

AXE 3: Description des documents numériques (Dublin Core)

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
Title = Titre Subject = Sujet Description = Description Coverage = Couverture Source = Source Relation = Relation Type = Type	Format = Format Date = Date Language = Langue Identifier = Identifiant	Creator = Créateur / Auteur Contributor = Collaborateur Publisher = Editeur Rights = Droits

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 : Éléments

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 : Éléments

Élément TITLE ou TITRE (Exemples de codage)

- Avec XML

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

Dublin Core : Éléments

Élément 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 : Éléments

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 : Éléments

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 : Éléments

Éléments 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 : Éléments

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 : Éléments

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 : Éléments

Élément 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 : Éléments

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 : Éléments

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 : Éléments

Éléments 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 Identifier) 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](http://dublincore.org/documents2001/04/11/dcmes-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)**
<http://www.ietf.org/rfc/rfc2731.txt> (Anglais)
- **Mentions [Meta](http://fr.selfhtml.org/html/entete/meta.htm) 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
- **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 ?

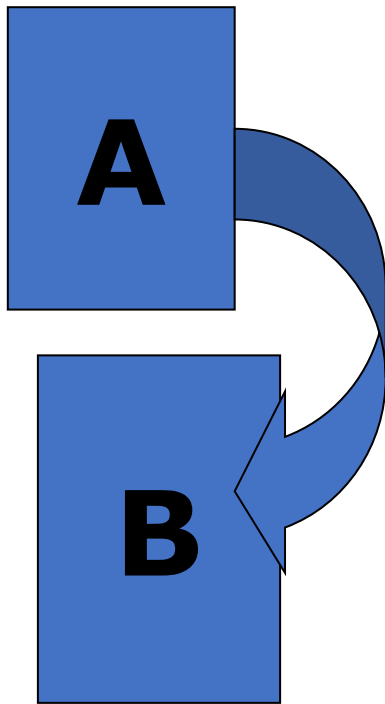
- Les groupes de travail soumettent des listes d'**éléments génériques** ou **spécifiques** (**Qualificatifs**) d'une discipline au comité de validation de l'Initiative de Métadonnées du Dublin Core (IMDC ou DCMI) qui évalue ces propositions et prend la décision finale.
- Cette procédure permet une évolution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



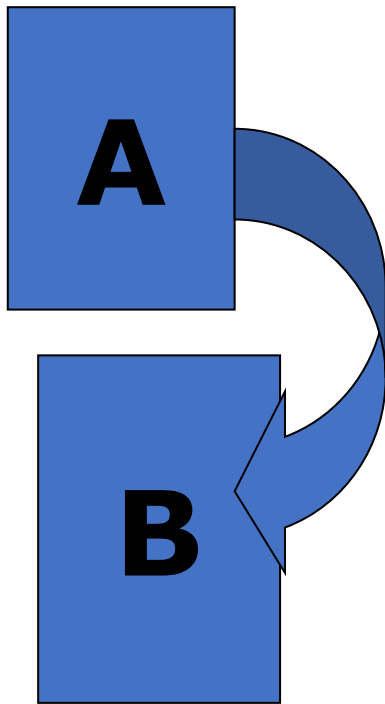
- Imaginons qu'on veuille mettre à jour les métadonnées d'une ancienne version d'un article en ligne (A) avec des informations concernant la nouvelle version (B).
- On peut utiliser l'élément **DC relation**, défini comme « une référence à une ressource liée ».

- 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 (français)

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.