

# Metadata Records Report

---

## Introduction

This paper is the result of an exercise in creating metadata records using the Electronic Resource Citation (ERC) metadata scheme as well as the Dublin Core (DC) scheme with NoteTab Light version 4.95 text editing tool. Five document objects were given to create these metadata records out of each; therefore, ten records exist for discovery learning which is discussed in the Discussion section of this document along with a Summary and Conclusion section of my experiences creating these records.

## Electronic Resource Citation Records

### ERC Record 1

erc:

who: Murtha Baca | Tony Gill | Anne Gilliland-Swetland | Mary Woodley | J. Paul Getty Trust

what: Introduction to Metadata: Pathways to Digital Information

when: 1998

where: Los Angeles

### ERC Record 2

erc:

who: Rachel Heery | Manjula Patel | Ariadne Issue 25

what: Application Profiles: Mixing and Matching Metadata Schemas

when: September 24, 2000

where: <http://www.ariadne.ac.uk/issue25/app-profiles/intro.html>

### ERC Record 3

erc:

who: Kopiaosto | CAL | Editeur | IFPI | Muze Inc. | IDF | <Indecs> Framework Ltd

what: Interoperability of Data in e-Commerce Systems: Putting Metadata to Rights

when: 1998

where: <http://www.indec.org/>

### ERC Record 4

erc:

who: CARL | Canadian Association of Research Libraries

what: What's New at CARL

when: 2003

where: Canada

### ERC Record 5

erc:

who: Jane Greenberg | Abe Crystal | W. Davenport Robertson | Ellen Leadem

what: Iterative Design of Metadata Creation Tools for Resource Authors

when: 2003

where: [http://www.siderean.com/dc2003/202\\_Paper82-color-NEW.pdf](http://www.siderean.com/dc2003/202_Paper82-color-NEW.pdf)

## **Dublin Core Records**

### DC Record 1

<META NAME="DC.Title" LANG="en" CONTENT="Introduction to Metadata: Pathways to Digital Information">

<META NAME="DC.Creator" LANG="en" CONTENT="J. Paul Getty Trust">

<META NAME="DC.Subject" LANG="en" CONTENT="metadata scheme">

<META NAME="DC.Subject" LANG="en" CONTENT="metadata crosswalk">

<META NAME="DC.Subject" LANG="en" CONTENT="metadata networks">

<META NAME="DC.Subject" LANG="en" CONTENT="taxonomy">

<META NAME="DC.Subject" LANG="en" CONTENT="Internet metadata">

<META NAME="DC.Description" LANG="en" CONTENT="Introduction to metadata, contains crosswalks of metadata schemes, version 2">

<META NAME="DC.Publisher" LANG="en" CONTENT="Getty Research Institute">

<META NAME="DC.Contributor" LANG="en" CONTENT="Murtha Baca">

<META NAME="DC.Contributor" LANG="en" CONTENT="Tony Gill">

<META NAME="DC.Contributor" LANG="en" CONTENT="Anne Gilliland-Swetland">

<META NAME="DC.Contributor" LANG="en" CONTENT="Mary Woodley">

<META NAME="DC.Date" LANG="en" CONTENT="1998">

<META NAME="DC.Type" LANG="en" CONTENT="text">

<META NAME="DC.Type" LANG="en" CONTENT="Physical object">

<META NAME="DC.Format" LANG="en" CONTENT="Book">

<META NAME="DC.Format" LANG="en" CONTENT="7" x 9 1/2"">

<META NAME="DC.Format" LANG="en" CONTENT="47 pages">

<META NAME="DC.Identifier" LANG="en"

CONTENT="http://www.getty.edu/research/conducting\_research/standards/intrometadata/1\_introduction/index.html">

<META NAME="DC.Source" LANG="en" CONTENT="ISBN 0-89236-533-1 {where "ISBN 0-89236-533-1" is the ISBN number of the print version of the resource, from which the excerpt version was taken."}>

<META NAME="DC.Language" LANG="en" CONTENT="en">

<META NAME="DC.Relation" LANG="en" CONTENT="IsBasedOn ISBN 0-89236-533-1">

### DC Record 2

<META NAME="DC.Title" LANG="en" CONTENT="Application Profiles: Mixing and Matching Metadata Schemas">

<META NAME="DC.Creator" LANG="en" CONTENT="Ariadne">

<META NAME="DC.Subject" LANG="en" CONTENT="application profile">

<META NAME="DC.Subject" LANG="en" CONTENT="metadata">

<META NAME="DC.Subject" LANG="en" CONTENT="DESIRE project">

<META NAME="DC.Subject" LANG="en" CONTENT="Warwick Framework">

<META NAME="DC.Subject" LANG="en" CONTENT="Dublin Core">

<META NAME="DC.Description" LANG="en" CONTENT="this paper introduces application profiles as a metadata scheme">

<META NAME="DC.Publisher" LANG="en" CONTENT="Ariadne">

<META NAME="DC.Contributor" LANG="en" CONTENT="Rachel Heery">

<META NAME="DC.Contributor" LANG="en" CONTENT="Manjula Patel">

<META NAME="DC.Date" LANG="en" CONTENT="2000-09-24">

<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.ariadne.ac.uk/issue25/app-profiles/">

<META NAME="DC.Source" LANG="en" CONTENT="ISBN: 1361 3200; http://www.ariadne.ac.uk/issue25/ (where this article is contained.)">

<META NAME="DC.Language" LANG="en" CONTENT="en">

<META NAME="DC.Relation" LANG="en" CONTENT="IsPartOf http://www.ariadne.ac.uk/issue25/">

<META NAME="DC.Rights" LANG="en" CONTENT="Joint Information Systems Committee (JISC)">

<META NAME="DC.Rights" LANG="en" CONTENT="Electronic Libraries ('eLib) Programme">

### DC Record 3

<META NAME="DC.Title" LANG="en" CONTENT="Interoperability of Data in e-Commerce Systems: Putting Metadata to Rights">

<META NAME="DC.Creator" LANG="en" CONTENT="Jane Dorner">

<META NAME="DC.Subject" LANG="en" CONTENT="Standards">  
<META NAME="DC.Subject" LANG="en" CONTENT="Standardisation">  
<META NAME="DC.Subject" LANG="en" CONTENT="identifiers">  
<META NAME="DC.Subject" LANG="en" CONTENT="metadata">  
<META NAME="DC.Subject" LANG="en" CONTENT="rights management">  
<META NAME="DC.Subject" LANG="en" CONTENT="intellectual property rights">  
<META NAME="DC.Subject" LANG="en" CONTENT="IP rights">  
<META NAME="DC.Subject" LANG="en" CONTENT="CIS Plan">  
<META NAME="DC.Subject" LANG="en" CONTENT="DOI">  
<META NAME="DC.Subject" LANG="en" CONTENT="RDF">  
<META NAME="DC.Subject" LANG="en" CONTENT="Resource Description Framework">  
<META NAME="DC.Subject" LANG="en" CONTENT="XML">  
<META NAME="DC.Subject" LANG="en" CONTENT="resource description">  
<META NAME="DC.Subject" LANG="en" CONTENT="commerce model">  
<META NAME="DC.Subject" LANG="en" CONTENT="metadata registries">  
<META NAME="DC.Subject" LANG="en" CONTENT="INFO2000">  
<META NAME="DC.Subject" LANG="en" CONTENT="MMRCS">  
<META NAME="DC.Subject" LANG="en" CONTENT="digital libraries">  
<META NAME="DC.Description" LANG="en" CONTENT="An international initiative of rights owners creating metadata standards for e-commerce">  
<META NAME="DC.Publisher" LANG="en" CONTENT="Indecs Framework Ltd.">  
<META NAME="DC.Contributor" LANG="en" CONTENT="Kopiosto">  
<META NAME="DC.Contributor" LANG="en" CONTENT="CAL">  
<META NAME="DC.Contributor" LANG="en" CONTENT="Editeur">  
<META NAME="DC.Contributor" LANG="en" CONTENT="IFPI">  
<META NAME="DC.Contributor" LANG="en" CONTENT="Muze Inc.">  
<META NAME="DC.Contributor" LANG="en" CONTENT="IDF ">  
<META NAME="DC.Contributor" LANG="en" CONTENT="<indecs> Framework Ltd.">  
<META NAME="DC.Date" LANG="en" CONTENT="1998">  
<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.indecs.org/">  
<META NAME="DC.Source" LANG="en" CONTENT="<Indecs>">  
<META NAME="DC.Language" LANG="en" CONTENT="en">  
<META NAME="DC.Coverage" LANG="en" CONTENT="Europe">

DC Record 4

<META NAME="DC.Title" LANG="en" CONTENT="Canadian Association of Research Libraries">

<META NAME="DC.Subject" LANG="en" CONTENT="Library Research">

<META NAME="DC.Subject" LANG="en" CONTENT="Library Association">

<META NAME="DC.Description" LANG="en" CONTENT="This Website is a portal to access CARLs membership area and provide public information.">

<META NAME="DC.Publisher" LANG="en" CONTENT="Canadian Association of Research Libraries">

<META NAME="DC.Contributor" LANG="en" CONTENT="Katherine McColgan">

<META NAME="DC.Date" LANG="en" CONTENT="2005-02-05">

<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.carl-abrc.ca/frames\_index.htm">

<META NAME="DC.Language" LANG="en" CONTENT="en">

<META NAME="DC.Language" LANG="en" CONTENT="fr">

<META NAME="DC.Coverage" LANG="en" CONTENT="Canada">

<META NAME="DC.Rights" LANG="en" CONTENT="Copyright Forum">

<META NAME="DC.Rights" LANG="en" CONTENT="Paul Whitney">

DC Record 5

<META NAME="DC.Title" LANG="en" CONTENT="Iterative Design of Metadata Creation Tools for Resource Authors">

<META NAME="DC.Subject" LANG="en" CONTENT="metadata generation">

<META NAME="DC.Subject" LANG="en" CONTENT="metadata applications">

<META NAME="DC.Subject" LANG="en" CONTENT="Dublin Core">

<META NAME="DC.Subject" LANG="en" CONTENT="Author-generated metadata">

<META NAME="DC.Subject" LANG="en" CONTENT="NIEHS">

<META NAME="DC.Subject" LANG="en" CONTENT="National Institute of Environmental and Health Sciences">

<META NAME="DC.Subject" LANG="en" CONTENT="Metadata Generation Research Project">

<META NAME="DC.Subject" LANG="en" CONTENT="NIEHS metadata application version 1.0">

<META NAME="DC.Subject" LANG="en" CONTENT="NIEHS metadata application version 2.0">

<META NAME="DC.Description" LANG="en" CONTENT="This paper discusses the overview of author-generated metadata, evaluates the two versions of the NIEHS metadata generation tool, and summarizes its findings.">

```

<META NAME="DC.Publisher" LANG="en" CONTENT="National Institute of Environmental
Health Sciences">
<META NAME="DC.Contributor" LANG="en" CONTENT="Jane Greenberg ">
<META NAME="DC.Contributor" LANG="en" CONTENT="Abe Crystal">
<META NAME="DC.Contributor" LANG="en" CONTENT="W. Davenport Robertson">
<META NAME="DC.Contributor" LANG="en" CONTENT="Ellen Leadem">
<META NAME="DC.Date" LANG="en" CONTENT="2003">
<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.siderean.com/dc2003/202_Paper82-color-NEW.pdf">
<META NAME="DC.Language" LANG="en" CONTENT="en">

```

## Discussion

After reading the required readings, both the ERC and DC metadata schemes appeared to be well defined in how to input at least basic data.

At the start of entering ERC metadata scheme records, the code itself—erc:, was found not to appear in the document coding in the NotesTab Light program. NoteTab Light has other templates such as the HTML template addressing the document structure as a whole that is machine readable. The ERC template addressed the metadata scheme elements and not the metadata scheme record as a whole. This omission required corrective editing to obtain a machine readable record. Course reading material added needed assistance in structuring the complete scheme e.g. erc: who:.

Regarding the assignment of the ERC elements, the basic anchoring story was used to provide a bibliographic representation of the information object. All document records have the required *who*, *what*, *when*, *where* elements. Although both the ERC and DC metadata schemas are descriptive metadata, the ERC scheme almost requires an additional story segment for “erc-about:” to address intellectual content to help the end user to find real value.

The approach taken on this assignment was one of the layman metadata coder using simplified semantics and approach. Acquiring the element values was relatively easy. Since these metadata schemes are geared for documents, simply the content and source code were the main source for all values entered. The pipe character “|” was used to delineate multiple values. Although a total of four elements existed per story, the way in which the values may be assigned to each element varied greatly, thus potentially leading to seemingly intricate and multiple stories.

The DC metadata scheme record entry in NoteTab Light addressed all of the coding required for placement in an HTML header. The elements were descriptive, enabling easy completion of the values. An easy comparison between the elements added further differentiation for evaluating placement of values. And the ability to add an element multiple times or leave out elements as needed provided much flexibility in meeting metadata needs.

Also, the question of format came about when searching for the correct MIME type to address the pdf application. The application/pdf MIME type was used for this assignment. However, in truth different browsers rely on differing MIME types.

## Summary and Conclusion

Although the ERC scheme contains only four elements per story, the average layman assigning values and story segments may get worsted into the finer nuances of story segments and get lost. As long as a non-librarian input user can conceptually enter into a story segment and not break out its context, the ERC story segment can easily work. The average laymen generally stay away from coding activities; however, the syntax is very easy to manipulate. Also, the permanence factor of the ERC scheme is a tremendous benefit to those who search for information according to the values.

The story segments and element concepts can be vague. People have a tendency to stay away from things they don't understand and vagueness of definition has a disorienting and exposing feel to any activity associated with it. The ERC elements, although easily interpretable by the input user, is highly vague when attempting to assign extra values that need a different story segment. An ERC record almost requires consisting of two stories: one for bibliographic and one for intellectual capital. A person could add many story segments in an attempt to provide a surrogate for the document contents. The key is the depth of indexing one wishes to provide. Also, my personal experience with teaching non-librarian users in inputting information into distributed content management corporate portals (and assigning metadata) has shown me that many students shy away from placing anything onto a Web environment that can point back to them as having made a mistake. Interestingly, the ERC metadata scheme is like a facet classification with each classification having a story. One thinks of the story and then the elements. However, as one is focused, for example, on the *who* element, multiple types of "who's" become apparent. Therefore, a high level of conceptual filtering is required. The person must pan out conceptually and focus in alternately to avoid getting lost in the details—an activity which is not intuitive for many people.

When entering the ERC elements was complete, I considered that I liked that metadata scheme best. However, after entering the DC elements, I find that I liked entering DC records even better since the DC template entered the syntax code for me. I did not need to think about appropriate story segments, and I'm accustomed to seeing through programming code. When I entered an element's value such as "contributor", I was already in tune with the "who" concept from previously entering the ERC records. I could quickly scan the elements and determine the difference between "contributor" and "publisher" or "creator" and assign values accordingly.

Assigning the DC elements made me really think about the meaning of the values and the extent to which I wanted to index. The ERC scheme was not compulsive in that way. Any time a person can be placed in a position to differentiate, he can provide himself with stable datum for follow through. However, I don't think the layman would like the DC metadata scheme because of the librarian feel it has. When I teach my students in corporate portals, I am often confronted with the aspect that the input users may see themselves as owners of information but not see themselves as Web designer or librarian. These metadata schemes are, however, easy enough to learn if such information input users wish to "sell" a product by promoting their information via metadata in a registry. They will follow through or pay for someone to do it.

I particularly like the ability to map DC elements into the ERC coding using the slash character i.e. *who/contributor*:. The virtue of the ERC metadata scheme truly shows its value in not being tied to a semantic. Both schemes have their place. I would like to see a Web environment using the ERC code to see how it plays out as an end user.

Also, the chief source for MIME types is difficult to determine. I suppose application/msword would be acceptable. However, in my experience, if an input user posts a Microsoft Word document using Netscape, the MIME type used would be vnd.msword. The Internet Explorer

would use the application/msword. Alas, the browser platforms at war! It appears that Microsoft is even winning the chief source of information for format metadata.

Finally, I have some commentary about the intriguing dialogue passed on by student discussions, namely how others are thinking through this process of obtaining metadata in the source code of a Web page. What they are stumbling into is an exercise in document management whereby each version has its own record and apparently the DC code provides for this easily within the list of metadata. The keynote is "live document" which is actively owned by someone or entity or creator, etc. They are attempting to recreate a history through metadata that would almost require them to obtain that history by contacting the creator, author, etc. The DC scheme certainly serves this endeavor well being geared towards document records with fifteen elements to address. Again, that's why I believe the DC scheme gives the input user a compelling initiative to place these attributes into an element value that the ERC scheme does not compel. My previous understanding of a record is that it would create a record for each version such as a book published 1<sup>st</sup> version, 2<sup>nd</sup>, 3<sup>rd</sup>, etc. It would result in three records, each having its own set of metadata with most values being redundant as they are inherited. However, with capturing a live document-ouch! That's high maintenance but comprises a majority of the documents that will be captured.