determining which format to use can be extremely challenging. Although the ability to create one's own syntax for a specific purpose makes the schema more accessible to a wider variety of people and organizations, interoperability of these systems becomes an almost unachievable task.

One of the major difficulties I had with creating the Dublin Core records was in deciding how much information was required in each field and in each record. This was a particular challenge with the subject fields. I would personally prefer to select subject terms from the Library of Congress Subject Headings or another controlled vocabulary source, but the difficulty with a controlled vocabulary is that the vast majority of users find them to be arcane and hard to use. However, without the use of controlled vocabulary, the finding task becomes incredibly more difficult to support. This issue, of course, is not unique to metadata created for DC or ERC records. The conflict between controlled vocabulary, which can be difficult to understand for the user but helps to ensure records aren't lost due to different wording, and the use of natural

language, which is easier for users to comprehend but most fully supports the finding task, is pervasive throughout library cataloging. However, the issue is complicated in these metadata schemes by the fact that the intended creators and users are not trained in information science. The records are meant to be created and used by anyone.

Another challenge was with specifying entity level. Neither schema provides a means to define the entity level of the information object. This is particularly problematic with ERC, which has no means to identify the entity level. DC does not provide a field to enter entity level information, but it can be included in the description of the object. So, unless the entity level happens to be included in the description field, there is no way to determine it.

Data retrieval was not a major issue for me in this project. For the particular resources we were given, finding the necessary data was not usually too difficult. However, there were several times when the data given was ambiguous or incomplete, which complicated the creation of accurate records with high quality metadata. An example of this can be found on the JISC Information Environment Architecture page. The notation at the bottom of the page gives the information, "Page by Andy Powell." However, many of the other pages within this site are by both Powell and Liz Lyon, and although Lyon's name does not appear on the webpage itself, her name is listed as a creator in the source metadata. The question then becomes, should Lyon's name be added as a creator/who, despite the fact that her name is not found on the page itself? Should she be added in a story other than the anchoring story? Or should she not be mentioned at all in the metadata for the site?

Another, related issue is in the source data itself. A statement such as "Page by Andy Powell" may indicate that Powell is the creator of the original work on the page, or it could indicate that he created the page from information written by someone else. The vagueness of terms such as this makes the job of creating accurate metadata that much more complex.

The simplicity of the semantics for ERC is both a blessing and a curse. Because the semantics for the elements are so basic, the system can be utilized without anything but the most basic training. However, that simplicity can also lead to ambiguous or inaccurate records, particularly when dealing with a complex information item. For instance, how would the anchoring story for a DVD of a film be created? There are many people and organizations that could be classified as responsible parties, each with a distinct and vital part in the creation of the object. The needs of the user should be the determining factor, but with a broad and undefined user base, those needs can be very difficult to ascertain. The when element is also challenging. Which date is the most important, and how do you select from among the multiple dates associated with an item? This field is particularly challenging when working with a website. Often the only dates given are the dates of the most recent update; however, if this date is used, the metadata will be outdated as soon as the page is updated again.

Summary and Conclusion

Although I do feel that the ERC has some excellent points, I found the story premise to be somewhat unwieldy at times. For instance, were the ERC records to be expanded beyond the anchoring story into segments such as erc-about or erc-from, often only one of the required fields

is necessary; the others are either repeats of the elements from the anchoring story, or are irrelevant. It seems unnecessary to enter so many null or repeated values in a system that purportedly exists to make the metadata creation process simpler.

The Dublin Core also has its strengths and weaknesses. I found the lack of syntactical definition frustrating, for it seems to me that this inevitably leads to imprecise and inconsistent records. This limits the searchability of the data, which impedes access by the user. I understand that in order to have an international standard, particularly one created by nonprofessionals, it is difficult if not impossible to regulate things such as subject terms; however, without some form of authority control, there will always be difficulties for the user in locating relevant information.

Although each system has its flaws, both do an admirable job in creating basic, understandable metadata with a relative minimum of complication or confusion. The Dublin Core's flexible structure allows the metadata records to be as complex or as simple as is needed by the creator and the user, while the more rigid structure of the ERC gives each information object a basic minimum of information while still leaving open the opportunity to add additional metadata as needed. Each schema has a different purpose and a different method, but both provide the modularity and flexibility necessary for a successful broad-based metadata scheme.