

Le Dublin Core (DC)

Élizabeth CHERHAL
Cellule MathDoc, UMS5638, CNRS/Université Joseph Fourier,
Grenoble.

Note : article rédigé par l'auteur à la suite de sa communication orale du 16 novembre 2004.

Résumé :

Nous rappellerons le contexte historique de l'initiative Dublin Core, puis verrons l'ensemble des éléments de base et la notion de Dublin Core qualifié. Nous verrons ensuite que la description bibliographique est indépendante des formats physiques et présenterons quelques-unes des représentations physiques les plus utilisées.

Mots clés :

Dublin Core, métadonnées, indexation, archives ouvertes, Open Archives Initiative, normes et standards pédagogiques, ressources pédagogiques numériques

Rappels

Le Dublin Core n'est pas une norme. Il s'agit d'un ensemble de recommandations qui peut être qualifié de standard de fait. Le Dublin Core est né en 1995, lors d'une conférence tenue à Dublin (Ohio). Il s'agissait de produire une recommandation pour la description bibliographique des documents ou objets se trouvant sur le web, dans une bibliothèque numérique ou même ailleurs. Les acteurs de cette première conférence, qui regroupait des personnes issues de la communauté des bibliothèques, des informaticiens et du web, ont spécifié un ensemble de quinze éléments de base, une sorte de dénominateur commun de la description bibliographique. Le Dublin Core, dans son vocabulaire propre, parle d'éléments et de termes. Éléments, parce qu'il s'agit ici d'un monde issu de SGML, celui de HTML (et plus tard XML). Les termes d'éléments et d'attributs sont donc utilisés là où MARC parle par exemple de zones ou de sous-zones et où d'autres encore pourraient parler de « champs ».

Il est établi que les quinze éléments du Dublin Core sont « gravés dans le marbre » et que tout élément supplémentaire appartient aux extensions.

Les quinze éléments de base

Nom de l'élément en français	Décri	Nom de l'élément en anglais
Titre Sujet Description Source Langue Relation Couverture	contenu	Title Subject Description Source Language Relation Coverage
Créateur Éditeur Contributeur Droits	propriété intellectuelle	Creator Publisher Contributor Rights
Date Type Format Identifiant	instance particulière	Date Type Format Identifier

Le Dublin Core qualifié et les « termes »

Les quinze éléments de base sont considérés comme un dénominateur commun et pour la plupart des cas sont insuffisamment précis. Dès le départ, les éléments de base ont été étendus (ou précisés) par un ensemble d'autres termes, parfois appelés *qualifiers*. Deux classes de *qualifiers* sont reconnues :

- les « raffinements d'éléments » qui rendent plus spécifique le sens d'un élément ;
- et les « schémas d'encodage » ou vocabulaires contrôlés.

Les éléments supplémentaires et « raffinements d'éléments »

Le DCMI¹ a défini trois éléments supplémentaires et une trentaine de raffinements d'éléments.

Les éléments supplémentaires : **Audience**, **Provenance** et **RightsHolder** étendent l'ensemble des quinze éléments de base.

Les raffinements d'éléments (« element refinements »), dont voici pour information la liste complète², précisent le sens d'un élément existant :

¹ DCMI : Dublin Core Metadata Initiative, organisme responsable de la définition du Dublin Core.

² Voir <http://dublincore.org/documents/dcmi-terms/>

Element refinement	Element refined
Abstract	description
Access Rights	rights
Alternative	title
Available	date
Bibliographic Citation	identifiant
Conforms To	relation
Created	date
Date Accepted	date
Date Copyrighted	date
Date Submitted	date
Audience Education Level	audience
Extent	format
Has Format	relation
Has Part	relation
Has Version	relation
Is Format Of	relation
Is Part Of	relation
Is Replaced By	relation
Issued	date
Is Version Of	relation
License	rights
Mediator	audience
Medium	format
Modified	date
References	relation
Replaces	relation
Requires	relation
Spatial	coverage
Table Of Contents	description
Temporal	coverage
Valid	date

Les schémas d'encodage (vocabulaires contrôlés)

Chaque élément ou raffinement d'élément peut disposer d'un ou plusieurs schémas d'encodage. Le DCMI, soucieux de ne pas réinventer la roue, fait référence à des schémas existants et en a inventé un, DCMIType, pour décrire les types (logiques et non physiques) d'objets, d'autres, DCMIBox et DCMIPoint, pour délimiter géographiquement un objet ou lieu et DCMIPeriod pour situer un événement dans le temps.

Nom du schéma (vocabulaire)	« Element refined »
Box (DCMI)	spatial
DCMI Type	type
DDC (Dewey Decimal Classification)	subject
IMT (Internet media type)	format
ISO 3166 (norme iso des noms de pays)	spatial
ISO 639-2 (norme iso pour les codes de langues)	language
LCC (Library of Congress Classification)	subject
LCSH (Library of Congress Subject Headings)	subject
MeSH (Medical Subject Headings)	subject
Period (DCMI)	date / temporal
Point (DCMI)	spatial
RFC 1766 (codes langues- pays)	language
RFC 3066 (codes langues- pays, remplace RFC 1766)	language
TGN (Getty Thesaurus of Geographic Names)	spatial
UDC (Universal Decimal Classification)	subject
URI (Uniform Resource Identifier)	identifiant/ source/ relation
W3C-DTF (dates et heures)	date/temporal

Le vocabulaire DCMIType

Le vocabulaire DCMIType contient une liste de différents types de ressources :

Collection	Groupe de documents (ou ressources)
Dataset	Base de données, liste, tableau
Event	Événement ponctuel (conférence...)
Image	Image, photo, (tous formats, générique)
Interactive Resource	Objet qui demande l'interaction de l'utilisateur
Moving Image	Image animée, vidéo...
Physical Object	Un objet lui-même (et non son image)
Service	
Software	Logiciel disponible pour installation sur une machine
Still Image	
Text	

Les formats physiques

Le Dublin Core ne dépend pas d'une expression physique précise. On peut l'exprimer en HTML ou XML/RDF³ ou encore en XML, selon l'utilisation à laquelle il est destiné. Il existe trois recommandations officielles :

- **Expressing Simple Dublin Core in RDF/XML**
<http://dublincore.org/documents/2002/07/31/dcmes-xml/>
- **Guidelines for implementing Dublin Core in XML**
<http://dublincore.org/documents/2003/04/02/dc-xml-guidelines/>
- **Expressing Qualified Dublin Core in HTML/XHTML meta and link elements**
<http://dublincore.org/documents/dcq-html/>.

Et une recommandation proposée :

- **Expressing Qualified Dublin Core in RDF/XML**
<http://dublincore.org/documents/2002/05/15/dcq-rdf-xml/>

³ RDF : Resource Description Framework, recommandation du W3C.

Voici un exemple simple d'utilisation du Dublin Core dans un fichier HTML :

```
<html>
<head>
  <title> A Dirge </title>
  <link rel = "schema.DC" href = "http://purl.org/DC/elements/1.0/">
  <meta name = "DC.Title" content = "A Dirge">
  <meta name = "DC.Creator" content = "Shelley, Percy Bysshe">
  <meta name = "DC.Type" content = "poem">
  <meta name = "DC.Date" content = "1820">
  <meta name = "DC.Format" content = "text/html">
  <meta name = "DC.Language" content = "en">
</head>
<body>
  <pre>Rough wind, that moanest loudGrief too sad for song ;Wild wind, when sullen
cloudKnells all the night long ;Sad storm, whose tears are vain,Bare woods, whose
branches strain,Deep caves and dreary main, -WailJfor the world's wrong !</pre>
</body>
</html>
```

Voici un exemple d'encodage en HTML, mais en utilisant les raffinements d'éléments (seul l'entête du document HTML est montré).

```
<head profile="http://dublincore.org/documents/dcq-html/">
  <title>Expressing Dublin Core in HTML/XHTML meta and link elements</title>
  <link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
  <link rel="schema.DCTERMS" href="http://purl.org/dc/terms/" />
  <meta name="DC.title" lang="en" content="Expressing Dublin Core in HTML/XHTML
meta and link elements" />
  <meta name="DC.creator" content="Andy Powell, UKOLN, University of Bath" />
  <meta name="DCTERMS.issued" scheme="DCTERMS.W3CDTF" content="
2003-11-01" />
  <meta name="DC.identifier" scheme="DCTERMS.URI" content="
http://dublincore.org/documents/dcq-html/" />
  <link rel="DCTERMS.replaces" hreflang="en" href="
http://dublincore.org/documents/2000/08/15/dcq-html/" />
  <meta name="DCTERMS.abstract" content="This document describes how qualified
Dublin Core metadata can be encoded in HTML/XHTML &lt;meta&gt; elements" />
  <meta name="DC.format" scheme="DCTERMS.IMT" content="text/html" />
  <meta name="DC.type" scheme="DCTERMS.DCMIType" content="Text" />
</head>
```

L'exemple le plus courant d'expression du Dublin Core en XML est celui d'un serveur OAI⁴ (le schéma XML « oai_dc » se trouve sur le site de l'initiative OAI). Voici un exemple extrait d'un serveur OAI.

```
<metadata>
  <oai_dc:dc xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:oai_dc="
http://www.openarchives.org/OAI/2.0/oai_dc" xsi:schemaLocation="
http://www.openarchives.org/OAI/2.0/oai_dc http://www.openarchives.org/OAI/2.0/oai_dc.xsd"
xmlns:dc="http://purl.org/dc/elements/1.1/">
  <dc:creator>Moret, M.</dc:creator>
  <dc:title>Notice nécrologique : Émile Cotton (1872--1950)</dc:title>
  <dc:date>1949</dc:date>
  <dc:identifier>AIF_1949__1__1_0</dc:identifier>
  <dc:identifier>http://www.numdam.org/item?id=AIF_1949__1__1_0</dc:identifier>
  <dc:identifier>oai:numdam.org:AIF_1949__1__1_0</dc:identifier>
  <dc:identifier>citation: Ann. Inst. Fourier 1, 1-4 (1949)</dc:identifier>
  </oai_dc:dc>
</metadata>
```

Conclusion

Précisons pour terminer que :

- Le Dublin Core (ses extensions) évolue assez souvent. Les documents cités en référence peuvent changer.
- La description logique, qui n'est pas liée à un format physique, fait qu'il est possible d'inventer d'autres expressions de celle-ci sans en changer le sens.
- Une des premières idées des acteurs du Dublin Core était d'améliorer la qualité des résultats de recherche des « moissonneurs » web. Le Dublin Core était destiné au moissonnage par un moteur de recherche web tel que Google, Yahoo ou encore un moteur plus spécialisé. Dans un premier temps, c'est cette utilisation spécialisée qui a prévalu, car dans les faits les moteurs web généralistes ont ignoré le Dublin Core (ou plus précisément ont ignoré l'utilisation d'éléments META dans l'entête des documents HTML, à cause du détournement à des fins commerciales de leur utilisation). Aujourd'hui cependant, le projet « Google scholar » préconise l'utilisation du Dublin Core.
- Le Dublin Core est aujourd'hui moissonné efficacement avant tout par le biais d'OAI, à partir de serveurs se conformant au protocole OAIPMH.
- Actuellement, si tous les serveurs OAI utilisent (c'est obligatoire) le schéma oai_dc, bien peu se sont lancés dans l'utilisation de schémas « Dublin Core qualifié ». L'avenir dira si cette tendance s'inversera.

⁴ Voir « État de l'art en matière de normes et approche critique : du Marc au Dublin Core » dans ces actes.

- Les métadonnées « Dublin Core » sont très souvent exportées d'une base de données ou application autre, dont les métadonnées sont plus précises ou plus spécialisées.

Bibliographie

HILLMAN, Diane, *Guide d'utilisation du Dublin Core*

<http://www.bibl.ulaval.ca/DublinCore/usageguide-20000716fr.htm>

Dublin Core Metadata Initiative Documents : <http://dublincore.org/documents/>

DCMI Metadata Terms: <http://dublincore.org/documents/dcmi-terms/>

DCMI Encoding Guidelines <http://dublincore.org/resources/expressions/>

DCMI Schemas <http://dublincore.org/schemas/>