



From *Metadata Fundamentals for All Librarians* by Priscilla Caplan.
Copyright © 2003 by the American Library Association. All rights reserved. Permission granted to reproduce for nonprofit, educational purposes.

[Check out the book at the ALA Online Store \(www.alastore.ala.org\).](http://www.alastore.ala.org)

8

The Dublin Core

The Dublin Core Metadata Element Set is a general-purpose scheme for resource description originally intended to facilitate discovery of information objects on the Web.

THE ELEMENT SET

The origin of the Dublin Core is by now nearly legendary. In the fall of 1994, the Second International World Wide Web Conference was held in Chicago with the theme “Mosaic and the Web.” Although most of the conference tracks addressed the potential of the infant Web for transforming the way knowledge is presented, many of the participants were concerned with how all this newly available content would be found. Three of the attendees—Stu Weibel of OCLC, Joseph Hardin of NCSA, and the late Yuri Rubinski of Softquad—took the initiative to convene a multidisciplinary workshop in March 1995 to address which descriptive data elements were essential for discovery of networked information resources. A stated goal of the meeting was “to achieve consensus on a core set of data elements for document analogs and explore elements for other network-specific object types.” Because the workshop was held at OCLC headquarters in Dublin, Ohio, the core set of data elements proposed there became known as the Dublin Core, and the workshop itself was retrospectively dubbed “DC1,” the first of an ongoing series of Dublin Core metadata workshops.

The development of official specifications related to the Dublin Core is managed by the Dublin Core Metadata Initiative (DCMI), which consists of a small, paid directorate advised by a board of trustees, and a large number of loosely organized volunteers. Over time the DCMI has developed a governance structure and formal procedures for the

approval of new specifications and the adoption of new terms. Most of the business of the DCMI is carried out in working groups, which provide a forum for discussion of specific issues and may draft requirements or specifications documents. Specifications follow a progression of statuses similar to that used by the World Wide Web Consortium (W3C), from “Working Draft,” through “Proposed Recommendation” to “Recommendation.” The status of Recommendation is equivalent to a standard; the specification is considered stable and supported for adoption by implementers. The approval of new metadata terms (elements or qualifiers) is the responsibility of a small, high-level committee called the Usage Board.

The Dublin Core Metadata Element Set (Dublin Core) itself consists of fifteen data elements. Identifiers and definitions of the elements are excerpted here from the reference definition of the Dublin Core:¹

Identifier: Title

Definition: A name given to the resource.

Identifier: Creator

Definition: An entity primarily responsible for making the content of the resource.

Identifier: Subject

Definition: The topic of the content of the resource.

Identifier: Description

Definition: An account of the content of the resource.

Identifier: Publisher

Definition: An entity responsible for making the resource available.

Identifier: Contributor

Definition: An entity responsible for making contributions to the content of the resource.

Identifier: Date

Definition: A date associated with an event in the life cycle of the resource.

Identifier: Type

Definition: The nature or genre of the content of the resource.

Identifier: Format

Definition: The physical or digital manifestation of the resource.

Identifier: Identifier

Definition: An unambiguous reference to the resource within a given context.

Identifier: Source

Definition: A reference to a resource from which the present resource is derived.

Identifier: Language

Definition: A language of the intellectual content of the resource.

Identifier: Relation

Definition: A reference to a related resource.

Identifier: Coverage

Definition: The extent or scope of the content of the resource.

Identifier: Rights

Definition: Information about rights held in and over the resource.

All elements are optional, and all elements are repeatable. The scheme itself is format-independent, meaning that it is not tied to any single data representation the way, for example, the TEI header is tied to SGML/XML. The scheme is also not tied to any particular set of content rules, although recommended best practice is noted in the comments attribute for some elements, and additional recommendations are given in an official usage guide.² Following these documents for recommended best practice, a simple Dublin Core description could look like this:

Title="The Electronic Text Center Introduction to TEI and Guide to Document Preparation"

Creator="Seaman, David"

Subject="Text Encoding Initiative"

Subject="SGML markup rules"

Description="Guidelines written by the University of Virginia Electronic Text Center for marking up electronic texts using the TEILITE.DTD, a subset of the TEI tagset."

Date="1995"

Type="text"

Language="en"

Identifier="http://etext.lib.virginia.edu/tei/uvatei.html"

Dublin Core Qualifiers is a companion specification to the Dublin Core Metadata Element Set.³ A *qualifier* either identifies the encoding scheme used in representing a Dublin Core element or refines the meaning of an element. An *encoding scheme* qualifier indicates the scheme or authority list used in representing the value of an element. An *element refinement* qualifier can narrow the meaning of an element but may not extend or change it. An important characteristic of element refinement qualifiers is that they can be ignored and the meaning of the value of the element will still make sense. This requirement, also known as the "dumb down principle," is based on the realization that not all applications processing Dublin Core metadata will necessarily recognize all qualifiers, so it must be possible to "dumb down" to the basic, unqualified meaning of the element.

Qualifiers are specific to individual elements. The element Date, for example, has five approved element refinement qualifiers (Created, Valid, Available, Issued, Modified) and two approved encoding scheme qualifiers (DCMI period, and W3C-DTF). The element Title has one element refinement qualifier (Alternative) and no encoding scheme qualifiers. Qualifiers have their own definitions that may reference other specifications or authority lists.

Despite the simplicity of the Dublin Core scheme, certain problems have arisen repeatedly in applications. One issue concerns the overlap in meaning in the definition of some elements. Creator can be seen as a particular type of Contributor, and Source

is a particular type of Relation. This has led to confusion among implementers about when it is appropriate to use one element rather than another. At one point, a proposal to combine the elements Creator, Contributor, and Publisher into a single element called “agent” was considered and rejected. It has also been suggested that use of Source be deprecated in favor of Relation. However, the reasons for including Creator and Source in the original specification remain valid to many implementers. The bibliographic community has always accorded authorship special status, as reflected in the AACR2 concept of main entry. Distinguishing the special role of Creator from other contributors can make logical or practical sense in some applications. Along the same line, an important use of Dublin Core is to describe electronic versions of resources created by retrospective conversion projects. A special element for recording the nondigital source of the electronic resource can be justified in this context.

A second persistent issue concerns the nature of description when multiple versions exist. For example, the name of the photographer will generally be recorded as the creator of a photograph. However, if the photograph has been digitized and exists as a JPEG image, it can be argued that the person who scanned the photo is the creator of the image. Some Dublin Core implementers feel that the scanning technician is intellectually meaningless and should be recorded, if at all, as a contributor. Others believe that a Dublin Core record should accurately describe the resource in hand, which implies that for the JPEG image, the scanner is the creator and the photographer is at best a contributor. This principle, known as “one-to-one,” prescribes that if multiple versions of a resource exist, each should be separately and accurately described.

SYNTAXES FOR DUBLIN CORE

Although the two Recommendations defining the Dublin Core and the Dublin Core Qualifiers are meant to convey semantics only, for a metadata scheme to be usable in practice it must have one or more generally accepted syntactical representations. The first encoding specification to reach Recommendation status was for HTML.⁴ This specification makes use of <meta> “name” and “content” attributes, in the generic format:

```
<meta name = “PREFIX.Element_name”
      content = “element_value”>
```

The prefix is arbitrary and used to link to the Dublin Core specification. This is represented in HTML by a set of attributes to the <link> element:

```
<link rel=“schema.PREFIX”
      href=“http://purl.org/dc/elements/1.1/”
      title=“Dublin Core Metadata Element Set, Version 1.1”>
```

The HTML Recommendation specifies using the capitalized “DC” as the prefix. A portion of the Seaman document description shown above could be represented in the following HTML:

```

<link rel="schema.DC"
      href="http://purl.org/dc/elements/1.1/"
      title="Dublin Core Metadata Element Set, Version 1.1">
<meta name="DC.Title"
      content="The Electronic Text Center Introduction to TEI and Guide to
      Document Preparation">
<meta name="DC.Creator"
      content="Seaman, David">
<meta name="DC.Identifier"
      content="http://etext.lib.virginia.edu/tei/uvatei.html">

```

Encoding scheme qualification is represented by use of the <meta> “scheme” attribute:

```

<meta name="DC.Type"
      scheme="DCMIType"
      content="text">

```

Element refinement qualification is represented in “dot” notation:

```

<meta name="DC.Date.created"
      content="1995">

```

Representing Dublin Core semantics in HTML is fairly straightforward and works particularly well in environments where web pages are spidered and indexed by search engines configured to take advantage of <meta> tags. There are, however, some limitations and drawbacks. For applications that require metadata records, as opposed to metadata embedded in documents, XML tends to be the preferred exchange syntax. Also, HTML cannot represent more complex constructions—for example, where sets of repeated elements need to be grouped to be meaningful.

Dublin Core can also be represented in XML. Several XML schemas have been developed for particular applications of Dublin Core, including one approved for use with Open Archives Initiative metadata harvesting applications. The DCMI home page links to a list of schemas that are supported by the Dublin Core community. In addition, general guidelines for representing both qualified and unqualified Dublin Core in XML have been issued by UKOLN.⁵ UKOLN recommends that implementers make use of the XML namespace facility to uniquely identify Dublin Core elements, which should be represented as XML elements. The Seaman document represented in simple Dublin Core according to the UKOLN specification might look like this:

```

<?xml version="1.0"?>
<metadata
      xmlns="http://myorg.org/myapp/"

```

```

xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://myorg.org/myapp/ http://myorg.org/myapp/
schema.xsd"
xmlns:dc="http://purl.org/dc/elements/1.1/">
<dc:title> The Electronic Text Center Introduction to TEI and Guide to
Document Preparation </dc:title>
<dc:creator> Seaman, David </dc:creator>
<dc:identifier> http://etext.lib.virginia.edu/tei/uvatei.html </dc:identifier>
</metadata>

```

In this example, the XML schema used is a (fictitious) schema referenced by the URI “http://myorg.org/myapp/schema.xsd.” The XML namespace for the Dublin Core metadata element set itself is referenced with the namespace statement that begins “xmlns:dc=” indicating that Dublin Core element names will be prefaced by “dc:” and that the definition of DC elements will be found in the document at http://purl.org/dc/elements/1.1/. The UKOLN specification recommends representing Dublin Core element names (property names) in lowercase (that is “dc:title” rather than “dc:Title”).

To encode qualified Dublin Core, a namespace statement for the reference definition of the Dublin Core Qualifiers must be added, shown in the following example as “xmlns:dcterms=” UKOLN recommends representing element refinement qualifiers as elements rather than as attributes, so that, for example, the Date qualifier “Created” would be represented as:

```
<dcterms:created>2002</dcterms:created>
```

rather than

```
<dc:date type="created">2002</dc:date>.
```

In contrast, encoding scheme qualifiers should be represented using a “scheme” attribute, and the language of a value should be represented using the XML “lang” attribute.

```

<?xml version="1.0"?>
<metadata
xmlns="http://myorg.org/myapp/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://myorg.org/myapp/ http://myorg.org/myapp/
schema.xsd"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:dcterms="http://purl.org/dc/terms">

```

```

<dc:title xml:lang="en"> The Electronic Text Center Introduction to TEI
and Guide to Document Preparation </dc:title>
<dc:creator> Seaman, David </dc:creator>
<dcterms:created> 1995 </dcterms:created>
<dc:identifier scheme="URI"> http://etext.lib.virginia.edu/tei/uvatei.html
</dc:identifier>
</metadata>

```

Dublin Core can also be represented in XML according to the rules of the Resource Description Framework (RDF). The Recommendation, “Expressing Simple Dublin Core in RDF/XML,” was approved by the DCMI in October 2002 (<http://www.dublincore.org/documents/2002/07/31/dcmes-xml/>). According to the Recommendation, the use of RDF must be declared with an `<rdf:RDF>` tag. A single RDF encoding can be used to represent multiple resources, as long as each resource to be described is encapsulated within a separate `<rdf:Description>` element. No qualifiers or locally defined elements can be used, and the resulting RDF/XML cannot be embedded in web pages.

The actual encoding of the Dublin Core elements is quite straightforward, as the example below shows.

```

<?xml version="1.0"?>
<!DOCTYPE rdf:RDF PUBLIC "-//DUBLIN CORE//DCMES DTD
2002/07/31//EN" "http://dublincore.org/documents/2002/07/31/dcmes-
xml/dcmes-xml-dtd.dtd">
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
<rdf:Description rdf:about="http://etext.lib.virginia.edu/tei/uvatei.html">
<dc:title>The Electronic Text Center Introduction to TEI and Guide
to Document Preparation</dc:title>
<dc:creator>Seaman, David</dc:creator>
<dc:date>2002-07-31</dc:date>
</rdf:Description>
</rdf:RDF>

```

Note that if the resource has a single URI, it is encoded as the value of the `rdf:about` attribute, rather than as the value of a `<dc:identifier>` element. If the resource has multiple URIs, the additional URI(s) may be given in `<dc:identifier>`.

There is no approved Recommendation for expressing qualified Dublin Core in RDF/XML, but a proposed Recommendation is working its way through the approval process.⁶ According to this document, the Dublin Core elements and element refinement qualifiers correspond to RDF properties and subproperties. Encoding scheme qualifiers, on the other hand, correspond to RDF “classes” or “types.” In the following

example, which follows the proposed Recommendation, two assertions are made. First, the value of the element refinement qualifier “created” is stated to be 1995. Second, “created” itself is noted to be a subproperty of the Dublin Core element Date.

```
<rdf:Description>
  <dcq:created>1995</dcq:created>
</rdf:Description>
<rdf:Description about="http://purl.org/dc/terms/created">
  <rdfs:subPropertyOf
    rdf:resource="http://purl.org/dc/elements/1.1/date"/>
</rdf:Description>
```

Another syntactical requirement, beyond the ability to represent simple and qualified Dublin Core, is some mechanism for combining elements from Dublin Core and other defined metadata element sets. From the beginning, implementers realized that the Dublin Core would have to be augmented by additional elements to be useful in specific application areas or domains. The element prefix serves this function by indicating the scheme from which an element was taken.

A good example of this is provided by a project called BIBLINK, funded by the European Commission. BIBLINK was designed to encourage publishers to contribute standard descriptive metadata for electronic documents to national bibliographic services that would in turn send enhanced metadata back to the publishers. BIBLINK defined a metadata element set with nineteen elements.⁷ Twelve of these were taken from the Dublin Core, and seven were defined specifically for BIBLINK, including a checksum, place of publication, frequency, and price. A BIBLINK-compliant description in HTML uses prefixes to distinguish standard Dublin Core from BIBLINK-specific elements:

```
<meta name="BIBLINK.Checksum"
content="fd66e37fb693491e84e184b092121265">
<meta name="DC.Title" content="Taylor-Schechter Unit Home Page">
```

The use of the namespace facility provides a more formal mechanism for extensibility in XML and RDF. (Note, however, that XML DTDs do not explicitly support namespaces, so use of XML schema is preferred for document definition.) The following example, taken from the UKOLN “Guidelines for Implementing Dublin Core in XML,” shows a record that includes both Dublin Core elements and the IEEE Learning Object Model (LOM) element “TypicalLearningTime.”

```
<?xml version="1.0"?>
<record
  xmlns="http://myorg.org/learningapp/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
```



```

xsi:schemaLocation="http://myorg.org/learningapp/ http://myorg.org/
learningapp/schema.xsd"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:ims="http://www.imsglobal.org/xsd/imsmd_v1p2">
<dc:title> Frog maths </dc:title>
<dc:identifier> http://somewhere.com/frogmaths/ </dc:identifier>
<dc:description> Simple math games for 5-7 year olds. </dc:description>
<ims:typicallearningtime>
    <ims:datetime> 0000-00-00T00:15 </ims:datetime>
</ims:typicallearningtime>
</record>

```

APPLICATION PROFILES

As demonstrated by the preceding BIBLINK example, when the Dublin Core is used to describe resources for a particular project or application, it is not uncommon for implementers to supplement it with additional elements or qualifiers needed by that application. Implementers may also feel the need for stricter limitations on usage (for example, to define some required elements) or more specific guidelines on content than appear in the Dublin Core itself.

Application profiles are one way to formalize the definition of metadata schemes based on Dublin Core. Formally, an application profile is a scheme designed for a particular application that consists of data elements from one or more previously defined schemes. It can refine the meaning of existing elements, but it cannot expand the meaning of elements or introduce new elements. It can also specify limits on the use of elements, such as mandating conditions of use (e.g., mandatory, nonrepeatable) or specifying permitted or required data representations or controlled vocabularies.

Application profiles are best implemented as XML schema, as namespaces are supported, and XML schemas support local usage constraints, such as authority lists of values, required elements, and limitations on repeatability. Application profiles can also be implemented in RDF with slightly less flexibility. However, conceptually, application profiles can be established as written implementers' agreements and encoded in any syntax, so long as machine-understandability and technical enforcement are not required. In some communities, the idea of the application profile is being expanded to include the type of information that would commonly appear in a user guide, including more guidelines for choice and form of content than even XML schema language can enforce.

The BIBLINK scheme mentioned earlier is an example of an application profile and has been represented as an XML schema for this purpose.⁸ Some application profiles are being developed under the auspices of the DCMI, such as the Libraries application profile being developed by the Libraries Working Group. This profile is being developed

to support library applications of the Dublin Core, such as use as an interchange format between systems using different metadata standards; use in metadata harvesting applications, such as those following the Open Archives Initiative Metadata Harvesting Protocol; and use in creation of simple library catalog records with Dublin Core semantics. Other application profiles are under development within the DCMI for government, education, and environmental domains.

USES AND ISSUES

With or without the addition of domain-specific terms, the Dublin Core has proven useful in several library contexts. It is often used in subject gateways or portals, where the description of a resource appearing in the web gateway is generated from a database of brief Dublin Core information. It is also popular for describing electronic texts and images created in retrospective digitization projects, particularly those involving large numbers of items and in which full library cataloging may not be affordable or warranted. Also, for certain types of materials, such as photographs or newspaper articles, the application of AACR2 rules may be problematic, discouraging the integration of these items into the main library catalog. Here, use of Dublin Core-based schemes allows the advantage of some standardization while giving project designers the leeway to identify data elements and guidelines that are meaningful to them. For example, a project could decide to use AACR2 rules and associated authority files only for name headings in the Creator and Contributor fields, and not for formulating titles or for other aspects of bibliographic description.

Another common situation occurs when metadata is stored in the local database according to some richer scheme, but is converted to Dublin Core for use in a union catalog, Internet search engine index, or other external database containing contributions from multiple sources. The Dublin Core serves as a least common denominator to which more complex schemes can be mapped, so that searching can take place over a consistent set of data elements. The prime example of this is the Open Archives Initiative protocol for metadata harvesting, which requires that, at a minimum, all participating sites have the ability to export unqualified Dublin Core.

It should be noted that the long-term significance of the Dublin Core may lie less in its utility as a resource description scheme than in the role of the DCMI in bringing together so many disparate communities of interest. The DCMI has created an organization that is truly international in scope and participation, and it has brought to the fore issues related to language and multilingual representation of both metadata and metadata schemes. DCMI workshops have created a venue for libraries, museums, and other cultural heritage institutions to exchange information with governmental organizations, scientific agencies, web developers, computer scientists, educators, and others, enriching all these communities.

The DCMI has also played an important role in making the library community aware of interoperability issues beyond the closed MARC environment. From developing an early theoretical architecture for combining metadata from diverse schema (the “Warwick Framework”) to current use of XML and RDF namespaces, researchers associated with the Dublin Core have always acknowledged the real-world need to integrate descriptive and administrative metadata originating from different sources at different

times. Crosswalks to Dublin Core have been developed from nearly all important descriptive metadata schemes. Various prototypes developed for the DCMI registry effort have attempted to incorporate terms from application profiles and related domain-specific schemes as well as from the official Dublin Core namespaces. The heavy involvement of some members of the DCMI in the development of RDF and the Semantic Web has helped to promote awareness of these initiatives within the library community. The DCMI itself has deliberately broadened its mission in recent years, seeking to become a general forum for issues related to cross-domain discovery and frameworks for interoperability.

At the same time, the DCMI has been criticized for taking too long to produce basic guidance for Dublin Core implementers. There are still no approved Recommendations for syntactical representation of qualified Dublin Core in XML and RDF. Guidelines for representing citations to journal articles in Dublin Core have been under development since 1998 and are still unfinished. Element refinement qualifiers for the Creator, Contributor, and Publisher elements were omitted from the Dublin Core Qualifiers because of lack of consensus within the Usage Board, and currently are still pending, despite a great need for these among implementers. It remains to be seen whether the needs of implementers will be satisfied more or less well as the interests of the DCMI focus increasingly on theoretical and practical issues of interoperability.

NOTES

1. *Dublin Core Metadata Element Set, Version 1.1: Reference Description*, available at <http://dublincore.org/documents/dces/>. Accessed 21 June 2002.
2. Diane Hillmann, *Using Dublin Core*, available at <http://dublincore.org/documents/2001/04/12/usageguide/>. Accessed 21 June 2002.
3. *Dublin Core Qualifiers*, available at <http://dublincore.org/documents/dcmes-qualifiers/>. Accessed 21 June 2002.
4. John Kunze, *Encoding Dublin Core in HTML* (IETF RFC2731), available at <http://www.ietf.org/rfc/rfc2731.txt>. Accessed 21 June 2002.
5. Andy Powell and Pete Johnson, "Guidelines for Implementing Dublin Core in XML," 12 March 2002, available at <http://www.ukoln.ac.uk/metadata/dcmi/dc-xml-guidelines/>. Accessed 21 June 2002.
6. Stefan Kokkelink and Roland Schwänzl, "Expressing Qualified Dublin Core in RDF/XML," 29 August 2001, available at <http://www.dublincore.org/documents/2001/08/29/dcq-rdf-xml/>. Accessed 24 June 2002.
7. "BIBLINK Core Field Semantics," available at <http://hosted.ukoln.ac.uk/biblink/wp8/fs/bc-semantics.html>. Accessed 24 June 2002.
8. Jane Hunter, *An XML Schema Approach to Application Profiles*, 3 October 2000, available at http://archive.dstc.edu.au/maenad/appln_profiles.html. Accessed 24 June 2002.

READINGS

Dekkers, Makx, and Stuart L. Weibel. "Dublin Core Metadata Initiative Progress Report and Workplan for 2002." *D-Lib Magazine* 8, no. 2 (February 2002). Available at <http://www.dlib.org/dlib/february02/weibel/02weibel.html>.

A "state of Dublin Core" report from 2002. *D-Lib Magazine* has published numerous articles on Dublin Core, including reports from most of the Dublin Core workshops. These reports, read in order, constitute a history of the evolution of the DCMI.

Dublin Core Metadata Initiative (home page). Available at <http://www.dublincore.org/>.

Many relevant documents are linked to from the Dublin Core home page, including the current reference description of the Dublin Core specification itself (<http://www.dublincore.org/documents/dces/>) and the Dublin Core qualifiers (<http://www.dublincore.org/documents/dcmes-qualifiers/>). Approved DCMI Recommendations and proposed Recommendations are listed under the link "Documents." The "Resources" link includes a bibliography of writings related to Dublin Core.

Guenther, Rebecca, and Priscilla Caplan. "Metadata for Internet Resources: The Dublin Core Metadata Elements Set and Its Mapping to USMARC." *Cataloging and Classification Quarterly* 22, no. 3/4 (1996).

A discussion of some of the issues raised in attempting to map from Dublin Core to MARC. The Library of Congress maintains an official crosswalk from MARC to Dublin Core at <http://lweb.loc.gov/marc/marc2dc.html> and between Dublin Core, MARC21, and GILS at <http://www.loc.gov/marc/dccross.html>.