



# **Normes et standards du traitement documentaire des ressources électroniques**

---

**Antonin Benoît DIOUF**  
Université Gaston Berger (Saint-Louis)  
Bibliothèque centrale  
Service des acquisitions et du traitement

**Atelier COBESS du 30 juin 2007, Dakar, SENEGAL**

**Mobile : +221 776413523**  
**Mail : [antoninbenoit@gmail.com](mailto:antoninbenoit@gmail.com)**  
**Weblog : <http://louisar.wordpress.com>**

# PLAN

## ○ INTRODUCTION

- Considérations générales
  - ✓ Norme & Standard (Définition)
  - ✓ Traitement documentaire
  - ✓ Notion de Métadonnées

## ○ APPROCHE CLASSIQUE

- ISBD & ISBD (ER)
- Formats MARC

## ○ APPROCHE MOINS CLASSIQUE

- Dublin Core
  - ✓ Éléments
  - ✓ Langages de balisage (HTML, XML)
  - ✓ Dublin Core qualifié
  - ✓ Raffinements d'éléments
  - ✓ Schémas d'encodage
  - ✓ Espace de noms
  - ✓ Registre
  - ✓ Profils d'application

## ○ CONCLUSION

## ○ LIENS

## TIMING

**Présentation : 1 h**

**Cas pratiques &  
Discussions : 1 h**

**Total : 2h**



# INTRODUCTION

## Considérations générales

---

### Norme & Standard (Définition)

- **Normes** : caractère *officiel, légal* (« de jure »). Ensemble de règles de conformité, édictées et élaborés par des organismes officiels de normalisation (ISO, AFNOR...)
- **Standards** : caractère *consensuel* (« de facto »). Ensemble de recommandations émanant d'un groupe représentatif d'utilisateurs : *sociétés* ou *consortia* (XML : W3C,...) ; *groupements de professionnels* (Dublin Core : Dublin Core Metadata Initiative) ; *institutions* (MarcXML : Library of Congress)
- Parfois un standard s'officialise en norme (Dublin Core)



# Considérations générales

---

## Traitement documentaire

**Traditionnellement, il s'agit de :**

- “**décrire**” (“cataloguer”, “inventorier”) les documents
- “**contrôler**” certains points d'accès à la description des documents
- “**indexer**” (“classifier”) le contenu sémantique des documents



# Considérations générales

---

## Traitement documentaire

Dans le contexte de la présente intervention :

- isoler et articuler des **éléments de description**, dont

certains peuvent servir de **points d'accès aux documents**

**mais pas forcément  
« décrire » soi-même**

**« points d'accès aux *documents* »  
(numériques) et non à leur seule  
description**



# Considérations générales

---

## Traitement documentaire

- **Documents traditionnels**
  - **séparation** physique ***inévitabile*** entre le document et sa description
  - points "d'accès" (vedettes matières, auteurs, etc.) = accès à la ***description***, non au document même !
- **Documents électroniques**
  - Apparition au milieu des années 90 de la notion de ***métadonnées***



# Considérations générales

---

## Métadonnées

- Sens originel : données (numériques) fournissant des renseignements sur le paquet de données (numériques) auquel elles appartiennent
- Par extension et abus de langage : simplement “données sur des données” (même non numériques... ?). Donc données de fiches de catalogage sont aussi des métadonnées.
- → Même pour les ressources électroniques, les métadonnées *peuvent* être **physiquement séparées** des données sur lesquelles elles renseignent.



# Une définition parmi tant d'autres

---

## Métadonnées

- « Les métadonnées sont des informations structurées qui décrivent, expliquent, localisent ou encore facilitent la découverte, l'utilisation ou la gestion d'une ressource informationnelle. »
- NISO (National Information Standards Organization), *Understanding metadata*, 2004, ISBN 1-880124-62-9  
<http://www.niso.org/standards/resources/UnderstandingMetadata.pdf>





# Considérations générales

---

## Métadonnées

- Typologie des métadonnées :
  - Métadonnées **administratives**  
(comprenant notamment : **métadonnées de gestion des droits** et des accès, et **métadonnées de préservation**)
  - Métadonnées **structurelles**  
(explicitation des **relations entre les composants** de la ressource, liens entre ces composants). Facilitent la navigation et la présentation des ressources électroniques
  - Métadonnées **descriptives**  
(but : **découverte** et **identification** des ressources ; seront les seules traitées dans la présente intervention)
- Ces catégories n'ont pas de frontières bien définies et se recouvrent souvent partiellement



# Considérations générales

---

## Métadonnées descriptives

Elles servent à :

- La **Localisation**. Les métadonnées de description peuvent indiquer où est localisée une ressource d'information, que ce soit physiquement ou virtuellement.
- L'**Identification**. Les métadonnées de description peuvent distinguer une ressource d'information d'une autre sans décrire toute la collection de ressources.
- La **Recherche d'information**. Les métadonnées de description peuvent relier la requête d'un utilisateur sur un sujet donné à des ressources d'information accessibles sur le même sujet.



# Métadonnées descriptives

---

## ○ **Localisation** (exemples)

- **URL** (Uniform Resource Locator) est un exemple de métadonnée de description utilisée pour localiser des données. L'URL distingue les ressources web les unes des autres. Même si la plupart des URL semblent identifier, ou nommer, des ressources web, elles donnent en réalité le nom d'un fichier sur un serveur, ou d'un serveur lui-même sur lequel la ressource web est située.
- Information sur la classification, par exemple la classification décimale de Dewey. (Cotation)



# Métadonnées descriptives

---

- **Identification** (exemples)
  - **ISBN** (Numéro International Normalisé du Livre).
  - L'adresse **PURL** (Persistent Uniform Resource Locator). Une PURL identifie une ressource quelle que soit sa localisation sur le web et quel que soit l'endroit où cette ressource est déplacée sur le web.  
([http://purl.fao.org/2002\\_sustainable\\_development.zimbabwe.html](http://purl.fao.org/2002_sustainable_development.zimbabwe.html))



# Métadonnées descriptives

---

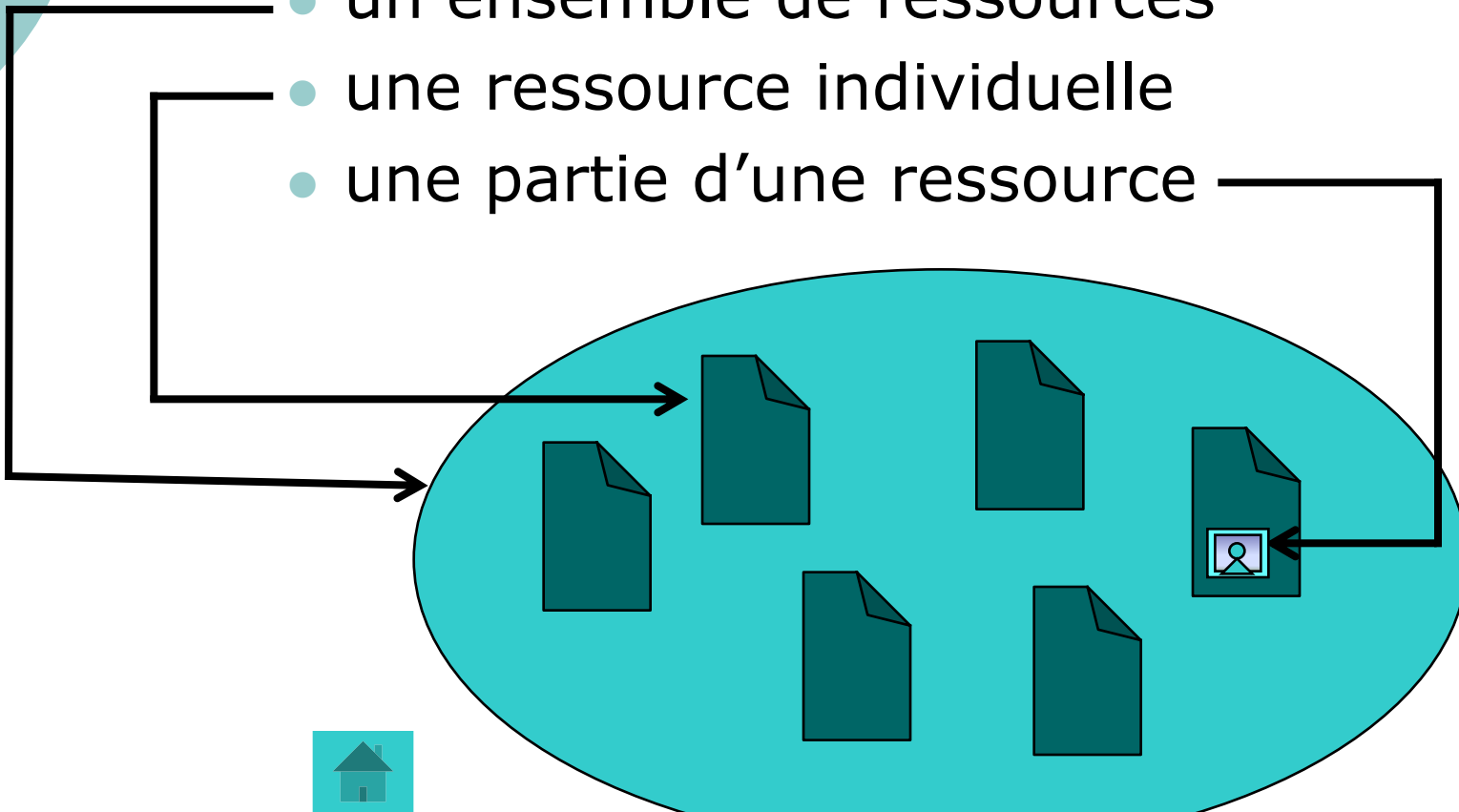
- **La recherche d'information** (exemples)
  - Mots matières ou vedettes-matières (d'un OPAC par exemple) qui renvoient à des documents traitant du même sujet.



# Considérations générales

## Métadonnées

- Peuvent concerner :
  - un ensemble de ressources
  - une ressource individuelle
  - une partie d'une ressource

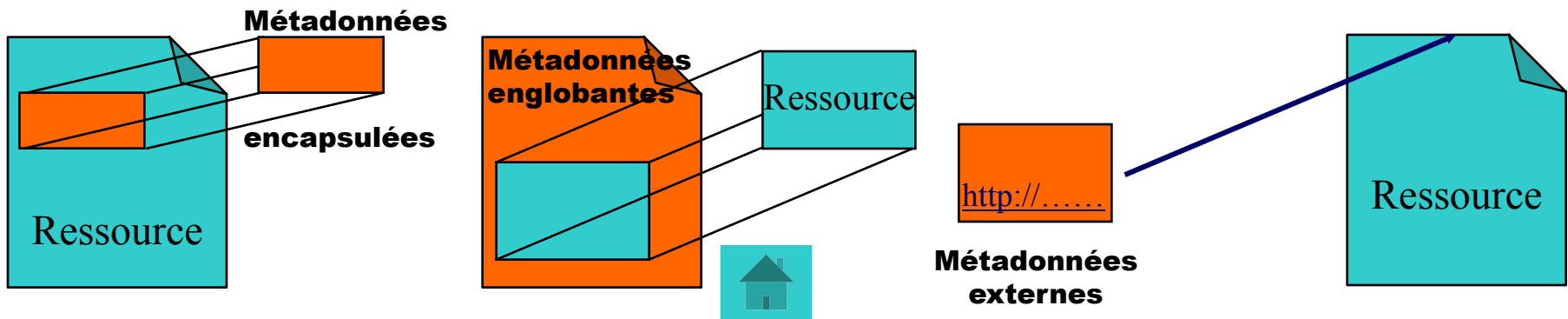


# Considérations générales

## Métadonnées : Où se trouvent-elles ?

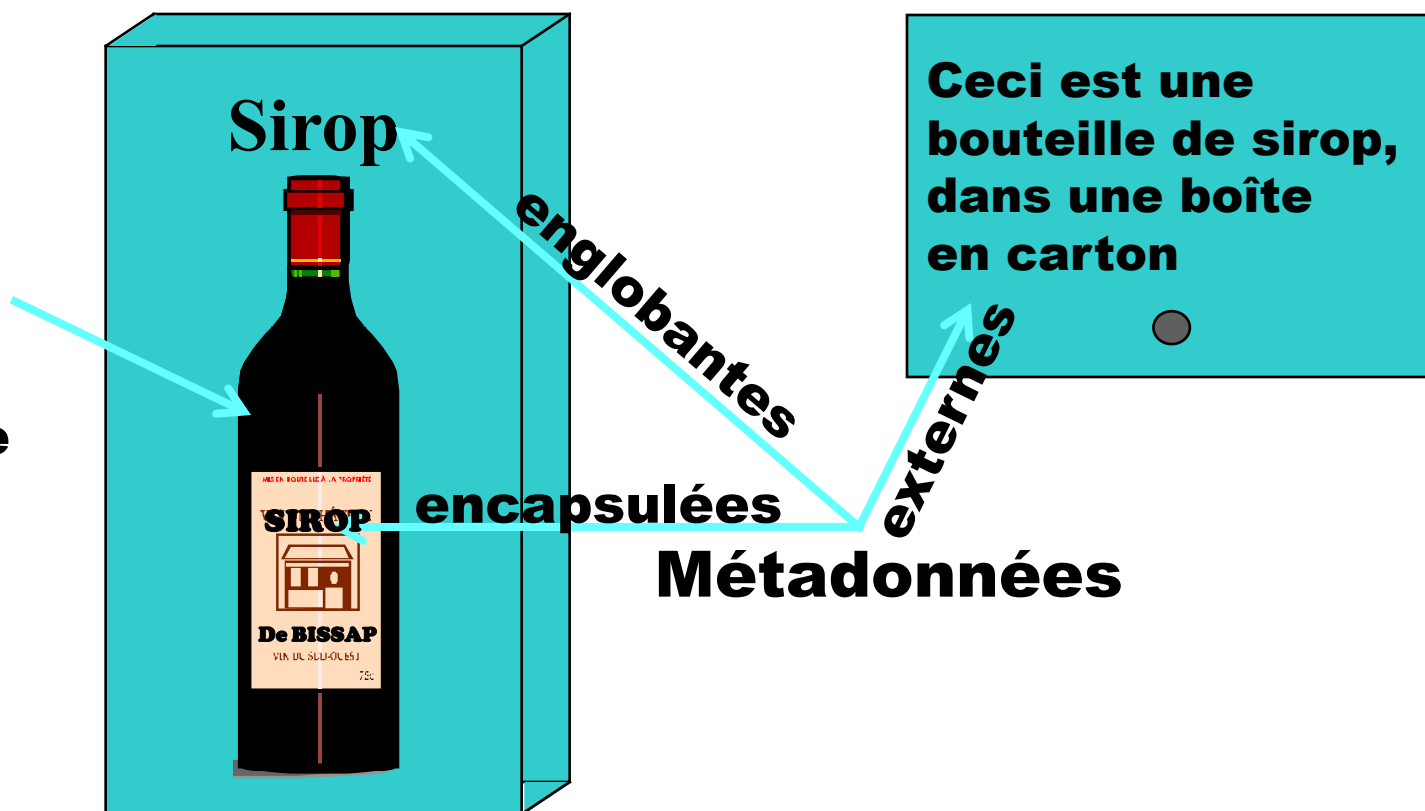
○ Elles peuvent être :

- **Encapsulées** (ex. : Dublin Core). Balises META des codes sources des pages ou sites web
- **Englobantes** (ex. : EAD = Encoded Archives Description)
- **Externes** (ex. : MARC) (séparées de la ressource qu'elles décrivent, peuvent être stockées dans des catalogues ou BDD et reliées à la ressource par une métadonnée de localisation comme **URL** ou **Cote**. Zone 856 UNIMARC)



# Une autre façon d'illustrer les 3 types d'emplacement des métadonnées...

**Données**  
Ou  
**Ressource**





# Considérations générales

---

## Métadonnées

### ☞ Remarque

- Pour les métadonnées, quelques-unes peuvent fournir des données sur le contenu ou sur le contexte d'une ressource.
  - Métadonnées de contenu : Mots matières, titre,...
  - Métadonnées de contexte : nom du créateur/auteur des données, la date et le lieu de création et les autres métadonnées administratives,...



# APPROCHE CLASSIQUE

---

○ **ISBD**

**&**

○ **MARC**



# Approche classique

---

## ISBD

- **ISBD** = **I**nternational **S**tandard **B**ibliographic **D**escriptions
- émanent de l'IFLA (Fédération internationale des associations de bibliothécaires et de bibliothèques)
- 1971 : ISBD(M) = le premier
- 1990 : ISBD(CF) CF = « Computer Files »
- **1997 : ISBD(ER)** ER = « **Electronic Resources** », d'où découle la norme AFNOR Z44-082 (Catalogage des ressources électroniques) de décembre 1999
- 2006 ou 2007 ? : ISBD(ER) révisé



# Approche classique

---

- Alors que les catalogues de bibliothèque commençaient à être disponibles électroniquement, l'ISBD initialement utilisé pour créer des catalogues papiers de bibliothèque a continué d'être utilisé comme préconisation pour la description des ressources électroniques.
- Donc création de l'ISBD (ER)



# Approche classique

---

## ISBD (ER)

- Couvre :
  - ressources électroniques *sur support* autonome (CD-ROMs...) ou *en ligne* (bases de données, sites Web...)
  - *interactives* ou *non interactives*
  - consistant en des *données* ou en des *programmes* (ou combinaison des deux)



# Approche classique

---

## ISBD (ER)

- Prévoit **deux modes de traitement** en fonction du type de ressource électronique (sur support / accès en ligne)
- toute ressource électronique en ligne est réputée « publiée » (Postulat)
- sont exclus : « jouets programmés, calculateurs et autres objets programmés »
- **notion de bibliothèque numérique *absente*** de l'ISBD(ER) : pas de distinction entre fonds numériques et reste des fonds, un seul catalogue.



# Approche classique

---

Particularités des ressources électroniques dans l'ISBD :

- **Zone 3 : Type et taille** de la ressource
  - zone facultative
  - Type : "Données textuelles", "Données cartographiques", "Données et programme" etc.
  - Taille : nombre de fichiers et/ou d'enregistrements et/ou d'octets et/ou d'instructions
- **Zone 4 : Adresse bibliographique**
  - Ressources électroniques en ligne : toutes réputées « publiées » → zone à remplir
- **Zone 5 : Description matérielle**
  - Essentiellement pour les ressources sur support ; facultatif pour les ressources en ligne
- **Zone 7 : Notes**
  - 1res notes = Configuration requise et Mode d'accès
- **URL pas explicitement prévus !**



# Approche classique

---

## Formats MARC

Nés du besoin d'un schéma **permettant à des bases de données bibliographiques de communiquer entre elles**, ce que ne permet pas l'ISBD (ER)

- Reposent tous sur la norme ISO 2709 "Format pour l'échange d'information"
  - norme adaptée à l'information bibliographique car elle **permet de gérer des zones** (et sous-zones)
    - de longueur variable
    - facultatives
    - répétables
  - définit la **structure** d'une notice
    - Label (zone fixe de 24 caractères)
    - Répertoire
    - Structure générique des zones





# Approche classique

---

## MARC 21

- Le format MARC 21 (Machine-Readable Cataloging ou Normes de Catalogage Lisibles par Machine) a été développé dans cette optique ; il met en forme les règles préconisées par l'ISBD de telle manière que les ordinateurs puissent les interpréter. Il fournit des champs pour stocker le titre de la ressource, l'individu ou l'organisme auteur, etc.
- Certaines données MARC 21 ne correspondent pas à une description bibliographique. Ces données permettent à différents systèmes bibliographiques d'échanger l'enregistrement.
- (Idem UNIMARC)

**Bloc d'identification**



**000** 01142cam 2200301 a 4500  
**001**92005291  
**003**DLC  
**005**19930521155141.9  
**008**920219s1993 caua j 000 0 eng  
**010|a** 92005291

# Approche classique

---

## UNIMARC

- créé par l'**IFLA**
- dans le cadre du programme du **CBU** (Contrôle bibliographique universel)
- pour sortir de la « babélisation » des formats nationaux (**USMARC** : format national aux USA, **CAN/MARC** : format national au Canada, **MARC 21** : fusion d'USMARC et de CAN/MARC, reconnue par l'IFLA comme format d'échange, **INTERMARC** utilisé par la Bibliothèque nationale de France etc...)
- première édition : **1977**
- format « pivot » des **échanges** d'information bibliographique
- adopté ensuite comme format de **saisie** par des bibliothèques



# Approche classique

---

## UNIMARC

Spécificités pour les ressources électroniques :

- **215** : **Description matérielle** : champ **omis** pour les ressources en ligne, remplacé par 230
- **230** : **Caractéristiques** de la ressource électronique (**obligatoire** pour les ressources en ligne)
- **336** : Note sur le **type** de la ressource électronique
- **337** : Note sur la **configuration requise**
- **856** : Localisation électronique et accès (**URL**)



# Approche classique

---

## ○ **Avantages :**

- **Pas de séparation** entre catalogue des ressources électroniques et catalogue des autres fonds
- **Richesse** et précision de la description

## ○ **Inconvénients majeurs :**

- **Inapplicable** à la totalité du Web !!!
- Pas en phase avec les autres développements en cours (TIC par exemple)
- **Isolement** des bibliothèques par rapport aux autres médiateurs de ressources (archives et musées entre autres)



# Approche moins classique « moderne »

---

- Dublin core (le seul traité dans cette intervention)
- MODS ;
- EAD ;
- MARCXML, etc.



# Dublin Core

---

- Créé en 1995 à Dublin, Ohio (USA), par OCLC et NCSA
- Maintenance assurée par Dublin Core Metadata Initiative (DCMI) et supervisée par OCLC – Office of Research & Special Projects
- Ancien standard devenu norme Z39.85 en 2001 par la National Information Standards Organization (NISO)
- Norme ISO 15836 depuis 2003 (DCMI = Agence de maintenance)

**NCSA** = National Center for Supercomputing Applications

**OCLC** = Online Computer Library Center



# Dublin Core

---

- Le Dublin Core inclut des éléments bibliographiques types comme le **titre**, l'**auteur**, l'**éditeur**, etc., mais aussi des éléments qui concernent plus particulièrement les ressources en réseau, par exemple le **type** et le **format** de la ressource, les **relations** entre ressources, et les **droits** de propriété intellectuelle. (Voir chapitres suivants)
- Le Dublin Core de base, utilisé pour décrire une ressource en réseau quelle qu'elle soit, a été enrichi de qualificatifs et d'autres éléments pour décrire des types particuliers de ressources. (Diapo 33 ; 51 et suivants)



# Dublin Core

---

## But initial :

- définir un ensemble de **métadonnées communes** à diverses communautés
- suffisamment **simples** pour que des non-spécialistes puissent les créer à n'importe quel point du cycle de vie de la ressource (créateur, propriétaire, gestionnaire, éditeur, utilisateur...)
- mais suffisamment **structurées** pour qu'elles puissent rendre les **moteurs de recherche** plus performants et donc faciliter la recherche et la récupération des ressources





# Dublin Core

---

Grands principes :

- 15 éléments
- tous optionnels
- tous répétables
- peuvent recevoir depuis 2000 des qualificatifs ( → « Dublin Core non qualifié » vs. « Dublin Core qualifié ») : raffinement d'éléments ou précision du schéma d'encodage

*Voir : (Diapo 51 et suivantes)*

- sont répartis en 3 groupes :



# Dublin Core : Éléments

- Tableau des 15 éléments (**Anglais=Français**)

CONTENU	VERSION	PROPRIETE INTELLECTUELLE
<b>Title=Titre</b> <b>Subject=Sujet</b> <b>Description=Description</b> <b>Coverage=Couverture</b> <b>Source=Source</b> <b>Relation=Relation</b> <b>Type=Type</b>	<b>Format=Format</b> <b>Date=Date</b> <b>Language=Langue</b> <b>Identifier=Identifiant</b>	<b>Creator=Créateur/Auteur</b> <b>Contributor=Collaborateur</b> <b>Publisher=Editeur</b> <b>Rights=Droits</b>



# Dublin Core : Éléments

---

- Ces éléments sont **optionnels**, peuvent être **répétés** et peuvent **apparaître dans n'importe quel ordre**. Leur application dépendra des besoins spécifiques du professionnel de l'information et de la ressource informationnelle pour laquelle il crée les métadonnées. Les éléments peuvent être représentés (Codage) dans différentes syntaxes, par exemple [HTML](#), [XML](#), XHTML, RDF.



# Dublin Core : Eléments

## Élément TITLE ou TITRE

**Titre du document** : il s'agit a priori du titre principal du document

○ **Exemple de codage avec HTML** :

utilisation des balises META ; NAME (se réfère au nom de l'élément utilisé. La valeur de « NAME » sera l'un des éléments du Dublin Core, écrit sous la forme suivante : « DC.[element] » ; « CONTENT » se réfère à la valeur effective de l'élément. La valeur de « CONTENT » sera celle que l'on a assignée à l'élément.

Exemple de création d'une métadonnée Titre pour le site web de l'EBAD.

```
<META NAME="DC.Title" CONTENT="Ecole de bibliothécaires, archivistes et documentalistes">
```

**Autres exemples :**

```
<meta name="DC.Title" content="Hamlet">
```

```
<meta name="DC.Title" content="L'aventure ambiguë">
```



# Dublin Core : Eléments

---

## Elément TITLE ou TITRE (Exemples de codage)

- Avec XML

```
<dc:title> Ecole de bibliothécaires,  
archivistes et documentalistes  
</dc:title>
```



# Dublin Core : Eléments

## Elément SUBJECT

- **Sujet et mots-clefs** : mots-clés ou mots matières généralement choisis dans un vocabulaire contrôlé, liste de termes conçue avant la création des enregistrements de métadonnées et utilisée de la même manière pour toutes les ressources de la collection (par exemple le thésaurus AGROVOC de la FAO ou le thésaurus de la United States National Agriculture Library, ou macrothésaurus de l'OCDE, Rameau, etc.). Indices de classification (Dewey, CDU, etc.)
- **Exemple :**  
<META NAME="DC.Subject"  
CONTENT="Agriculture"> <META  
NAME="DC.Subject" CONTENT="510">



# Dublin Core : Eléments

---

## Élément DESCRIPTION

- **Description du document** : description en texte libre de la ressource, telle qu'un résumé, et ne s'appuie pas sur un vocabulaire contrôlé. La balise Description du DC peut même être utilisée pour fournir une table des matières de la ressource décrite.
- Exemple :  
<META NAME="DC.Description"  
CONTENT="Voici la page d'accueil du site de l'EBAD">



# Dublin Core : Eléments

---

## Élément **COVERAGE**

- **Portée du document** : peut-être un domaine géographique, un laps de temps, ou le nom d'une entité administrative). Il est recommandé d'utiliser des mots-clés ou des mots matières généralement issus d'un type spécifique de vocabulaire contrôlé (des représentations normalisées de ces types de données, par exemple un vocabulaire relatif à la couverture spatiale ou temporelle de la ressource).

- Exemple :

<META NAME="DC.Coverage" CONTENT="Sénégal">

- ☞ **les éléments Subject (sujet), Description (description), et Coverage (couverture) sont utilisés pour accéder aux ressources. On peut considérer ces trois éléments comme étant indissociables.**





# Dublin Core : Eléments

---

## Eléments RELATION & SOURCE

- Les éléments **Source** et **Relation** du Dublin Core sont utilisés pour mettre en relation une ressource avec une autre.
- L'**élément Source** contient de l'information sur une deuxième ressource de laquelle **dérive** la ressource courante.
- L'**élément Relation** indique que la ressource est **reliée** d'une manière quelconque **à une autre ressource**. Par exemple, pour connecter le site principal de la FAO au site d'une division.
- Exemple : `<META NAME="DC.Relation" CONTENT="http://www.fao.org/ag/fr/default.htm">`



# Dublin Core : Eléments

---

## Élément TYPE

- **Nature ou genre du contenu** : L'élément **Type** décrit la nature de la ressource d'information. Il est recommandé de choisir les valeurs de Type au sein d'un vocabulaire contrôlé, par exemple la liste provisoire des types du Dublin Core (DCMItypes) : **Collection**, **Dataset** (données), **Event** (événement), **Image**, **Service**, **Software** (logiciel), **Sound** (son), **Text** (texte), **Interactive Resource** (ressource interactive)
- De nombreuses ressources web, comme les pages d'accueil, doivent être décrites avec plus d'un élément Type, selon le type d'informations qu'elles contiennent. La plupart des sites web contiennent du texte et des images :
- Exemples : `<META NAME="DC.Type" CONTENT="Text">`  
`<META NAME="DC.Type" CONTENT="Image">`



# Dublin Core : Eléments

---

## Élément FORMAT

- **Format** est le premier élément utilisé pour fournir de l'information sur la ressource elle-même.
- **Format du document** : Information sur le matériel ou l'équipement nécessaire pour utiliser ou accéder à la ressource. Choisir les valeurs de Format au sein d'un vocabulaire contrôlé, par exemple la liste des types de médias Internet (MIME) qui définit les formats de médias informatiques.
- Exemples : `<META NAME="DC.Format" CONTENT="text/html">` `<META NAME="DC.Format" CONTENT="image/gif">` `<META NAME="DC.Format" CONTENT="video/quicktime">`
- La ressource contient, en partie, du **texte** codé en **HTML** et des **images** au format **GIF**. Il inclut aussi des **vidéos** au format **Quicktime**. Dans la description des métadonnées, on répètera ainsi plusieurs fois la balise Format du DC.



# Dublin Core : Eléments

---

## Elément DATE

- La balise Date est utilisée pour indiquer les dates de publication, de création, de mise à disposition, ou de modification de la ressource. Il est recommandé de saisir la valeur de la date selon le format défini par l'[ISO 8601](#), sous la forme YYYY-MM-DD (AAAA-MM-JJ).
- Pour un site donné, tout ce que l'on connaît est l'année de la dernière mise à jour du site : 30 juin 2007.). La balise Date du DC doit ressembler à ceci :
- `<META NAME="DC.Date" CONTENT="2007-06-30">`



# Dublin Core : Eléments

---

## Élément LANGUAGE

- La balise Language (langue) est utilisée pour décrire la langue dans laquelle la ressource a été créée. Il est recommandé d'utiliser comme valeurs de l'élément Language celles définies par la [RFC 3066](#) (remplace [RFC 1766](#)) avec un code langue à deux caractères, éventuellement suivi d'un code pays à deux lettres. Par exemple, « en » pour l'anglais, « fr » pour le français, ou « en-uk » pour l'anglais utilisé au Royaume-Uni. Par exemple le site web à décrire est disponible en anglais, en français et en espagnol. On ajoutera donc trois balises Language dans l'enregistrement de métadonnées :
- `<META NAME="DC.Language" CONTENT="en"> <META NAME="DC.Language" CONTENT="fr"> <META NAME="DC.Language" CONTENT="es">`



# Dublin Core : Eléments

---

## Élément IDENTIFIER

- L'élément Identifier (identifiant) est utilisé pour **identifier de manière non ambiguë** une ressource informationnelle dans un contexte donné. Il est recommandé d'utiliser comme identifiant une chaîne de caractères ou un nombre conforme à un système d'identification officiel. Voici des exemples de systèmes officiels d'identification : l'Identifiant de Ressource Uniforme (URI/URL), l'Identificateur d'Objet Numérique (DOI) et le Numéro International Normalisé du Livre (ISBN).
- L'URL du site web de l'EBAD est <http://www.ebad.ucad.sn> ; la balise Identifier du DC sera donc de la forme :
- `<META NAME="DC.Identifier" CONTENT="http://www.ebad.ucad.sn">`



# Dublin Core : Eléments

---

## Eléments CREATOR, CONTRIBUTOR, PUBLISHER & RIGHTS

- **Le Dublin Core est aussi utilisé pour fournir de l'information sur les responsables de la ressource.**
- Le **creator** (auteur), le **contributor** (collaborateur), et le **publisher** (éditeur) peuvent être soit des individus, soit des organisations ; habituellement, l'éditeur est une organisation.
- L'élément **Rights** (droits) fournit de l'information sur les droits de propriété intellectuelle, les autres droits patrimoniaux et le copyright.



# Dublin Core : Éléments

## Éléments **CREATOR, CONTRIBUTOR, PUBLISHER & RIGHTS**

- **Creator** : Créateur du document (auteur principal) : nom de la personne, de l'organisation ou du service à l'origine de la rédaction du document.
- **Contributor** : Contributeur au document (auteur secondaire) : nom d'une personne, d'une organisation ou d'un service qui contribue ou a contribué à l'élaboration du document.
- **Publisher** : Editeur du document (Editeur) : nom de la personne, de l'organisation ou du service à l'origine de la publication du document.
- **Rights** : Droits relatifs à la ressource : permet de donner des informations sur le statut des droits du document, par exemple la présence d'un copyright, ou un lien vers le détenteur des droits. L'absence de cet élément ne présume pas que le document est libre de droits.





# Pourquoi ces 4 éléments ensemble ?

---

L'élément **Rights** concerne la plupart du temps les 3 autres éléments, c'est-à-dire que ces derniers détiennent le plus souvent les droits relatifs à une ressource.

- Dans notre exemple, l'EBAD est l'auteur, l'éditeur et détient aussi les droits patrimoniaux. Par conséquent, il est important d'utiliser la même valeur que celle utilisée dans le titre : "Ecole de bibliothécaires, archivistes et documentalistes". Les collaborateurs de la création du site web de l'EBAD n'apparaissent pas clairement. Il n'est donc pas nécessaire d'inclure une balise DC Contributor.
- On aura donc la description suivante :  
<META NAME="DC.Creator" CONTENT="Ecole de bibliothécaires, archivistes et documentalistes" >  
<META NAME="DC.Publisher" CONTENT="Ecole de bibliothécaires, archivistes et documentalistes" >  
<META NAME="DC.Rights" CONTENT="© Ecole de bibliothécaires, archivistes et documentalistes" >



# Dublin Core & Langages de balisage

---

## HTML / XML

- Les exemples que nous avons utilisé sont codés en HTML, mais la même description de métadonnées Dublin Core peut être codée en XML.
- Dans l'exemple qui suit, les mêmes éléments (Format, Date, Language et Identifiant) sont codés en utilisant les deux syntaxes :



# Dublin Core

---

## HTML

```
<META NAME="DC.Format" CONTENT="text/html">  
<META NAME="DC.Format" CONTENT="image/gif">  
<META NAME="DC.Format" CONTENT="video/quicktime">  
<META NAME="DC.Date" CONTENT="2004-06-30">  
<META NAME="DC.Language" CONTENT="en">  
<META NAME="DC.Language" CONTENT="fr">  
<META NAME="DC.Language" CONTENT="es">  
<META NAME="DC.Identifier" CONTENT="http://www.ebad.ucad.sn">
```

## XML

```
<dc:format>text/html</dc:format>  
<dc:format>image/gif</dc:format>  
<dc:format>video/quicktime</dc:format>  
<dc:date>2004-06-30</dc:date>  
<dc:language>en</dc:language>  
<dc:language>fr</dc:language>  
<dc:language>es</dc:language>  
<dc:identifier>http://www.ebad.ucad.sn</dc:identifier>
```

**Autres syntaxes de codage utilisables : RDF, XHTML**



# Quelques liens sur HTML et XML & XHTML

---

- **Beckett, David.** An **XML** encoding of simple Dublin core metadata (Encodage en XML des métadonnées simples du Dublin core)  
<http://dublincore.org/documents2001/04/11/dcmes-xml> (anglais)
- **Encoding Dublin core metadata in **HTML****  
<http://www.ietf.org/rfc/rfc2731.txt> (Anglais)
- **Mentions **Meta** concernant le contenu**  
<http://fr.selfhtml.org/html/entete/meta.htm> (Français, très bon site)
- **Passer du **HTML** au **XHTML****  
[http://openweb.eu.org/articles/html\\_au\\_xhtml](http://openweb.eu.org/articles/html_au_xhtml)
- **Le **XHTML****  
<http://openweb.eu.org/xhtml>



# Dublin Core qualifié

---

- L'ensemble des éléments de métadonnées Dublin Core (DC) fournit l'information nécessaire à la description de ressources comme les ouvrages, les articles et les pages web.
- Toutefois, comme diverses communautés appliquaient le DC différemment, des groupes de travail furent mis en place pour examiner la manière dont les éléments sont **qualifiés** plus précisément pour chaque application locale.
- Parmi ces groupes, on trouve le DC-Education, le [DC-Libraries](#), le DC-Government, qui étudient chacun les besoins de leur propre domaine.



# Dublin Core qualifié

---

Comment travaillent ces groupes ?

- Les groupes de travail soumettent des listes d'**éléments génériques** ou **spécifiques (Qualificatifs)** d'une discipline au comité de validation de l'Initiative de Métadonnées du Dublin Core (IMDC ou DCMI) qui évalue ces propositions et prend la décision finale.
- Cette procédure permet une évolution concertée de la Série d'Éléments de Métadonnées du Dublin Core (DCMES).



# Dublin Core qualifié

---

- Ces qualificatifs supplémentaires prennent la forme de :
  - **raffinements d'éléments**, ([voir liste](#))
  - **schémas d'encodage**. ([Voir liste](#))
- Ces deux sortes de qualificatifs permettent de décrire plus finement les éléments, comme le font les adjectifs dans notre langue naturelle.



# Dublin Core : Raffinements déléments

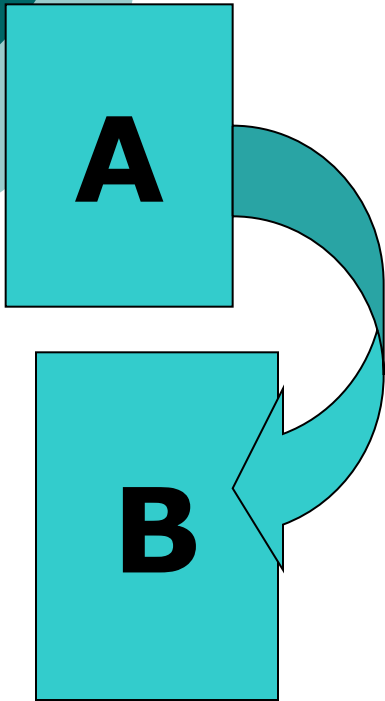
---

## Exemple de raffinement

- Imaginons qu'on veuille mettre à jour les métadonnées d'une ancienne version d'un article en ligne (A) avec des informations concernant la nouvelle version (B).
- On peut utiliser l'élément **DC relation**, défini comme « une référence à une ressource liée ».
- Le code HTML de la métadonnée de la ressource **A** serait :

```
<META NAME="DC.Relation" CONTENT="B">
```

- L'instruction ci-dessus indique que la ressource **A** est en relation avec une ressource **B**.
- Pour autant, cela ne donne pas d'information sur « comment » les deux ressources sont liées.



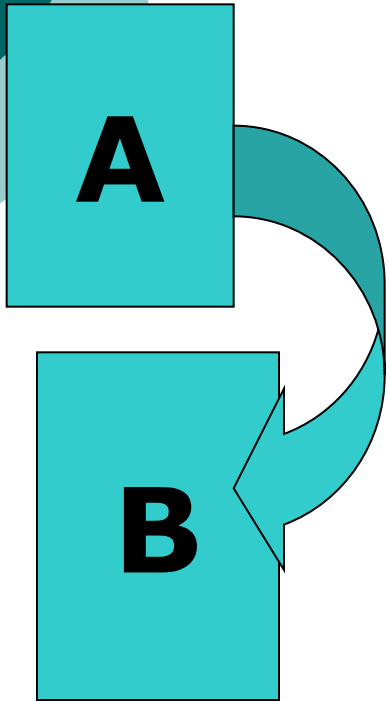


# Dublin Core : Raffinements déléments

- On souhaiterait informer l'utilisateur que la ressource **A** a été **remplacée** par la ressource **B**.
- Regarder la liste des qualificatifs pour Relation. (Diapo suivante)
- La paire de raffinements « **Replaces/isReplacedby** » (**remplace/est remplacé par**) semble la plus adéquate pour indiquer la relation « comment » !
- Le code HTML de la métadonnée pour la ressource **A** serait alors le suivant :

```
<META NAME="DC.Relation.isReplacedBy" CONTENT="B" >
```

- L'instruction ci-dessus indique deux choses :
  1. **A est reliée à B ;**
  2. **A est remplacée par B**
- Dans ce cas, le qualificatif « **isReplacedby** » précise la signification de l'élément « Relation » en spécifiant le type de relation.



# Raffinements de l'élément RELATION

---

- Is Version Of/ Has Version (est la version de/a pour version)
- Is Replaced By/Replaces (est remplacé par/remplace)
- Is Required By/Requires (est requise/requiert)
- Is Part Of/Has Part (fait partie de/a pour partie)
- Is Referenced By/References (est référencé par/référence)
- Is Format Of/Has Format (est une présentation de/a pour présentation)



# Dublin Core : Raffinements d'éléments

---

- En résumé, les raffinements d'éléments sont des qualificatifs permettant soit de **restreindre**, soit de **préciser** le **sens d'un élément**.
- Il est important de se souvenir qu'un élément raffiné a la même signification que l'élément non qualifié, mais avec une portée plus restreinte.
- Une application cliente ou un système qui ne pourrait interpréter le terme raffinant un élément devrait être capable d'ignorer le qualificatif et de traiter la valeur de la métadonnée comme s'il s'agissait d'un élément non qualifié.

Voir : <http://dublincore.org/usage/terms/dc/current-elements>

ou

[http://openweb.eu.org/articles/dublin\\_core](http://openweb.eu.org/articles/dublin_core) (français)



# Dublin Core : Schémas d'encodage

---

Autres types de qualificatifs qui

- **identifiant des schémas** qui facilitent l'interprétation de la valeur de l'élément (ou de ses raffinements)

Ils peuvent être soit

- des **vocabulaires contrôlés**, soit
- des **notations formelles**.



# Dublin Core : Schémas d'encodage

---

## EXEMPLE DE VOCABULAIRE CONTROLE

- L'instruction de métadonnée suivante permet d'interpréter la valeur « **Video games and teenagers** » comme une vedette matière du LCSH (Library of Congress Subject Headings ou Vedettes matières de la Bibliothèque du Congrès).

```
<META NAME="DC.Subject" SCHEME="LCSH" CONTENT="Video games and teenagers">
```

Schéma d'encodage



# Dublin Core : Schémas d'encodage

---

## EXEMPLE DE NOTATION FORMELLE

- Cette date est écrite selon le format YYYY-MM-DD (AAAA-MM-JJ), connu sous le nom de ISO 8601 (W3CDTF ou W3 Consortium Date and Time Formats). Ainsi, si on utilise ce format, l'instruction de métadonnée devrait être écrite en indiquant le schéma « W3CDTF ».



2007-06-20

```
<META NAME="DC.Date"  
  SCHEME="W3CDTF" CONTENT="2007-  
  06-20">
```




# Dublin Core : Schémas d'encodage

---

- En résumé, les schémas d'encodage facilitent l'interprétation de la valeur d'un élément.
- Même si une application ne comprend pas le schéma d'encodage, la valeur peut toujours être utile à un lecteur. Il pourra ainsi voir, comme dans l'exemple précédent, que l'expression « Video games and teenagers » est extraite des vedettes matières de la Bibliothèque du Congrès.
- Tableau indiquant certains schémas agréés par le Dublin Core pour l'élément Subject.(Diapo suivante)
- Une liste complète des schémas d'encodage supportés pour les autres éléments, ainsi que leurs définitions sont disponibles à :  
<http://dublincore.org/documents/2002/10/06/current-schemes> (en anglais) ou
- [http://openweb.eu.org/articles/dublin\\_core](http://openweb.eu.org/articles/dublin_core) (français)



# Dublin Core : Schémas d'encodage

<b>Élément DCMES</b>	<b>Schémas d'encodage</b>
Subject	<ul style="list-style-type: none"><li>- LCSH [Library of Congress Subject Headings] (vedettes matières de la Bibliothèque du Congrès)</li><li>- MeSH [Medical Subject Headings] (vocabulaire médical MeSH)</li><li>- DDC [Dewey Decimal Classification] (classification décimale de Dewey)</li><li>- LCC [Library of Congress Classification] (classification de la Bibliothèque du Congrès)</li><li>- UDC [Universal Decimal Classification] (CDU : classification décimale universelle)</li></ul>
	



# Dublin Core : Espace de noms

---

- Dans l'univers des métadonnées, les espaces de noms sont utilisés pour identifier les éléments « nouvellement définis » et leurs qualificatifs.
- Un espace de noms a une **autorité d'enregistrement**. Il s'agit de l'organisme autorisé à enregistrer les nouveaux éléments et qualificatifs dans un espace de noms donné.
- Toute organisation peut créer son propre espace de noms à partir du moment où elle s'engage à sa maintenance.

**L'IMDC est l'autorité d'enregistrement pour les éléments et qualificatifs du Dublin Core.**



# Dublin Core : Espace de noms

---

- Ces **nouveaux éléments**, **raffinements** et **schémas d'encodage** permettent de rendre la signification des éléments du Dublin Core plus claire et plus **spécifique au domaine couvert.**



# Dublin Core : Espace de noms

---

## Comment les reconnaître ?

- Les intitulés des espaces de noms sont indiqués par la **chaîne de caractères** située devant l'**élément**.

### Exemple

```
<rdf:Description rdf:about="http://www.ebad.ucad.sn">  
<dc:title> EBAD : Ecole de Bibliothécaires, Archivistes et documentalistes,  
  Université Cheikh Anta Diop, Dakar, Sénégal </dc:title>  
<dc:creator> EBAD </dc:creator> <dc:date> 2007-05-19 </dc:date>  
</rdf:Description>
```

Cet exemple d'enregistrement de métadonnées est codé en **XML** et **RDF** et **Dublin Core** sont les espaces de noms

- **Namespaces in XML 1.0** (Second Edition) : W3C Recommendation 16 August 2006 : <http://www.w3.org/TR/REC-xml-names/>
- **Namespaces in XML 1.1'** (4 février 2004) en version française. [2006/11/24]. <http://www.yoyodesign.org/doc/w3c/xml-names11>



# Dublin Core : Registre

---

## DC Registry

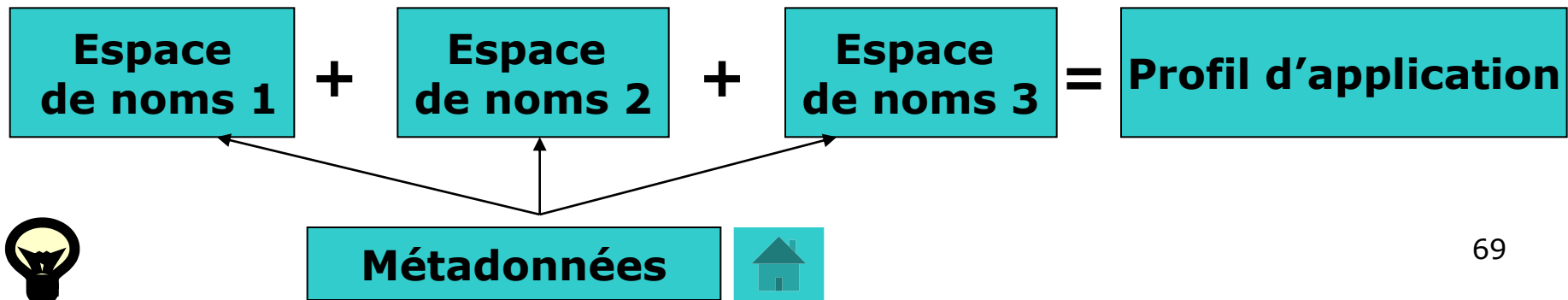
### Créé par l'IMDC

- contient **tous** les éléments et qualificatifs du Dublin Core.
- Un registre de métadonnées contient la définition des termes (éléments, raffinements d'éléments et schémas d'encodage), informe sur les nouveaux termes disponibles, contrôle les changements de version des termes, fait la promotion des termes en vue de leur réutilisation.
- Ces registres ont pour objectif de fournir une vue unique des éléments actuellement disponibles et de leurs définitions.
- <http://dublincore.org/dcregistry/> (registre Dublin core en anglais)



# Dublin Core : Profils d'application

- Si on a besoin d'éléments de métadonnées du Dublin Core pour décrire correctement des ressources, on peut consulter le registre du Dublin Core (DC Registry) qui contient les éléments déjà déclarés et choisir les éléments qui répondent à ses besoins (ne pas oublier que tous les éléments sont **optionnels**).
- L'emprunt d'éléments issus de différents espaces de noms (dans le DC ou autre espace de noms) aboutit à la création d'un **profil d'application**.
- De cette façon, on économise le temps que l'on aura passé à créer son modèle de données



# Dublin Core : Profils d'application

---

- Existe un profil d'application spécialement destiné aux bibliothèques, « **DC-Libraries** »  
Voir :  
<<http://dublincore.org/documents/2004/09/10/library-application-profile/>>
- Ajout des éléments : **Audience**, **Edition**, **Location**
- Possibilité de qualifier les éléments Contributor et Publisher au moyen d'un **rôle**
- Possibilité de qualifier l'élément **Date** par son schéma d'encodage (ISO 8601 recommandée)
- Etc.



# CONCLUSION

---

- L'objectif du Dublin Core et des autres standards de métadonnées est de promouvoir l'**interopérabilité** par la **réutilisation** d'un ensemble de métadonnées communes. Cela facilite l'échange et le partage d'information dans un environnement en réseau.
- Pour se comprendre mutuellement, on doit utiliser les mêmes balises de métadonnées, ou au moins quelques métadonnées communes.
- Par conséquent : lorsque cela est possible, utiliser un standard de métadonnées existant et largement accepté comme le DC (méthode de codage le plus « populaire » utilisé sur le web).
- On peut sans peine inclure de telles métadonnées dans ses propres pages : il suffit de faire figurer les balises où elles sont encodées dans l'en-tête HTML de source d'une ressource. Grâce à ces informations, des systèmes automatiques peuvent réellement tirer parti des pages, car elles contiennent leur propre descriptif synthétique.

**Plus grand sera le nombre de communautés adoptant un standard unique, plus forte sera l'interopérabilité.**



# LIENS : (générateurs de métadonnées Dublin Core) (support des exercices)

<http://dublincore.org/tools/> (**source**) [Consultée le 20 juin 2007]

- <http://www.fao.org/aims/tools/metamaker.jsp?lang=fr> . (Cet outil fournit la possibilité de créer des AgMES/DC métadonnées pour encoder des pages web.)
- **MKDoc** <http://www.mkdoc.org/> (MKDoc Ltd., United Kingdom) **Outil gratuit à télécharger et installer localement**
- **Mozilla Firefox Dublin Core Viewer Extension** <http://www.splintered.co.uk/experiments/73/> (Splintered, United Kingdom) **Téléchargement gratuit et installation locale (Mozilla Firefox 1.5 - 3.0a1 uniquement)**
- **UKOLN**. Dublin Core metadata editor. <http://www.ukoln.ac.uk/cgi-bin/dcdot.pl> (Générateur de métadonnées en Dublin core) ou <http://www.ukoln.ac.uk/metadata/dcdot/>
- **Editor-Converter Dublin Core metadata** (outil en ligne gratuit, éd. Par Kirovohrad Regional Universal Research Library, Ukraine ) <http://www.library.kr.ua/dc/dcreditunie.html> (Générateur de métadonnées DC et convertisseur vers UNIMARC)
- **DescribeThis** <http://www.describethis.com/> (eSand, Spain) Générateur automatique de métadonnées DC (Gratuit)
- Etc.





# LIENS

---

**Teasdale G. (trad.).** 2001. Guide d'utilisation du Dublin Core. [Consultée le 16 juin 2007].

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

**Jacquet, Christophe.** Métadonnées et Dublin Core. [Consultée le 21 juin 2007] (Complet et précis sûrement l'un des meilleurs liens sur la question)

[http://openweb.eu.org/articles/dublin\\_core/](http://openweb.eu.org/articles/dublin_core/) (Liste de tous les éléments et de leurs qualificatifs (raffinements et schémas) en français)

**IMotep.** Dublin Core. Métadonnées. [Consultée le 15 juin 2007].

<http://metadata.imotep.com/dc-fr/>

**Verscoustre A. M. (trad.).** 2002. Eléments de métadonnées du Dublin Core.[Consultée le 16 juin 2007].

**Cherhal, Élisabeth.** Le Dublin Core (DC). *Babel - edit -*, L'indexation des ressources pédagogiques numériques. ENSSIB - janvier 2006 [en ligne] <http://babel.enssib.fr/document.php?id=145> (\*)



# LIENS

---

Principles and operation of the Dublin Core Metadata Initiative. [Consultée le 12 juin 2007] (en anglais).  
<http://www.dublincore.org/about/organization/>

**ISO** (Organisation Internationale de Normalisation).  
[Consultée le 13 juin 2007].  
<http://www.iso.ch/iso/fr/ISOOnline.frontpage>

**BnF**. Professionnels. [Consultée le 25 novembre 2006].  
<http://www.bnf.fr/pages/zNavigat/frame/infopro.htm>

The Dublin Core Metadata Element Set. ANSI/NISO Z39.85-200. [Consultée le 15 juin 2007] (en anglais).  
<http://www.niso.org/standards/resources/Z39-85.pdf>



# LIENS

---

**DCMI Education Working Group.** [Consultée le 15 juin 2007]  
(en anglais).  
<http://dublincore.org/groups/educatio/>

**BnF** (Bibliothèque nationale de France). [Normes, formats, modélisation]. [Consultée le 25 novembre 2006] (en anglais).  
<http://www.bnf.fr/pages/zNavigat/frame/infopro.htm>

**Deschatelets J.** 2006. Dossier sur les métadonnées. [Consultée le 19 juin 2007].  
[http://bibliodoc.francophonie.org/article.php3?id\\_article=172](http://bibliodoc.francophonie.org/article.php3?id_article=172)

**Noy N. F. MCGuinness D. L. Angjeli A. (trad.).** Développement d'une ontologie 101 : guide pour la création de votre première ontologie. [Consultée le 25 novembre 2006].  
<http://www.bnf.fr/PAGES/infopro/normes/pdf/no-DevOnto.pdf>

Site de traductions du W3C :  
<http://www.w3.org/2005/11/Translations/TranslationSearch.html>





**FIN**

---

**Merci**

**de votre**

**aimable attention**