The Dublin Core Metadata Element Set

Abstract: Defines fifteen metadata elements for resource description in a cross-disciplinary information environment.

Contents

Foreword

The Dublin Core Metadata Initiative (DCMI) began in 1995 with an invitational workshop in Dublin, Ohio that brought together librarians, digital library researchers, content providers, and text-markup experts to improve discovery standards for information resources. The original Dublin Core emerged as a small set of descriptors that quickly drew global interest from a wide variety of information providers in the arts, sciences, education, business, and government sectors.

Since the original workshop there has been steadily growing interest in resource descriptions that are easy to create and that almost anyone can understand. The potential to increase visibility of resources in a collection across sectors and subject domains, and to do so at low cost, is broadly appealing. Services needing semantically rich descriptions would continue to provide them, but would attract cross-disciplinary discovery by also providing universally understandable descriptions common across disciplines. The digital tourist metaphor is apt. Internet travellers seeking information in foreign disciplines can use the Dublin Core's constrained vocabulary to obtain basic guidance in a language that they understand. Full accessibility to the culture and its services still requires mastery of the local vocabulary and environment, but a set of simple facts inscribed in Dublin Core can bring to the tourist's attention a foreign information portal that might otherwise have escaped notice.

The interest in cross-domain discovery fueled growing participation in a series of subsequent DCMI workshops. The Dublin Core metadata element set described here is a set of 15 descriptors that resulted from this effort in interdisciplinary and international consensus building. As of June 2000 the Dublin Core exists in over 20 translations, has been adopted by CEN/ISSS (European Committee for Standardization / Information Society Standardization System), and is documented in two internet RFCs (Requests for Comments). It also has official standing within the WWW Consortium and the Z39.50 standard. Dublin Core metadata is endorsed formally by governments in three countries for promoting discovery of government information in electronic form, and Dublin Core is under consideration as a national information standard in at least five others.

The Dublin Core is not intended to displace any other metadata standard. Rather it is intended to co-exist — often in the same resource description — with metadata standards that offer other semantics. It is fully expected that descriptive records will contain a mix of elements drawn from various metadata standards, both simple and complex. Examples of this kind of mixing and of HTML encoding of Dublin Core in general are given in RFC 2731 [RFC2731].

The simplicity of Dublin Core can be both a strength and a weakness. Simplicity lowers the cost of creating metadata and promotes interoperability. On the other hand, simplicity does not accommodate the semantic and functional richness supported by complex metadata schemes. In effect, the Dublin Core element set trades richness for wide

visibility. The design of Dublin Core mitigates this loss by encouraging the use of richer metadata schemes in combination with Dublin Core. Richer schemes can also be mapped to Dublin Core for export or for cross-system searching. Conversely, simple Dublin Core records can be used as a starting point for the creation of more complex descriptions.

This standard was processed and approved for submittal to ANSI by the National Information Standards Organization. It was balloted by the NISO Voting Members July 1, 2000 - August 15, 2000. It will next be reviewed in 2006. Suggestions for improving this standard are welcome. They should be sent to the National Information Standards Organization, 4733 Bethesda Avenue, Suite 300, Bethesda, MD 20814. NISO approval of this standard does not imply that all Voting Members voted for its approval. At the time it approved this standard, NISO had the following Voting Members:

NISO Voting Members

3M
Jerry Karel
Carolyn Egeberg (Alt)

Academic Press
Anthony Ross
Bradford Terry(Alt)

ALSi
Martin Sach
Tony O'Brien (Alt)

American Association of Law Libraries
Robert L. Oakley
Mary Alice Baish (Alt)

American Chemical Society
Robert S. Tannehill, Jr.

American Library Association
Paul J. Weiss

American Society for Information Science
Mark H. Needleman

American Society of Indexers
Charlotte Skuster
Marie Kascus (Alt)

American Theological Library Association
Myron Chace

ARMA International
Diane Carlisle

Art Libraries Society of North America
David L. Austin

Association for Information and Image Management
Betsy A. Fanning

Association of Information and Dissemination Centers
Bruce H. Kiesel

Association of Jewish Libraries
Pearl Berger
Elizabeth Vernon (Alt)

Association of Research Libraries
Duane E. Webster
Julia Blixrud (Alt)

Baker & Taylor
Robert H. Doran

Book Industry Communication
Brian Green

Broadcast Music Inc.
Edward Oshanani
Robert Barone (Alt)

Cambridge Scientific Abstracts
Matthew Dunie
Anthea Gotto (Alt)

CARL Corporation
Ward Shaw

College Center for Library Automation
J. Richard Madaus
Ann Armbrister (Alt)

Committee on Institutional Cooperation
Thomas Peters

Congressional Information Service, Inc.
Robert Lester

Data Research Associates, Inc.
Michael J. Mellinger
Mark H. Needleman (Alt)

EBSCO Information Services
Melanie Watts

Elsevier Science Inc.
John Mancia

Endeavor Information Systems, Inc.
Verne Coppi
Cindy Miller (Alt)

NISO Voting Members *(continued)*

epixtech, inc.
John Bodfish
Ricc Ferrante (Alt)

Ex Libris
James Steenbergen
Carl Grant (Alt)

Follett Corp.
D. Jeffrey Blumenthal
Don Rose (Alt)

Fretwell-Downing Informatics
Robin Murray

Gale Group
Katherine Gruber
Justine Carson (Alt)

Gaylord Information Systems
James English
William Schickling (Alt)

GCA Research Institute
Jane Harnad

Geac Computers

Indiana Cooperative Library Services Authority
Millard Johnson
Dea Szatkowski (Alt)

Innovative Interfaces, Inc.
Gerald M. Kline
Sandy Westall (Alt)

Institute for Scientific Information
Helen Atkins
Richard Newman (Alt)

The International DOI Foundation
Norman Paskin

Kluwer Academic Publishers
Mike Casey

Library Binding Institute
J. Wesley Moyer

The Library Corporation
Mark Wilson
Nancy Capps (Alt)

Library of Congress
Winston Tabb
Sally H. McCallum (Alt)

Los Alamos National Laboratory
Richard E. Luce

Lucent Technologies
M.E. Brennan

Medical Library Association
Nadine P. Ellero
Carla J. Funk (Alt)

MINITEX
Cecelia Boone
William DeJohn (Alt)

Motion Picture Association of America
William M. Baker
Axel aus der Muhlen (Alt)

Music Library Association
Lenore Coral
Geraldine Ostrove (Alt)

National Agricultural Library
Pamela Q. J. Andre
Gary K. McCone (Alt)

National Archives and Records Administration
Alan Calmes

National Federation of Abstracting and
  Information Services
Marion Harrell

National Library of Medicine
Betsy L. Humphreys

OASIS
Laura Walker

OCLC, Inc.
Donald J. Muccino

Openly Informatics, Inc.
Albert Simmonds
Eric Hellman (Alt)

R.R. Bowker
Albert Simmonds

Recording Industry Association of America
Linda R. Bocchi
Michael Williams (Alt)

Research Libraries Group, Inc.
Kathleen Bales
Wayne Davison (Alt)

RoweCom, Inc.
Robert Boissy
Marilyn Geller (Alt)

SilverPlatter Information, Inc.
Marwan Sabbouh

SIRS, Inc.
Leonardo Lazo
Harry Kaplanian (Alt)

Society for Technical Communication
Annette Reilly
Kevin Burns (Alt)

Society of American Archivists
Lisa Weber

Special Libraries Association
Marjorie Hlava

SUNY/OCLC Network
Mary-Alice Lynch
Jane Neale (Alt)

Triangle Research Libraries Network
Jordan M. Scepanski
Mona C. Couts (Alt)

(continued)

NISO Voting Members *(continued)*

U.S. Department of Commerce,
  National Institute of Standards and
  Technology, Office of Information
  Services

U.S. Department of Defense, Defense
  Technical Information Center
Gretchen A. Schlag
Gopalakrishnan Nair (Alt)

U.S. National Commission on Libraries
  and Information Science
Denise Davis

VTLS, Inc.
Vinod Chachra

The H. W. Wilson Company
George I. Lewicky
Ann Case (Alt)

NISO Board of Directors

At the time NISO approved this standard, the following individuals served on its Board
of Directors:

Donald J. Muccino, Chair
OCLC
Online Computer Library Center, Inc.

Beverly P. Lynch, Vice-Chair/Chair-elect
University of California - Los Angeles

Joel H Baron, Immediate Past Chair
Healthgate Date Corp

Jan Peterson, Treasurer
Infotrieve

Patricia R. Harris, Executive Director
National Information Standards Organization

Pieter S. H. Bolman
Academic Press

Priscilla Caplan
Florida Center for Library Automation

Carl Grant
Ex Libris (USA), Inc.

Brian Green
Book Industry Communication

Richard E. Luce
Los Alamos National Laboratory

Deanna B. Marcum
Council on Library and Information Resources

Norman Paskin
The International DOI Foundation

Steven Puglia
U.S. National Archives and Records Administration

Jordan M. Scepanski
Triangle Research Libraries Network

Albert Simmonds
Openly Informatics, Inc.

NISO Committee AS

The following individuals served on Committee AS, The Dublin Core Metadata
Element Set:

John Kunze, Chair
University of California/National Library of
Medicine

Rebecca Guenther
Library of Congress

Marjorie Hlava
Access Innovations, Inc.

Clifford Morgan
John Wiley & Sons Ltd.

John Perkins
CIMI Consortium

The Dublin Core Metadata Element Set

1. Scope and Purpose

The Dublin Core metadata element set is a standard for cross-domain information resource description. Here an information resource is defined to be anything that has identity; this is the definition used in Internet RFC 2396, "Uniform Resource Identifiers (URI): Generic Syntax," by Tim Berners-Lee et al. For Dublin Core applications a resource will typically be an electronic document.

This standard is for the element set only, which is generally used in the context of a specific project or application. Local or community based requirements and policies may impose additional restrictions, rules, and interpretations. It is not the purpose of this standard to define the detailed criteria by which the element set will be used with specific projects and applications.

This standard supersedes Internet RFC 2413, which was the first published version of the Dublin Core.

2. Referenced Standards

[DCT] DCMI Type Vocabulary. DCMI Recommendation, 11 July 2000.
http://dublincore.org/documents/dcmi-type-vocabulary/

[ISO3166] ISO 3166 - Codes for the representation of names of countries.
http://www.din.de/gremien/nas/nabd/iso3166ma/

[ISO639] ISO 639-2 - Codes for the representation of names of languages, Alpha-3 code (ISO 639-2:1998). http://www.loc.gov/standards/iso639-2/langhome.html

[MIME] Internet Media Types.
http://www.isi.edu/in-notes/iana/assignments/media-types/media-types

[RFC3066] Tags for the Identification of Languages, Internet RFC 3066. http://www.ietf.org/rfc/rfc3066.txt

[RFC2396] Uniform Resource Identifiers (URI): Generic Syntax, Internet RFC 2396.
http://www.ietf.org/rfc/rfc2396.txt

[RFC2413] Dublin Core Metadata for Resource Discovery. Internet RFC 2413.
http://www.ietf.org/rfc/rfc2413.txt

[RFC2731] Encoding Dublin Core Metadata in HTML. Internet RFC 2731.
http://www.ietf.org/rfc/rfc2731.txt

[TGN] Getty Thesaurus of Geographic Names.
http://www.getty.edu/research/tools/vocabulary/tgn/index.html

[W3CDTF] Date and Time Formats, W3C Note. http://www.w3.org/TR/NOTE-datetime

[XML] Extensible Markup Language. http://www.w3.org/TR/REC-xml

3. Definitions

DCMI — Dublin Core Metadata Initiative, the maintenance agency for the Dublin Core.

Information resource — anything that has identity (the same definition as in Internet RFC 2396).

Lifecycle of an information resource — a sequence of events that mark the development and use of an information resource. Some examples of events in a lifecycle are: Conception of an invention, Creation of a draft, Revision of an article, Publication of a book, Acquisition by a library, Transcription to magnetic disk, Migration to optical storage, Translation into English, and Derivation of a new work (e.g., a movie).

4. The Element Set

In the element descriptions below, each element has a descriptive label intended to convey a common semantic understanding of the element, as well as a unique, machine-understandable, single-word name intended to make the syntactic specification of elements simpler for encoding schemes.

Although some environments, such as HTML, are not case-sensitive, it is recommended best practice always to adhere to the case conventions in the element names given below to avoid conflicts in the event that the metadata is subsequently extracted or converted to a case- sensitive environment, such as XML (Extensible Markup Language) [XML].

Each element is optional and repeatable. Metadata elements may appear in any order. The ordering of multiple occurrences of the same element (e.g., Creator) may have a significance intended by the provider, but ordering is not guaranteed to be preserved in every system.

To promote global interoperability, a number of the element descriptions suggest a controlled vocabulary for the respective element values. It is assumed that other controlled vocabularies will be developed for interoperability within certain local domains.

5. The Elements

Element Name: `Title`

Label: Title

Definition: A name given to the resource.

Comment: Typically, Title will be a name by which the resource is formally known.

Element Name: `Creator`

Label: Creator

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

Comment: Examples of Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity.

Element Name: `Subject`

Label: Subject and Keywords

Definition: A topic of the content of the resource.

Comment: Typically, Subject will be expressed as keywords, key phrases, or classification codes that describe a topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.

Element Name: `Description`

Label: Description

Definition: An account of the content of the resource.

Comment: Examples of Description include, but are not limited to, an abstract, table of contents, reference to a graphical representation of content, or free-text account of the content.

Element Name: `Publisher`

Label: Publisher

Definition: An entity responsible for making the resource available.

Comment: Examples of Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity.

Element Name: `Contributor`

Label: Contributor

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

Comment: Examples of Contributor include a person, an organization, or a service. Typically, the name of a Contributor should be used to indicate the entity.

Element Name: `Date`

Label: Date

Definition: A date of an event in the lifecycle of the resource.

Comment: Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF] and includes (among others) dates of the form YYYY-MM-DD.

Element Name: `Type`

Label: Resource Type

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

Comment: Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary (for example, the DCMI Type Vocabulary [DCT]). To describe the physical or digital manifestation of the resource, use the Format element.

Element Name: `Format`

Label: Format

Definition: The physical or digital manifestation of the resource.

Comment: Typically, Format will include the media-type or dimensions of the resource. Format may be used to identify the software, hardware, or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [MIME] defining computer media formats).

Element Name: `Identifier`

Label: Resource Identifier

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

Comment: Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Formal identification systems include but are not limited to the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI), and the International Standard Book Number (ISBN).

Element Name: `Source`

Label: Source

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

Comment: The present resource may be derived from the Source resource in whole or in part. Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.

Element Name: `Language`

Label: Language

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

Comment: Recommended best practice is to use RFC 3066 [RFC3066], which, in conjunction with ISO 639 [ISO639], defines two- and three-letter primary language tags with optional subtags. Examples include "en" or "eng" for English, "akk for Akkadian, and "en-GB" for English used in the United Kingdom.

Element Name: `Relation`

Label: Relation

Definition: A reference to a related resource.

Comment: Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.

Element Name: `Coverage`

Label: Coverage

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

Comment: Typically, Coverage will include spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range), or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [TGN]) and to use, where appropriate, named places or time periods in preference to numeric identifiers such as sets of coordinates or date ranges.

Element Name: `Rights`

Label: Rights Management

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

Comment: Typically, Rights will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the Rights element is absent, no assumptions may be made about any rights held in or over the resource.

Appendix A:

Further Reading

(This Appendix is not part of the American National Standard Dublin Core Metadata Element Set, ANSI/NISO Z39.85-2001. It is included for information only.)

Further information about the Dublin Core metadata element set is available at the URL,

http://dublincore.org/

This web site contains information about workshops, reports, working group papers, projects, and new developments concerning the Dublin Core Metadata Initiative (DCMI).

Appendix B:

Maintenance Agency

(This Appendix is not part of the American National Standard Dublin Core Metadata Element Set, ANSI/NISO Z39.85-2001. It is included for information only.)

The Dublin Core Metadata Initiative (DCMI) is responsible for the development, standardization and promotion of the Dublin Core metadata element set. Information on DCMI is available at the URL,

http://dublincore.org/