



Scientific Research and High Studies

Jinan University



Dublin Core

prepared by
Dr. Ammar Yakan

Metadata for Interoperability

Web search

Anyone who has attempted to **find information online** using one of today's popular Web search services has likely experienced the frustration of **retrieving hundreds**, if not thousands, of "**hits**" with **limited ability to refine** or make a more precise search. The wide scale adoption of **descriptive standards** and **practices for electronic resources** will **improve retrieval** of relevant resources in any venue where information retrieval is critical.

What is Metadata?

Metadata can be thought of as **data about other data**, which most commonly refers to **descriptive information about Web resources**. A metadata record consists of a set of **attributes**, or **elements**, necessary to **describe the resource** in question. For example, a metadata system common in libraries – the **library catalog** -- contains a set of **metadata records** with elements that describe a book or other library item: **author, title, date of creation or publication, subject coverage**, and the call number specifying location of the item on the shelf.

The Dublin Core Metadata Initiative

- The Dublin Core Metadata Initiative (DCMI) is an **organization** dedicated to promoting the widespread **adoption of interoperable metadata standards** and **developing specialized metadata vocabularies** for describing resources that enable more **intelligent information discovery systems**.
- The **mission** of DCMI is to make it **easier to find resources using the Internet** through the following activities:
 1. **Developing metadata** standards for **discovery** across domains,
 2. **Defining frameworks for the interoperation** of metadata sets, and,
 3. **Facilitating the development of community- or disciplinary-specific** metadata sets that are consistent with items 1 and 2

Dublin Core defined

- The **Dublin Core Metadata Element Set** is a **vocabulary** of **fifteen properties** for use in **resource description**. The name "**Dublin**" is due to its origin at a 1995 invitational workshop in Dublin, Ohio; "**core**" because its elements are **broad** and **generic, usable** for describing a wide range of resources.
- The Dublin Core Metadata brought together **librarians, digital library researchers, content experts**, and **text-markup experts** to promote better discovery standards for electronic resources. The resulting metadata element set defines fifteen metadata elements for resource description in a **cross-disciplinary information environment**.
- The fifteen element "Dublin Core" described in this standard **is part of a larger set of metadata vocabularies** and technical specifications maintained by the **Dublin Core Metadata Initiative (DCMI)**.
- Since **1998**, when these fifteen elements entered into a **standardization** track, notions of best practice in the **Semantic Web** have evolved to include the assignment of **formal domains and ranges** in addition to definitions in **natural language**.

Dublin Core defined (continued)

Domains and ranges express the meanings **implicit in natural-language** definitions in an **explicit form that is usable for the automatic processing of logical inferences**. When a given property is encountered, an **inferencing application** may use information about the domains and ranges assigned to a property in order to make inferences about the resources described thereby.

- [ANSI/NISO Z39.85-2012](#)
- [ISO 15836:2009](#)
- [Translations of DCMI Documents](#)
- [DCMI Metadata Terms](#)

The Elements

1- Term Name: **contributor**

URI:	http://purl.org/dc/elements/1.1/contributor
Label:	Contributor
Definition:	An entity responsible for making contributions to the resource.
Comment:	Examples of a Contributor include a person , an organization , or a service . Typically, the name of a Contributor should be used to indicate the entity.

2- Term Name: **coverage**

URI:	http://purl.org/dc/elements/1.1/coverage
Label:	Coverage
Definition:	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.
Comment:	Spatial topic and spatial applicability may be a named place or a location specified by its geographic coordinates . Temporal topic may be a named period , date , or date range . A jurisdiction may be a named administrative entity or a geographic place to which the resource applies. Recommended best practice is to use a controlled vocabulary such as the Thesaurus of Geographic Names [TGN] . Where appropriate, named places or time periods can be used in preference to numeric identifiers such as sets of coordinates or date ranges.
References:	[TGN] http://www.getty.edu/research/tools/vocabulary/tgn/index.html

The Elements^(continued)

3- Term Name: **creator**

URI:	http://purl.org/dc/elements/1.1/creator
Label:	Creator
Definition:	An entity primarily responsible for making the resource.
Comment:	Examples of a Creator include a person , an organization , or a service . Typically, the name of a Creator should be used to indicate the entity.

4- Term Name: **date**

URI:	http://purl.org/dc/elements/1.1/date
Label:	Date
Definition:	A point or period of time associated with an event in the lifecycle of the resource.
Comment:	Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme, such as the W3 CDTF profile of ISO 8601 [W3CDTF].
References:	[W3CDTF] http://www.w3.org/TR/NOTE-datetime

The Elements^(continued)

5- Term Name: **description**

URI:	http://purl.org/dc/elements/1.1/description
Label:	Description
Definition:	An account of the resource.
Comment:	Description may include but is not limited to: an abstract , a table of contents , a graphical representation , or a free-text account of the resource.

6- Term Name: **format**

URI:	http://purl.org/dc/elements/1.1/format
Label:	Format
Definition:	The file format , physical medium, or dimensions of the resource.
Comment:	Examples of dimensions include size and duration . Recommended best practice is to use a controlled vocabulary such as the list of Internet Media Types [MIME].
References:	[MIME] http://www.iana.org/assignments/media-types/

The Elements^(continued)

7- Term Name: **identifier**

URI:	http://purl.org/dc/elements/1.1/identifier
Label:	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 conforming to a formal identification system.

8- Term Name: **language**

URI:	http://purl.org/dc/elements/1.1/language
Label:	Language
Definition:	A language of the resource.
Comment:	Recommended best practice is to use a controlled vocabulary such as RFC 4646 [RFC4646].
References:	[RFC4646] http://www.ietf.org/rfc/rfc4646.txt

The Elements^(continued)

9- Term Name: **publisher**

URI:	http://purl.org/dc/elements/1.1/publisher
Label:	Publisher
Definition:	An entity responsible for making the resource available.
Comment:	Examples of a Publisher include a person , an organization , or a service . Typically, the name of a Publisher should be used to indicate the entity.

10- Term Name: **relation**

URI:	http://purl.org/dc/elements/1.1/relation
Label:	Relation
Definition:	A related resource.
Comment:	Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system .

The Elements^(continued)

11- Term Name: **rights**

URI:	http://purl.org/dc/elements/1.1/rights
Label:	Rights
Definition:	Information about rights held in and over the resource.
Comment:	Typically, rights information includes a statement about various property rights associated with the resource, including intellectual property rights.

12- Term Name: **source**

URI:	http://purl.org/dc/elements/1.1/source
Label:	Source
Definition:	A related resource from which the described resource is derived.
Comment:	The described resource may be derived from the related resource in whole or in part. Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system .

The Elements^(continued)

13- Term Name: **subject**

URI:	http://purl.org/dc/elements/1.1/subject
Label:	Subject
Definition:	The topic of the resource.
Comment:	Typically, the subject will be represented using keywords, key phrases, or classification codes . Recommended best practice is to use a controlled vocabulary .

14- Term Name: **title**

URI:	http://purl.org/dc/elements/1.1/title
Label:	Title
Definition:	A name given to the resource.
Comment:	Typically, a Title will be a name by which the resource is formally known .

The Elements^(continued)

15- Term Name: type	
URI:	http://purl.org/dc/elements/1.1/type
Label:	Type
Definition:	The nature or genre of the resource.
Comment:	Recommended best practice is to use a controlled vocabulary such as the DCMI Type Vocabulary [DCMITYPE]. To describe the file format , physical medium , or dimensions of the resource, use the Format element.
References:	[DCMITYPE] http://dublincore.org/documents/dcmi-type-vocabulary/

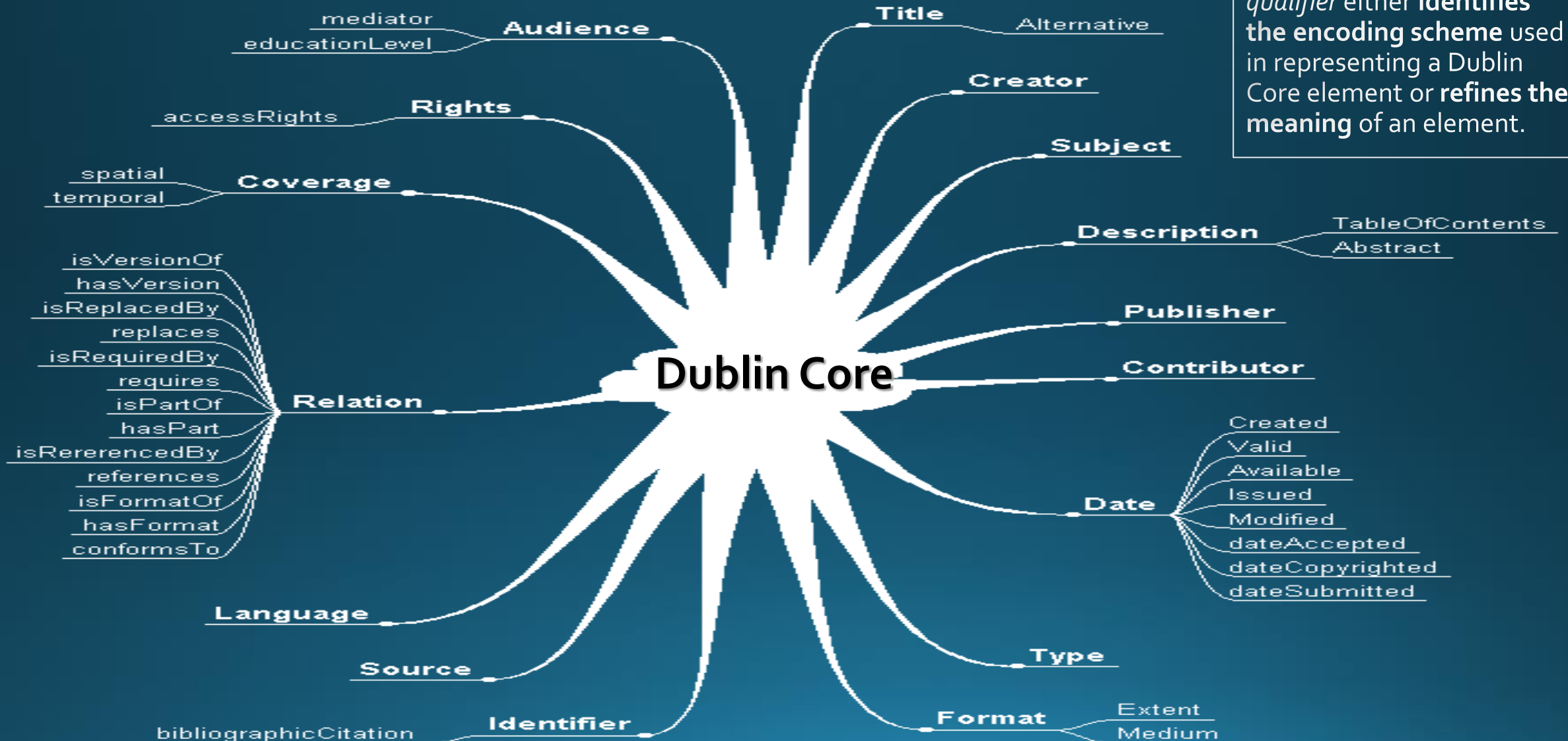
Content	Intellectual Property	Instantiation
Coverage	Contributor	Date
Description	Creator	Format
Type	Publisher	Identifier
Relation	Rights	Language
Source		
Subject		
Title		

Typical Example (can be repeated endlessly)

1. **Contributor**="Association of the Aircraft Pilots"
2. **Coverage**="Lebanese Aviation Jurisdiction"
3. **Creator**="Riad T."
4. **Date**="1998-02-16"
5. **Description**="Illustrated guide to airport markings and lighting signals"
6. **Format**="book/pdf 24 x 17 cm"
7. **Identifier**="H-A-X 5690B"
8. **Language**="en;fr"
9. **Publisher**="University of Miami. Faculty of Aviation"
10. **Relation**="HasVersion 13th Edition, 1998" | Relation ="IsPartOf University Faculty Program"
11. **Rights**="Access limited to University students."
12. **Source**="RC607.A26W574 1996" (call no. of the "Encyclopedia of Aviation")
13. **Subject**="Aircraft leasing and renting"
14. **Type**="text"
15. **Title**="A Pilot's Guide to Aircraft Control Panel"

Dublin Core Qualifiers

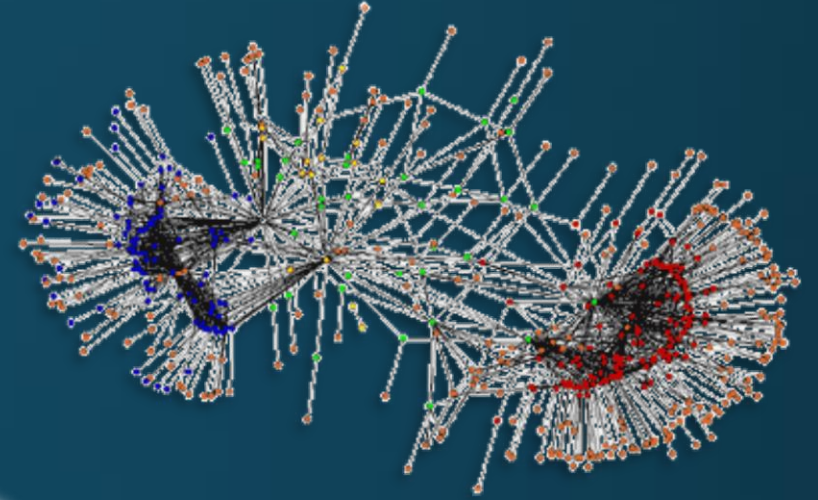
Dublin Core Qualifiers is a companion specification to the Dublin Core Metadata Element Set. A *qualifier* either **identifies the encoding scheme** used in representing a Dublin Core element or **refines the meaning** of an element.



Dublin Core element set goals:

1. Simplicity of creation and maintenance (e.g. non-specialist).
2. Commonly understood semantics (e.g. creator).
3. International scope (languages).
4. Extensibility (communities).
5. All elements are optional, and all elements are repeatable.

Ontologies (by Dublin Core)



Dublin Core Applications

1. Library of Congress (USA)
2. American Museum of Natural History.
3. Support the Description and Discovery of Language Resources
 - Open Language Archives Community(OLAC)
 - **Discourse Type:** drama, formulaic discourse, interactive discourse, language play, oratory, narrative, procedural discourse, report, singing, and unintelligible speech.
 - **Role:** annotator, artist, author, compiler, consultant, depositor, developer, editor, illustrator, interviewer, participant, performer, photographer, recorder, researcher, respondent, signer, speaker, sponsor, transcriber, and translator.
 - **Linguistic Subject:** anthropological linguistics, applied linguistics, cognitive science, computational linguistics, discourse analysis, forensic linguistics, general linguistics, historical linguistics, history of linguistics, language acquisition, language documentation, lexicography, linguistics and literature, linguistic theories, mathematical linguistics, morphology, neurolinguistics, philosophy of language, phonetics, phonology, pragmatics, psycholinguistics, semantics, sociolinguistics, syntax, text and corpus linguistics, translating and interpreting, typology, and writing systems.
4. The Publication and Citation of Research Data ([see ResDC](#))
5. Geospatial Multistate Archive
6. Librarians / archivist / museum professionals / scientists /government specialists / educators / business/corporate workers

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.

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">
```

XML

```
<?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>
```

Thank you!