

Dublin Core

Dublin Core: background

- ◆ Necessità di dare una rappresentazione formalizzata alle risorse di informazione accessibili in rete via WEB
- ◆ Varietà delle risorse: testi, suoni, immagini tridimensionali, video, fotografie, film ...
- ◆ Necessità di una rappresentazione capace di comprendere i vari tipi di risorse

Dublin Core: background

- ◆ Esigenze accolte da un gruppo di bibliotecari, e di informatici....
- ◆ “Finding relevant information on the World Wide Web has become increasingly problematic due to the explosive growth of networked resources. Current Web indexing evolved rapidly to fill the demand for resource discovery tools, but that indexing, while useful, is a poor substitute for richer varieties of resource description.”
- ◆ Dublin Core Metadata Initiative <http://www.ietf.org/rfc/rfc2413.txt>
- ◆ **Dublin Core Metadata Set. Un insieme di elementi descrittivi capaci di rappresentare qualsiasi risorsa di informazione accessibile in rete**

Dublin Core: background

- ◆ **Il *D.C. metadata set* e' il formato di metadati piu' diffuso oggi, risultato di un progetto iniziato nel 1995 [Dublin Core Metadata Workshop Series] con l'obiettivo di affrontare il problema della ricerca di risorse di informazione disponibili sulla rete**
 - documenti testuali, programmi, video....ma anche basi di dati, archivi, etc..
- ◆ **La sua caratteristica - frutto di una scelta ben precisa - è la generalità e la semplicità. Esso cioè è stato pensato come un insieme minimo di elementi descrittivi capaci di rappresentare qualsiasi risorsa di informazione -pagina web, un audio, un video etc. - che abbia un indirizzo WEB**

Caratteristiche di Dublin Core

- ◆ ***Si applicano a tutti i tipi di risorse di informazione***
- ◆ ***Sono facili da creare, anche da parte di personale non specializzato (a differenza di quanto succede per i record catalografici di tipo MARC)***
- ◆ ***Sono di uso generale per la “scoperta” delle risorse di informazione (a differenza delle descrizioni catalografiche che in genere seguono normative specifiche del settore di applicazione: biblioteche, musei, videoteche...)***
- ◆ ***Il modello Dublin Core è “piatto”, cioè non permette di descrivere metadati strutturati, ed ha un numero di elementi limitato (15)***

Linee guida nella definizione di Dublin Core

◆ **Semplicità**

- Anche i non esperti devono poter creare e gestire i metadati
- Insieme di metadati sufficiente per una ricerca efficace

◆ **Semantica universalmente accettata**

- Semantica identica per i catalogatori e per chi formula le richieste

◆ **Possibilità di utilizzo in ambito internazionale**

- Esistono diverse versioni nazionali

◆ **Estendibilità**

- Possibilità di estendere il modello per soddisfare esigenze di particolari categorie di utenti
- Si utilizzano gli “application profiles”

Dublin Core: principi generali

- ◆ **Ogni elemento è opzionale e può essere ripetuto**
 - Semplifica l'interoperabilità con altri modelli di metadati
 - Ha lo svantaggio che si possono creare record DC poco significativi

- ◆ **Ogni record di metadati DC descrive una manifestazione della risorsa**
 - Non deve esistere confusione tra la risorsa e quello che essa descrive (ad es. i metadati di una foto della Torre di Pisa dovranno descrivere la risorsa foto e non la Torre di Pisa)

- ◆ **I qualificatori sono opzionali**
 - I qualificatori fanno parte di una estensione di DC che potrebbe non essere supportata da tutti i sistemi
 - Un record DC con i qualificatori deve essere valido anche se questi vengono trascurati

Caratteristiche di Dublin Core

- ◆ **Vi sono diverse ragioni per adottare Dublin Core:**
 - I metadati Dublin Core sono facili da creare
 - Possono essere facilmente indicizzati
 - Permettono una precisione maggiore rispetto ai sistemi che utilizzano full text
 - Permettono l'interoperabilità

- ◆ **Dublin Core fornisce gli strumenti per**
 - Descrivere una risorsa (location, form etc.)
 - Fornire semplici descrizioni e permetterne l'indicizzazione
 - Individuare una risorsa
 - Ottenere ed accedere una risorsa
 - Migliorare la qualità della gestione delle risorse

Dublin Core metadata set [1/18]

- ◆ L'insieme degli elementi definiti nel Dublin Core Metadata Element Set, Version 1.1: Reference Description è presentato nella traduzione italiana, reperibile all'indirizzo <http://www.iccu.sbn.it/metadati.htm>

- ◆ Esiste un mapping tra UNIMARC e Dublin Core
 - ◆ Title
 - ◆ Creator
 - ◆ Subject
 - ◆ Description
 - ◆ Publisher
 - ◆ Contributor
 - ◆ Date
 - ◆ Type
 - ◆ Format
 - ◆ Identifier
 - ◆ Source
 - ◆ Language
 - ◆ Relation
 - ◆ Coverage
 - ◆ Rights

Dublin Core metadata set [2/18]

◆ Elemento: Titolo

| | |
|-----------------|---|
| Nome: | Titolo |
| Identificatore: | Title |
| Definizione: | Un nome dato alla risorsa. Tipicamente, Title è il nome con il quale la risorsa è conosciuta |
| Esempi: | Title = "La Divina Commedia" Title = "La Nona Sinfonia" Title = "Pietà" Title = "2001: Odissea nello Spazio" |
| Commenti: | Il titolo da solo non permette di individuare in modo univoco la risorsa |

Dublin Core metadata set [3/18]

◆ Elemento: **Creatore**

Nome: Creatore

Identificatore: Creator

Definizione: Un'entità che ha la responsabilità principale della produzione del contenuto della risorsa. Esempi di creator: una persona, un'organizzazione, un servizio software.

Regole generali: I vari creator vanno riportati separatamente.

Il cognome deve precedere il nome.

Nel caso di organizzazioni strutturate, elencare le varie componenti, a partire da quella principale.

Esempi: Creator = "Alighieri, Dante"

Creator = "Internal Revenue Service. Customer Complaints Unit"

Dublin Core metadata set [4/18]

◆ Elemento: **Soggetto**

Nome: Soggetto e Parole chiave

Identificatore: Subject

Definizione: L'argomento della risorsa. Tipicamente un Soggetto viene espresso da parole chiave o frasi o codici di classificazione che descrivono l'argomento della risorsa. È buona norma selezionare il valore da un vocabolario controllato o da uno schema di classificazione.

Esempio: Subject = "Poesia"
Subject = "Sinfonia"

Commento: Non è possibile avere soggetti strutturati, quali ad esempio "software/biblioteche digitali"
In questi casi bisogna utilizzare due Soggetti

Dublin Core Metadata Set [5/18]

◆ **Elemento:** **Descrizione**

Nome: Descrizione

Identificatore: Description

Definizione: Una spiegazione del contenuto della risorsa

Commento: Una descrizione può includere ma non è limitata solo a: un riassunto analitico, un indice, un riferimento al contenuto di una rappresentazione grafica o un testo libero del contenuto.

Molto utile in fase di ricerca. Normalmente vengono utilizzate frasi complete.

Può essere impreciso. Inoltre classificatori diversi possono fornire descrizioni diverse della stessa risorsa.

Esempio: Description = “Questa è una foto di Piazza dei Miracoli a Pisa. A destra si vede la Torre di Pisa, al centro il Duomo ed a sinistra il Battistero”

Dublin Core Metadata Set [6/18]

- ◆ **Elemento:** **Editore**
 - Nome: Editore
 - Identificatore: Publisher
 - Definizione: Un'entità responsabile della produzione della risorsa, disponibile nella sua forma presente
 - Commento: Esempi di un Publisher includono una persona, un'organizzazione o un servizio. Normalmente il nome di un Publisher è usato per indicare l'entità stessa
Questo campo viene specificato per identificare l'entità che fornisce accesso alla risorsa.
 - Esempi: Publisher="Mondadori"

Dublin Core Metadata Set [7/18]

◆ **Elemento: Autore di contributo subordinato**

Nome: Autore di contributo subordinato

Identificatore: Contributor

Definizione: Un'entità responsabile della produzione di un contributo al contenuto della risorsa

Commento: Esempi di un Contributor includono una persona, un'organizzazione o un servizio. Normalmente il nome di un Contributor è usato per indicare l'entità stessa

Dublin Core Metadata Set [8/18]

◆ Elemento: Data

Nome: Data

Identificatore: Date

Definizione: Una data associata a un evento del ciclo di vita della risorsa. Normalmente viene associata con la data di creazione o di disponibilità della risorsa. Si raccomanda l'uso di codifiche standard per le date. Se non si conosce la data completa, si può usare solo il mese-anno o solo l'anno.

Dublin Core Metadata Set [9/18]

◆ Elemento: Tipo

Nome: Tipo di risorsa

Identificatore: Type

Definizione: La natura o il genere del contenuto della risorsa

Commento: Tipo include termini che descrivono categorie generali, funzioni, generi o livelli di aggregazione per contenuto. Si raccomanda di selezionare un valore da un vocabolario controllato (ad esempio, la lista dei “Dublin Core Types”). Per descrivere la manifestazione fisica o digitale della risorsa, si usa l’elemento Format.

Se una risorsa ha più tipi, bisogna utilizzare più campi Type.

Esempi: Type=“Text”

Type=“Sound”

Type=“Image”

Dublin Core Metadata Set [10/18]

- ◆ <http://dublincore.org/usage/terms/dcmitype/>
- ◆ **Dublin Core Types** - Dizionario dei termini usabili quali “valore” dell’elemento “TYPE” per identificare il genere di una risorsa:
 - Collection – Aggregazione di elementi.
 - Dataset – Informazione codificata in una struttura definita (lista, tabella, ...)
 - Event – Avvenimento non persistente nel tempo (conferenza, intervista, ...)
 - Image – Include anche foto, disegni, grafici, ecc.
 - Interactive Resource – richiede l’interazione con l’utente
 - Moving Image – Ad es. animazioni, film, programmi televisivi, ecc.
 - Physical Object – Oggetto tridimensionale, ad es. scultura, palazzo, ecc.
 - Service – Sistema che fornisce una o più funzioni ad un utente.
 - Software
 - Sound
 - Still image
 - Text

Dublin Core Metadata Set [11/18]

| ◆ Elemento: | Formato |
|-----------------|--|
| Nome: | Formato |
| Identificatore: | Format |
| Definizione: | La manifestazione fisica o digitale della risorsa. |
| Commento: | <p>Normalmente Format può includere il tipo di supporto o le dimensioni della risorsa.</p> <p>Format può essere usato per determinare il software, l'hardware o altro apparato necessario alla visualizzazione o all'elaborazione della risorsa.</p> <p>Esempi di dimensioni includono grandezza e durata. Si raccomanda di selezionare un valore da un vocabolario controllato (ad esempio la lista di Internet Media Types [MIME]) che definisce i formati dei supporti elettronici.</p> <p>Se vi sono più formati per una stessa risorsa, il campo Format deve essere ripetuto.</p> |
| Esempi: | <pre>Title="Dublin Core icon" Identifier="http://purl.org/metadata/dublin_core/images/dc2.gif" Type="Image" Format="image/gif" Format="4 kB" Subject="Saturn" Type="Image" Format="image/gif 6" Format="40 x 512 pixels" Identifier="http://www.not.iac.es/newwww/photos/images/satnot.gif"</pre> |

Dublin Core Metadata Set [12/18]

INTERNET MEDIA TYPES (<http://www.iana.org/assignments/media-types/>)

Application

Audio (g722, mpeg, mpa, ...)

Image (e.g. jpeg, gif, tiff, ...)

Message

Model

Multipart

Text (e.g. rtf, sgml, richtext, xml, ...)

Video (e.g. quicktime, mpeg, H264, ...)

Dublin Core Metadata Set [13/18]

◆ Elemento: **Identificatore**

Nome: Identificatore della risorsa

Identificatore: Identifier

Definizione: Un riferimento univoco alla risorsa nell'ambito di un dato contesto

Commento: Si raccomanda di identificare la risorsa utilizzando un sistema formale di identificazione (ad.es. URL, URI, ISBN). Si può anche utilizzare un identificatore locale.

Esempi: Identifier="http://www.intratext.com/X/ITA0191.htm"
Identifier="ISBN:0382364759"

Dublin Core Metadata Set [14/18]

◆ Elemento: Fonte

Nome: Fonte

Identificatore: Source

Definizione: Un riferimento a una risorsa dalla quale è derivata la risorsa in oggetto

Commento: La risorsa può essere derivata dalla risorsa Source in tutto o in parte. Si raccomanda di identificare la risorsa per mezzo di una sequenza di caratteri alfabetici o numerici secondo un sistema di identificazione formalmente definito

Esempi: Source="RC607.A26W574 1996" (Call Number della versione a stampa)

Source="Immagine da pagina 32 dell'edizione pubblicata nel 1922 della Divina Commedia"

Dublin Core Metadata Set [15/18]

◆ Elemento: **Lingua**

Nome: Lingua

Identificatore: Language

Definizione: La lingua del contenuto intellettuale della risorsa. Si raccomanda l'uso di valori presi da una lista standard (RCF 3066)

Esempi: Language="it"

Language="Primarily English, with some abstracts also in French"

Dublin Core Metadata Set [16/18]

◆ Elemento: **Relazione**

Nome: Relazione

Identificatore: Relation

Definizione: Un riferimento alla risorsa correlata. Si raccomanda l'uso di un sistema di identificazione. Può indicare relazioni di tipo Parte/Tutto, Versioni, Formato derivato, citazioni, ecc.

Esempio: Title="La Divina Commedia. Inferno"
Relation="Parte 1/3"

Dublin Core Metadata Set [17/18]

◆ **Elemento:** **Copertura**

Nome: Copertura

Identificatore: Coverage

Definizione: L'estensione o scopo del contenuto della risorsa

Commento: Normalmente Coverage include la localizzazione spaziale (il nome di un luogo o le coordinate geografiche), il periodo temporale (l'indicazione di un periodo, una data o un range di date) o una giurisdizione (ad esempio il nome di un'entità amministrativa).

Si raccomanda di selezionare un valore da un vocabolario controllato (ad esempio il Thesaurus of Geographic Names [TGN]) e, se possibile, di utilizzare i nomi di luogo o i periodi di tempo piuttosto che identificatori numerici come serie di coordinate o range di date

Esempi:

Coverage="1995-1996"

Coverage="Pisa, IT"

Coverage="17th century"

Dublin Core Metadata Set [18/18]

◆ **Elemento: Gestione dei diritti**

Nome: Gestione dei diritti

Identificatore: Rights

Definizione: Informazione sui diritti esercitati sulla risorsa

Commento: Normalmente un elemento Rights contiene un'indicazione di gestione dei diritti sulla risorsa, o un riferimento ad un servizio che fornisce questa informazione. L'informazione sui diritti spesso comprende i diritti di proprietà intellettuale Intellectual Property Rights (IPR), Copyright e vari diritti di proprietà. Se l'elemento Rights è assente, non si può fare alcuna ipotesi sullo stato di questi o altri diritti in riferimento alla risorsa

Esempio: Rights="Accesso limitato ai membri dell'associazione"

Qualifiers [1/5]

- ◆ **Gli elementi di Dublin Core possono essere poco specifici ed ambigui in molti casi, per cui è in corso un processo di estensione basato sull'uso di *qualifiers* da associare ad ogni Dublin Core element**
- ◆ **Ogni programma che utilizza una descrizione Dublin Core deve essere in grado di ignorare i qualifiers, oppure di interpretarli**
- ◆ **Vi sono attualmente due classi di qualifiers**
 - *Element refinement*
 - *Rendono il significato di un elemento più specifico*
 - *Encoding Scheme*
 - *Semplificano l'interpretazione dell'elemento.*
 - *Possono includere vocabolari controllati o regole di elaborazione (ad es. per le date)*

Qualifiers [2/5]

- ◆ **Esempi di “element refinements”**. La lista completa si trova in <http://www.dublincore.org/documents/usageguide/qualifiers.shtml>
- ◆ **Refinement per “Title”**
 - Alternative – titolo aggiuntivo o alternativo
 - Ad es. Title = “ACM Transaction on Information Systems”
Alternative = “ACM TOIS”
- ◆ **Refinements per “Description”**
 - Table of Contents
 - Abstract
- ◆ **Refinements per “Date”**
 - Created – data di creazione della risorsa
 - Valid – data di validità della risorsa
 - Available – data di validità
 - Issued
 - Modified
 - Accepted
 - Submitted

Qualifiers [3/5]

- ◆ **Gli “Encoding Schema” sono schemi di codifica per elementi specifici di DC**
- ◆ **In particolare per Subject, Date, Type, Format, Identifier, Source, Language, Relation, Coverage sono stati definiti degli Encoding Schema**
- ◆ **Esempi:**
 - **Subject: DDC (Dewey Decimal Classification)**
 - **<http://www.oclc.org/dewey/>**
 - **Sistema di classificazione usato nel mondo librario**
 - **Fornisce un sistema per la classificazione della conoscenza**
 - **Suddivisione in 10 classi (Computers, Filosofia, Religione, ecc.)**
 - **Ogni classe è a sua volta suddivisa in sottoclassi**

Qualifiers [4/5]

◆ Esempi:

- Subject: LCC (Library of Congress Classification)
 - <http://www.loc.gov/catdir/cpsol/lcco/lcco.html>
 - Diverse classi (da A a Z)
 - A – General Work
 - B – Philosophy, Psychology, Religion
 - Ecc.
 - Ogni classe è a sua volta diviso in sottoclassi
- Date: DMCI Period – Codifica per specificare un periodo di tempo
 - <http://dublincore.org/documents/dcmi-period/>
 - Il periodo è specificato da [start, end, scheme, name]
 - Tutte le componenti sono opzionali
 - “start” ed “end” sono gli istanti di inizio e fine del periodo
 - “scheme” è lo schema di codifica per la rappresentazione di “start” ed “end”
 - “name” è il nome dell’intervallo di tempo
 - Esempi
 - name=The Great Depression; start=1929; end=1939;

 - name="1999 AFL Grand Final": start=1999-09-25T14:20+10:00;
end=1999-09-25T16:40+10:00; scheme=W3C-DTF;

Qualifiers [5/5]

◆ Esempi:

- Date: W3C-DTF – Regole di codifica per le date ed i tempi
 - ➔ <http://www.w3.org/TR/NOTE-datetime>
 - ➔ Sono consentite sei diverse tipologie
 - Anno: YYYY (eg 1997)
 - Anno e mese: YYYY-MM (eg 1997-07)
 - Data completa: YYYY-MM-DD (eg 1997-07-16)
 - Data completa più ore e minuti: YYYY-MM-DDThh:mmTZD (eg 1997-07-16T19:20+01:00)
 - Data completa più ore, minuti e secondi : YYYY-MM-DDThh:mm:ssTZD (eg 1997-07-16T19:20:30+01:00)
 - Data completa più ore, minuti, secondi e decimi di secondo YYYY-MM-DDThh:mm:ss.sTZD (eg 1997-07-16T19:20:30.45+01:00)
 - > dove TZD = indicatore time zone (Z o +hh:mm o -hh:mm)
 - > Il tempo viene espresso come UTC (Coordinated Universal Time) (indicato con Z) o come tempo locale più offset (+- hh:mm)
 - Esempi
 - > 1994-11-05T08:15:30-05:00 corrisponde al 5 Novembre, 1994, 8:15:30 am, US Eastern Standard Time.
 - > 1994-11-05T13:15:30Z corrisponde allo stesso istante.

Esempio di record Dublin Core

Supponiamo di voler creare un record DC per il libro di Witten, Bainbridge su Greenstone. Un possibile record Dublin Core è il seguente (si può usare l'applicazione disponibile in <http://www.lub.lu.se/cgi-bin/nmdc.pl>):

| | |
|--------------------------------|---|
| DC.Title | <i>How to Build a Digital Library</i> |
| DC.Creator.PersonalName | <i>Witten, Ian H.</i> |
| DC.Creator.PersonalName2 | <i>Bainbridge, David</i> |
| DC.Subject | <i>Greenstone</i> |
| DC.Subject | Information Storage and Retrieval |
| DC.Subject | Documentation (CCS) |
| DC.Subject | H. Information Systems (CCS) |
| DC.Subject | H.4 Information Systems Application (CCS) |
| DC.Description | Il libro descrive come creare una |
| DC.Publisher | Morgan Kaufmann Publishers |
| DC.Contributor.CorporateName | University of Waikato in New Zeland |
| DC.Date | 2003-01-01 (ISO8601) |
| DC.Type | Text |
| DC.Format | text/html (IMT) |
| DC.Identifier | http://pc-erato2.isti.cnr.it/savino/Greenstone-book.html |
| DC.Source | 1-55860-790-0 (ISBN) |
| DC.Language | it (ISO639-1) |
| DC.Date.X-MetadataLastModified | 2004-10-22 (ISO8601) |

Rappresentazione di Dublin Core

- ◆ **Dublin Core può essere rappresentato utilizzando diversi formati**
- ◆ **Sono state sviluppate rappresentazioni sia di “simple DC” che di “qualified DC” in**
 - XML: <http://dublincore.org/documents/2003/04/02/dc-xml-guidelines/>
 - RDF: <http://dublincore.org/documents/2002/07/31/dcmes-xml/> e <http://dublincore.org/documents/dcq-rdf-xml/>

Riferimenti

- ◆ **Informazione aggiornata su Dublin Core**
<http://dublincore.org/index.shtml>
- ◆ **La documentazione aggiornata si trova a partire dall'indirizzo: <http://dublincore.org/usage/>**
- ◆ **Diane Hillman, Using Dublin Core,**
<http://dublincore.org/documents/usageguide/>
- ◆ **Dublin Core Metadata Template**
<http://www.lub.lu.se/cgi-bin/nmdc.pl>
- ◆ **Crea dublin core da pagine web**
<http://www.ukoln.ac.uk/metadata/dcdot/>