

Représentation de ressources II: Dublin Core

CERN-UNESCO School on Digital Libraries
Dakar, Nov 21-25, 2011

Annette Holtkamp
CERN

Dublin Core

- collection minimale d'éléments de metadonnées
 - conversion de MARC21 pas possible sans perte d'information
- pour décrire des objets digitales et physiques
- 1995 workshop a Dublin, Ohio
- Dublin Core Metadata Initiative (DCMI)
 - forum ouvert (W3C, bibliothécaires, techniciens)
 - principe de standards ouverts et interopérabilité

<http://dublincore.org>

Simple Dublin Core

- 15 éléments
- 3 groupes
 - contenu
 - version
 - propriété intellectuelle

Éléments: contenu

- titre (title)
 - le nom d'un objet
- sujet (subject)
 - mots-clés, souvent d'un vocabulaire contrôlé
- description
 - texte libre, résumé, table de matières...
- couverture (coverage)
 - e.g. géographique ou temporelle
- source
 - ressource reliée de laquelle la ressource décrite est dérivée
- relation
 - autre objet lié a cette ressource (e.g. traduction d'un livre)
- type
 - nature ou genre d'un objet (texte, video, événement)

Éléments: version

- format
 - information matérielle (médium physique, file format, dimensions d'un objet...)
- date
 - dates de création, publication, dernière modification...
- langue (language)
- identifiant (identifier)
 - référence sans équivoque (DOI, URI/URL, ISBN...)

Éléments: Propriété intellectuelle

- créateur (creator)
 - entité principalement responsable pour créer la ressource
 - auteur principale
- collaborateur (contributor)
 - auteur secondaire
- éditeur (publisher)
 - personne ou organisation responsable pour la publication
- droits (rights)
 - copyright, droits de propriété intellectuelle

Simple DC: Exemple

title: New Light on Dark Matter from the LHC

creator: Ellis, John

subject: Astrophysics

identifier: <http://inspirehep.net/record/875145>

description: The prospects for detecting a candidate supersymmetric dark matter particle at the LHC are reviewed...

date: 2010-11-02

type: text

Simple DC: Exemple en XML

<dc:title>New Light on Dark Matter from the LHC</dc:title>

<dc:creator>Ellis, John</dc:creator>

<dc:subject> Astrophysics</dc:subject>

<dc:identifier><http://inspirehep.net/record/875145>
</dc:identifier>

<dc:description>The prospects for detecting a candidate supersymmetric dark matter particle at the LHC are reviewed...</dc:description>

<dc:date>2010-11-02</dc:date>

<dc:type>text</dc:type>

Attributs de DC éléments

- optionnel
- répétable
- pas d'ordre défini
- recherche améliorée par vocabulaires contrôlés
- défini dans une recommandation du DCMI
- identifié par un URI dans l'espace de noms de DMCI

DCMI recommendation

- 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/>

Application: OAI Harvesting

- Open Archives Initiative (OAI):
 - cadre d'interopérabilité entre archives numériques facile à utiliser
 - pour établir services combinant des métadonnées de plusieurs sources
- Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)
 - infrastructure technique
 - mécanisme pour fournisseurs de métadonnées d'offrir leurs données
 - représentation en Dublin Core (XML) obligatoire
 - autres représentations facultatif
 - largement utilisé par bibliothèques et archives

OAI-PMH: exemple

http://export.arxiv.org/oai2?verb=GetRecord&identifier=oai:arXiv.org:1011.0077&metadataPrefix=oai_dc

<header>

<identifier>oai:arXiv.org:1010.0077</identifier>

<datestamp>2010-11-102</datestamp> </header>

<metadata>

<oai:dc_dc

xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/...>

<dc:title>New Light on Dark Matter from the LHC</dc:title>

<dc:creator>Ellis, John</dc:creator>

<dc:subject>: Astrophysics</dc:subject>

<dc:description>The prospects for detecting a ...</dc:description>

<dc:date>2010-10-30</dc:date>

<dc:identifier>http://inspirebeta.net/record/875145</dc:identifier>

</oai_dc:dc>

</metadata>

Dublin Core Qualifié

- depuis 2000
 - améliorer l'utilité pour applications locales
 - Dublin Core Simple
- plus
- 3 éléments supplémentaires
 - qualificatifs pour chaque élément

DC Qualifié : éléments additionaux

- Audience
 - pour qui la ressource est destinée
- Provenance
 - e.g. changements en propriétaire
- Propriétaires de droits (RightsHolder)
 - e.g. propriétaire de copyright

DC Qualifié : Qualificants

- raffinement d'éléments

- préciser ou borner le sens d'un élément

- description.tableOfContents, description.abstract

- date.modified

- schéma d'encodage

- pour interpréter les valeurs d'un élément

- e.g. vocabulaire contrôlé, notation formelle

- <meta name="DC.subject" scheme="UDC" content="539.12">

- <meta name="DC.date" scheme="W3CDTF" content="2004-01-01">

Principes de DC

- simplicité de créer et maintenir des notices
- sémantique globalement compris
 - createur: auteur, artiste...
- internationalization
 - versions en beaucoup de langues
- extensibilité
 - extensions spécifiques pour une discipline via des profiles d'application

Dublin Core Application Profile (DCAP)

cadre pour définir notices de metadonnées par delà de DC

- taillé sur mesure pour des applications spécifiques
- peut combiner des éléments de metadonnées de différents espaces de noms
 - espace de noms: collection spécifique d'éléments et qualificatifs, maintenu par une organisation autorisée
- utilisant vocabulaires globalement définies
- interopérabilité sémantique
- intégrable dans l'environnement de linked data

<http://dublincore.org/documents/profile-guidelines/>

DCAP: composants

- besoins fonctionnels
 - buts lesquelles une communauté veut atteindre avec cette application
- modèle de domaine
 - caractérise les types d'objets représentés par les métadonnées et leurs relations
- Description Set Profile + directives d'usage
 - éléments à utiliser et règles pour leur usage
- directives de syntaxe + formats des données

Exemple d'UCAD

Titre: La confiance en Dieu ou le poisson providentiel

Auteur(s): Maïga, Sigoungou

Mots-clés: fable
conte
foi

Date de publication: 0

Résumé: C'est l'histoire d'un captif qui avait une totale confiance en Dieu et qui fut miraculeusement libéré par un poisson.

URI/URL: <http://bca.ucad.sn/jspui/handle/01/386>

Collection(s) : Ajami

Exemple en DC

dc.contributor.author **Maïga, Sigoungou**
dc.date.accessioned **0**
dc.date.available **0**
dc.date.issued **0**
dc.identifier.uri **<http://bca.ucad.sn/jspui/handle/01/386>**
dc.description.abstract **C'est l'histoire d'un captif qui avait une totale confiance en Dieu. Menacé par le roi,**
dc.format.medium **Papier**
dc.language.iso **pulaar**
dc.subject **fable**
dc.subject **conte**
dc.subject **foi**
dc.subject.other **Foi**
dc.title **La confiance en Dieu ou le poisson providentiel**
dc.rights.holder **IFAN-CAD**
bc.cote **C.12 I Niger F.Viel**
bc.numSupportNumerique **MNAJ_A_00003**
bc.serviceDetenteur **LaboIslamologie**
bc.typeBiblio **manuscrit**
bc.collation **2 f. ; 21.5 cm X 17 cm**

Conversion DC - MARC21

<http://www.loc.gov/marc/dccross.html>

Description -- An account of the resource.

Unqualified:

520 ##\$a (Summary, etc. note)

Qualified:

Abstract: 520 ##\$a (Summary, etc. note)

TableofContents: 505 0#\$a (Formatted Contents Note)

Exemple en MARC21

0410_\$apulaar

100__\$aMaïga, Sigoungou

245__\$aLa confiance en Dieu ou le poisson providentiel

340__\$aPapier

520__\$aC'est l'histoire d'un captif qui avait une totale confiance en Dieu. Menacé par le roi, il fut sauvé miraculeusement par un poisson.

542__\$dIFAN-CAD

653__\$afable

653__\$aconte

653__\$afoi

85640\$u<http://bca.ucad.sn/jspui/handle/01/386>

Tableau de conversion: Experiments@Inspire

Field Mapping

MARC tag	subfield	SPIRES elem	notes
119	a	experiment	??? no appropriate marc field
419	a	other.name	???
520	a	abstract	description
		spokesperson	
046	q	date-proposed	not marc standard
046	r	date-approved	not marc standard
046	s	date-started	
046	t	date-completed	
245	a	title	
510	a	related	name of related exp
	0		recid of related exp
	w		type of relation: a predecessor, b successor
520	a	abstract	description

MODS

Metadata Object Description Schema

- MARC21 derivative
- subset of key MARC21 elements
- a few elements without MARC21 equivalent
- no field/subfield structure
- verbal instead of numerical tags
- uses XML

MODS example

```
<mods ID="xxx">  
  <titleInfo>  
    <title>New light... </title>  
  </titleInfo>  
  <name type="personal">  
    <namePart type="given">John</namePart>  
    <namePart type="family">Ellis</namePart>  
    <role>  
      <roleTerm type="text">author</roleTerm>  
    </role>  
  </name>  
</mods>
```