

## Creating Metadata Records

### Introduction

The purpose of this document is to demonstrate a basic understanding of the Electronic Resource Citation (ERC) and Dublin Core (DC) Metadata Schemes by creating records using both schemes for the five items from Group E. In order to keep the records as simplistic as possible, only the anchoring story was used for the ERC records and all DC records are devoid of any qualifiers.

### Electronic Resource Citation Records

erc:

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

what: Metadata Demystified: A Guide for Publishers

when: 2003

where: [http://www.niso.org/standards/resources/Metadata\\_Demystified.pdf](http://www.niso.org/standards/resources/Metadata_Demystified.pdf)

erc:

who: DCMI-Libraries Working Group

what: DC-Library Application Profile (DC-Lib)

when: 2004-09-10

where: <http://dublincore.org/documents/2004/09/10/library-application-profile/>

erc:

who: Powell, Andy

what: JISC Information Environment Architecture - Distributed Systems - UKOLN

when: 2003-07-03

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

erc:

who: Lynch, Clifford A.

what: ARL 226: Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age

when: February 2003

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

erc:

who: The Library of Congress

what: METS: An Overview & Tutorial

when: (:unap)

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

### Dublin Core Records

```

<META NAME="DC.Creator" LANG="en" CONTENT="Barbara Meyers">
<META NAME="DC.Subject" LANG="en" CONTENT="metadata">
<META NAME="DC.Subject" LANG="en" CONTENT="publishing">
<META NAME="DC.Description" LANG="en" CONTENT="This guide presents an overview
of evolving metadata conventions in publishing as well as related initiative designed to
standarize how metadata is structured and disseminated online. ">
<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="Pat Harris">
<META NAME="DC.Contributor" LANG="en" CONTENT="Susan Parante">
<META NAME="DC.Contributor" LANG="en" CONTENT="Kevin Pirkey">
<META NAME="DC.Contributor" LANG="en" CONTENT="Greg Suprock">
<META NAME="DC.Contributor" LANG="en" CONTENT="Mark Witkowski">
<META NAME="DC.Date" LANG="en" CONTENT="2003">
<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.niso.org/standards/resources/Metadata_Demystified.pdf">
<META NAME="DC.Source" LANG="en" CONTENT="ISBN 1-880124-59-9 ">
<META NAME="DC.Language" LANG="en" CONTENT="en">
<META NAME="DC.Relation" LANG="en" CONTENT="IsVersionOf ">

<META NAME="DC.Subject" LANG="en" CONTENT="Dublin Core Metadata Element Set
">
<META NAME="DC.Subject" LANG="en" CONTENT="library application profile">
<META NAME="DC.Description" 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.">
<META NAME="DC.Publisher" LANG="en" CONTENT="CEN MIMI-DC Workshop">
<META NAME="DC.Language" LANG="en" CONTENT="en">
<META NAME="DC.Contributor" LANG="en" CONTENT="Robina Clayphan">
<META NAME="DC.Contributor" LANG="en" CONTENT="Rebecca Guenther">
<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.Source" LANG="en"
CONTENT="http://dublincore.org/documents/2002/09/24/library-application-profile/">
<META NAME="DC.Language" LANG="en" CONTENT="en">

```

```

<META NAME="DC.Title" LANG="en" CONTENT="JISC Information Environment
Architecture - Distributed Systems - UKOLN">
<META NAME="DC.Creator" LANG="en" CONTENT="Andy Powell">
<META NAME="DC.Subject" LANG="en" CONTENT="JISC Information Environment (JISC
IE)">
<META NAME="DC.Subject" LANG="en" CONTENT=" JISC IE Architecture Standards
Framework. ">
<META NAME="DC.Description" LANG="en" CONTENT="The JISC Information
Environment (JISC IE) technical architecture specifies a set of standards and protocols that
support the development and delivery of an integrated set of networked services that allow the
end-user to discover, access, use and publish digital and physical resources as part of their
learning and research activities. ">
<META NAME="DC.Publisher" LANG="en" CONTENT="UKOLN">
<META NAME="DC.Publisher" LANG="en" CONTENT="University of Bath">
<META NAME="DC.Date" LANG="en" CONTENT="2003-07-03">
<META NAME="DC.Type" LANG="en" CONTENT="text">
<META NAME="DC.Type" LANG="en" CONTENT="image">
<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">

```

```

<META NAME="DC.Title" LANG="en" CONTENT="ARL 226: Institutional Repositories:
Essential Infrastructure for Scholarship in the Digital Age">
<META NAME="DC.Creator" LANG="en" CONTENT="Clifford A. Lynch">
<META NAME="DC.Subject" LANG="en" CONTENT="institutional repositories ">
<META NAME="DC.Subject" LANG="en" CONTENT="scholarly publishing ">
<META NAME="DC.Subject" LANG="en" CONTENT="digital preservation ">
<META NAME="DC.Subject" LANG="en" CONTENT="scholarly communication ">
<META NAME="DC.Subject" LANG="en" CONTENT="DSpace institutional repository
system">
<META NAME="DC.Description" LANG="en" CONTENT="This web page discusses the
development of institutional repositories emerged as a new strategy that allows universities to
apply serious, systematic leverage to accelerate changes taking place in scholarship and scholarly
communication.">
<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.Type" LANG="en" CONTENT="image">
<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.Rights" LANG="en" CONTENT="ARL policy is to grant blanket
permission to reprint any article in the Bimonthly Report for educational use as long as full
attribution is made. Exceptions to this policy may be noted for certain articles. This is in addition
to the rights provided under sections 107 and 108 of the Copyright Act. For commercial use, a
reprint request should be sent to ARL Director of Information Services, Julia Blixrud. ">
```

```
<META NAME="DC.Title" LANG="en" CONTENT="METS: An Overview & Tutorial">
<META NAME="DC.Subject" LANG="en" CONTENT="Metadata Encoding and Transmission
Standard (METS)">
<META NAME="DC.Description" LANG="en" CONTENT="Discusses how the Metadata
Encoding and Transmission Standard can be used 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="The Library Of Congress">
<META NAME="DC.Type" LANG="en" CONTENT="text">
<META NAME="DC.Type" LANG="en" CONTENT="image">
<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

While the process of record creation is certainly worth mentioning, one must first talk about the catalyst by which the user is able to generate these records. The Generic Metadata Editor is an extremely useful tool for the user that has no prior experience in record creation. There is however one major flaw when using the ERC scheme and that is the exclusion of “erc:” at the very top of the record. This is easily remedied by manually adding in the information, but for the inexperienced user, this discrepancy might go completely overlooked, resulting in an incomplete record.

There are several features of the Generic Metadata Editor that prove to be quite useful. All of the pop-up boxes provided for the purpose of entering data for a specific element in the DC metadata scheme either contain a brief definition of the element, or a description of how best to enter the information. For example, the pop up box used for the Date element provides the correct format for entering in this information. On occasion, the most commonly entered data is provided by default, such as in the case with “text/html” for the Type element, “English” for the Language element and “text” for the Type element. The pop up boxes for the Type, Relation, and Language elements also offer a drop down menu with a selection based on the most frequently used data. The pop up box for the Relation element even goes so far as to remind the user to enter in a phrase after using the drop down menu if applicable. All of these features greatly aid the user when selecting data for entry into a particular DC element.

While only the anchoring story was used in the creation of all of the ERC records, there were some challenges. In the first record for the article entitled “Metadata Demystified: A Guide for Publishers,” there were several authors listed. According to Kunze, a vertical line can be used to

separate different values for the same element. Since all of these individuals contributed to the creation of this text, it was important that they be entered into this element, and a series of vertical lines were used to separate each individual name. The title information given at the top of the document was not the full title article. Scrolling down to the table of contents, the full title could be seen. This did cause some confusion when this record was being created, so when the method for discovering the full title of the article for the DC record (this will be discussed later) was used, the results were added to the ERC records. The most difficult decision was in determining which date to enter for the When element. There were two choices. The first was the copyright date of 2003 and the second was the printed date of July, 2003. Since the copyright date is the most recognized date for any item, 2003 was entered into this element.

One of the many limitations of the ERC metadata scheme is the small number of elements. There is no element for contributors of a document. There were several contributors listed for the document entitled “DC-Library Application Profile (DC-Lib),” immediately after the creator information is given. This document presented a challenge regarding its current URL. The URL given to locate the document for this assignment and the URL posted next to the Identifier heading at the top of the document are completely different. The question posed was which URL should be used. Under the Identifier heading was a Replaces heading with yet another URL. In the end, the URL next to the Identifier heading was used for the Where element since the document recognizes that URL as its most current location marker.

The last three items presented challenges in the location of certain pieces of information. The document entitled “JISC Information Environment Architecture - Distributed Systems – UKOLN,” is a website that offered no information as to who was responsible for the content. The webmaster was identified and it was used for the Who element due to his involvement with the web page. Whether that involvement included the content or not is unknown. The copyright date for the image used in the document “ARL 226: Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age” was used and entered into the When element, because the date for the document itself was unavailable. As mentioned before, it is common to locate a document using the copyright date (if a date is going to be used), and therefore it was deemed appropriate to use this particular date. Ambiguity played a part in the document entitled “METS: An Overview & Tutorial.” The only date available was on the bottom of the document in parenthesis next to a link for the Library of Congress Help Desk. It was difficult to ascertain whether this date represented the document or the most recent update of that link. As a result, “(:unap)” was used since an ERC element cannot be left blank.

Creating records for these items is much easier using the DC metadata scheme. There are several reasons for this. The larger number of elements allow for more information to be used to represent an item. Each element is clearly defined and guidelines are given as how to properly enter the data. While there is still some ambiguity in some of these elements, there is a greater amount of specificity that surpasses that of the ERC. Another advantage is that not all elements need to be used. Any elements not used need not be put into the record. This eliminates having to put in any abbreviation to explain the absence of any solid data.

Going back to “Metadata Demystified: A Guide for Publishers.” When creating a record for this article using ERC, the title was originally thought to be “Metadata Demystified” until scrolling

down to the table of contents. There was some uncertainty as to the actual title, until reading the guideline for determining the title for a DC record. According to the guidelines, if an item is in HTML, the source is viewed and use the title included in the title header is used for the Title element. Realizing the full title of the document, the ERC record was changed to match that of the Title element in the DC record for the sake of full representation. This was done on all ERC records for this assignment.

Another advantage of the DC scheme not previously mentioned is that elements may be repeated as many times as necessary. For example, there are several contributors for “DC-Library Application Profile (DC-Lib).” Instead of using a vertical line, a Contributor element was used for each individual person. It is this user’s opinion that this makes for easier reading of the record.

Two of the most confusing elements of the DC metadata scheme are Source and Relation. This is because while the element names differ in their descriptions, the method of referencing a resource by means of a string or number conforming to a formal identification system are quite similar. Another problem is that in some cases DC qualifiers are required to make the best use of the Relation element. There were two instances when both of these elements were considered.

The document “Metadata Demystified: A Guide for Publishers,” is a scanned version of a printed resource. The ISBN number of the printed resource is provided within the document. The ISBN number represents the original source of this document and this number is included in the Source element. Since there is a distinct relation between the two, the Relation element is used as well. As mentioned before, when entering for this element, a drop down menu is provided. Looking at the options to best represent the relationship to the original document, "IsVersionOf " is selected.

“DC-Library Application Profile (DC-Lib),” is a revision of a previous article. Unfortunately, there is no ISBN number to link to the original document. Instead, a URL of the previous article is provided. Therefore, the previous URL is entered into the Source element. Since simple DC is being used here, according to the guideline the refinement labels within the drop down menu may be used. This document is an obvious revision of the previous article, but this relationship is not available in the menu. According to the guideline for the Source element, information about a previous version which does not fit easily into the Relation element is contained here. Since this record is not expanding beyond the simple DC structure, this information stays within the Source element.

Beyond the advantages and disadvantages of the General Metadata Editor, the ERC and DC metadata schemes, the location of pertinent information, and the decision making process regarding reference entry, is the problem of the quality of available information. No other document better highlights the problem than “METS: An Overview & Tutorial.” This is the only document that did not contain any information about the creator of the item. As a result of this, The Library of Congress was entered for the Who element in the ERC record, and for the Publisher Element in the DC since the nature of responsibility is ambiguous and an organization is involved. The unclear meaning of the date was already mentioned in the discussion about the ERC record for this time. Again, because not every item must be used for the DC scheme, the Date element was not included. This lack of information results in an incomplete record and an

item that cannot be located via an author or date search. It is unclear and frustrating to this user why those responsible for this document felt it was not important to include this crucial information for the benefit of anyone interested in locating this document.

### **Summary and Conclusion**

This has been an informative learning experience in creating ERC and DC metadata records. The Generic Metadata Editor is an invaluable tool for this purpose. While there are pro and cons to each metadata scheme, both may be used to form a complete record as long as there is a clear understanding of the definition and scope of each element and the individual creating the record reads and follows the guidelines provided. The only obstacle in the way of complete representation is the lack of pertinent information in a document.