

Métadonnées & Dublin Core

Richard Walter

Métadonnées

Une métadonnée (mot composé du préfixe grec *meta*) est une donnée servant à définir ou décrire une autre donnée quel que soit son support (papier ou électronique).

Un exemple type est d'associer à une donnée la date à laquelle elle a été produite ou enregistrée, ou à une photo les coordonnées GPS du lieu où elle a été prise.

(Wikipédia)

=> Une multitude de types de métadonnées <=

Métadonnées

Possibilités de définitions ?

⇒ Données à propos de données

⇒ Ensemble structuré d'informations servant à décrire une ressource

⇒ Fiches, notices (en langue vernaculaire)

« Les métadonnées sont aux données ce que l'étiquette est au médicament »

Métadonnées

Exemple de métadonnées :

- Forme de la bouteille
- Etiquette
- Capsule représentation des droits
 - Droits d'assises

La donnée

Couleur verte : Vins tranquilles ou mousseux répondant à la définition des vins de qualité produits dans des régions déterminées (VQPRD) qui comprennent, notamment, les appellations d'origine contrôlée (AOC) et les vins délimités de qualité supérieure (VDQS).



Métadonnées

Exemple de métadonnées :

Étiquette

- Obligatoire (dans la langue du pays d'origine)
 - *Catégorie (AOC, IGP, Vin de table)*
 - *Nom de l'appellation (Côtes du Rhône)*
 - *Degré d'alcool*
 - *Embouteilleur*
- Facultatif
 - *Année de récolte*
 - *Cépages*
 - *Taux de sucre*



Métadonnées

Règles d'utilisation de ces métadonnées :

- Règlements communautaires
(OCM unique + règlements d'application)
- Réglementation nationale
Imposition du logo femme enceinte



+ Cahiers des charges des appellations

Forme de la bouteille

(Vins d'Alsace dans une bouteille « Vins du Rhin »)

Vendanges tardives en Vins d'Alsace, cépages (uniquement 5)
et millésimes sont obligatoires



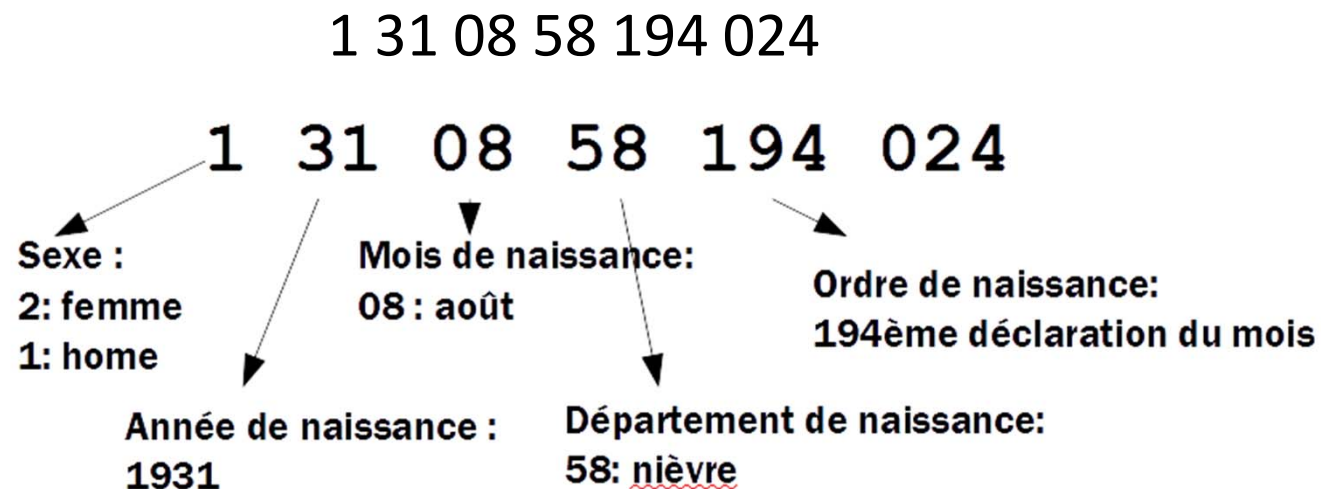
Métadonnées

Classement des différents types de métadonnées en fonction :

- de ce qu'elles décrivent (le contenu)
- de la façon dont elles sont créées (leur provenance)
- du moment où on les crée (leur historique)
- de l'endroit où on les trouve (leur localisation)
- de l'aspect qu'elles ont (leur forme)
- de l'usage qu'on en fait (leur objectif)

Métadonnées pour quoi faire ?

- Pour comprendre l'information



Métadonnées pour quoi faire ?

- Pour retrouver une ressource dans un ensemble très vaste...



Charlie est un jeune homme grand et mince aux cheveux bruns portant des lunettes de vue. Il est habillé avec un bonnet et un sweat shirt à rayures blanches et rouges, un pantalon bleu et des chaussures marrons.

des chaussures marrons

Métadonnées pour quoi faire ?

- Pour gérer des aspects administratifs
(droits, suivi des évolutions du document, suivi des évolutions des métadonnées)
- Pour piloter les migrations vers d'autres formats ou codages

⇒ Très (trop) nombreux formats d'expression des métadonnées

⇒ Choisir un format de métadonnées

À privilégier : « les spécifications publiques, sans entraves juridiques, normalisées, disposant de plusieurs outils de lecture et de conversion vers d'autres formats et codages. »

Dublin Core



Type de document

Livres 3

Image

Presse

Partitions

Cartes

+

Auteurs

Paris (

Paris,

Binete

Dessir

La Ma

Carto

Casin

+

Dates

XIXe s

XXe s

XVIIIe

XXIe s

XVIIe

+

Langues

Franç

fr (4 5

franc

Latin

Franç

+

Thèmes



Sommier des biens nationaux de la ville de Paris : conservé aux archives de la Seine. De la première à la quatrième municipalité (articles 1 à 1740) / publié par H. Moran & L. Lazard - L. Cerf (Paris)-1920 [Informations détaillées](#)

Livre en mode image et en mode texte, recherche plein texte disponible

Sujet: Paris (France) -- Administration -- Histoire Auteur : Paris Éditeur : L. Cerf (Paris)

Contributeur: Archives de Paris. Éditeur scientifique Bibliothèque historique de la Ville de Paris,

Ville de Paris. Carte d'ensemble - Avril frères (Paris)-18.. [Informations détaillées](#)

Notice complète

Titre : Ville de Paris. Carte d'ensemble

Auteur : Avril Frères (Paris). Auteur du texte [Ne voir que les résultats de cet auteur](#)

Éditeur : Avril frères (Paris)

Date d'édition : 18..

Sujet : Paris

Type : document cartographique, carte, image fixe

Langue : Français

Format : 1 fll ; 59 x 41 cm

Format : image/jpeg

Droits : domaine public

Identifiant : ark:/12148/btv1b8444443j

Source : Bibliothèque nationale de France, GED-5481

Relation : <http://catalogue.bnf.fr/ark:/12148/cb407249952>

Couverture : France - Île-de-France - Paris

Provenance : bnf.fr

Description : Échelle(s) : 1:25 000

Date de mise en ligne : 25/07/2011

[Fermer la notice complète](#)

Cartes et plans

Sujet: Paris Auteur : Avril Frères (Paris). Auteur du texte Éditeur : Avril frères (Paris)

Légendes: Ville de Paris. Carte d'ensemble Couverture: France - Île-de-France - Paris

Dublin Core

Initiative pour l'échange de documents

- *« Le Dublin Core est un schéma de métadonnées générique qui permet de décrire des ressources numériques ou physiques et d'établir des relations avec d'autres ressources. Il comprend officiellement 15 éléments de description formels (titre, créateur, éditeur), intellectuels (sujet, description, langue...) et relatifs à la propriété intellectuelle. »*
- http://fr.wikipedia.org/wiki/Dublin_Core



Dublin Core

- Des métadonnées « généralistes »

Dublin Core Metadata Initiative, 1995

Norme ISO-15836:2003

<http://dublincore.org>

Pour toute type de ressources

Utile pour l'interopérabilité



Dublin Core

« Une description bibliographique simplifiée »

- 15 éléments
- Tous les éléments sont facultatifs
- Tous les éléments peuvent être répétés
- Il existe des syntaxes d'écriture en XML, en HTML, en RDF



Dublin Core

Les 15 descripteurs (1/3)

- **Title** : « nom habituellement donné à la ressource ».
- **Subject** : « description du contenu « intellectuel ». Mots-clés, classification. Un vocabulaire spécifique ou contrôlé est souhaitable, pour l'interopérabilité...
- **Coverage** : couverture temporelle ou spatiale de la ressource. Exemples : ville, pays, époque, etc. Pas de vocabulaire spécifique conseillé ici, mais ce serait mieux (idem).
- **Description** : description du contenu « intellectuel » sous forme de texte libre.
- **Language** : langue(s) de la ressource. Code ISO conseillé.



Dublin Core



Les 15 descripteurs (2/3)

- **Source** : ressource dont dérive la ressource actuelle, totalement ou partiellement.
- **Creator** : « responsable principal » de la création de l'œuvre. Statut à débattre.
- **Type** : type intellectuel du document. Remplir en grandes catégories ou autres catégories connues de l'organisation ; vocabulaire contrôlé conseillé.
- **Contributor** : nom d'un autre agent (organisme, personne) qui a contribué à la ressource.
- **Publisher** : responsable de la publication de la ressource dans sa forme actuelle.

Dublin Core



Les 15 descripteurs (3/3)

- **Rights** : propriété, responsabilité, droits d'usage. Au choix.
- Exemple : une des licences Creative Commons (ex : Paternité-Pas d'Utilisation Commerciale-Pas de Modification 2.0)
- **Date** : une ou des dates marquantes dans le cycle de vie de la ressource. A débattre.
- **Format** : format informatique ou physique de la ressource décrite.
- **Identifier** : un identifiant non ambigu de la ressource.
- **Relation** : indique une ressource liée à la ressource qu'on décrit, en termes de format, version, partie/appartenance, etc.

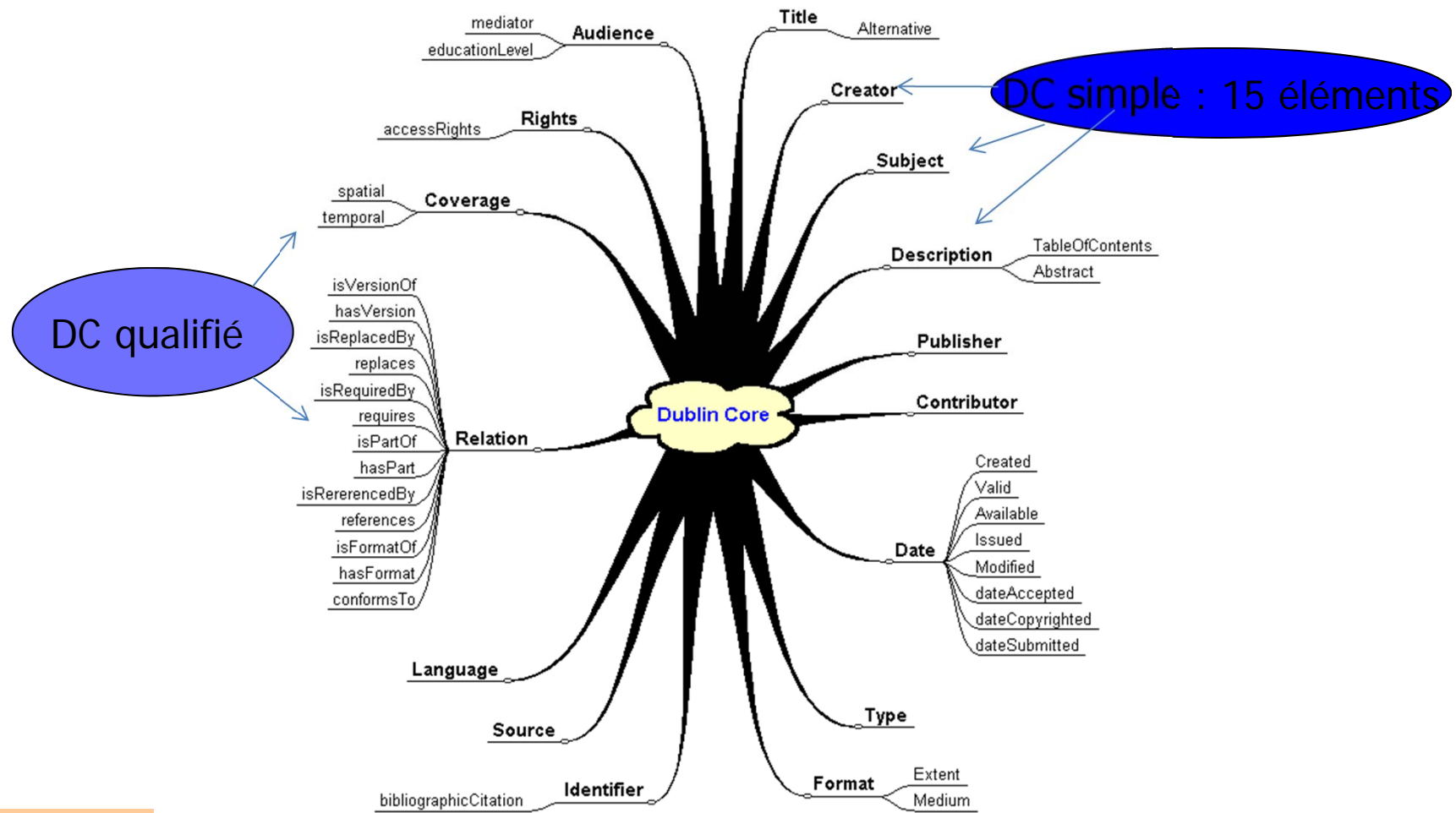
Dublin Core



Le Dublin core « qualifié »

La signification des éléments peut être précisée par des
« refinements »

Dublin Core



<http://ganesha.fr/index.php?post/2008/03/31/Dublin-Core>

Dublin Core



Quelques bonnes pratiques

- Pour beaucoup d'éléments : préciser la langue dans laquelle est le contenu de l'élément
- Il ne doit y avoir qu'un contenu par élément. Si plusieurs contenus alors dupliquer l'élément autant que nécessaire

Quand cela est possible :

- Utiliser une syntaxe ou un vocabulaire contrôlé normalisé

⇒ On peut/doit aller plus loin

Dublin Core

Avantages et inconvénients

Tous les éléments ne sont pas obligatoires

L'utilisation des raffinements est facultative

C'est un format strict mais qui laisse de la liberté

⇒ Nécessité de bien définir les champs Dublin Core

⇒ Selon le type d'utilisation des documents



Conclusion sur les métadonnées

- Il y a un continuum entre données et métadonnées

Une évidence :

« Plus les métadonnées sont riches et précises plus les données seront facilement exploitables et réutilisables »

Mon projet et les métadonnées

Adaptation aux standards

- Se confronter à des syntaxes spécifiques, « versionnables »
- Se confronter à des formulaires « réducteurs »

Une difficulté :

Gérer une tensions « habituelle » entre :

- Sophistication de la description *versus* catégorisation des cases du formulaire (forcément réducteur)
- Indexation pour la sauvegarde *versus* indexation pour l'exploitation