

## 2. Notion de métadonnées

### 2.1. Que sont les métadonnées?

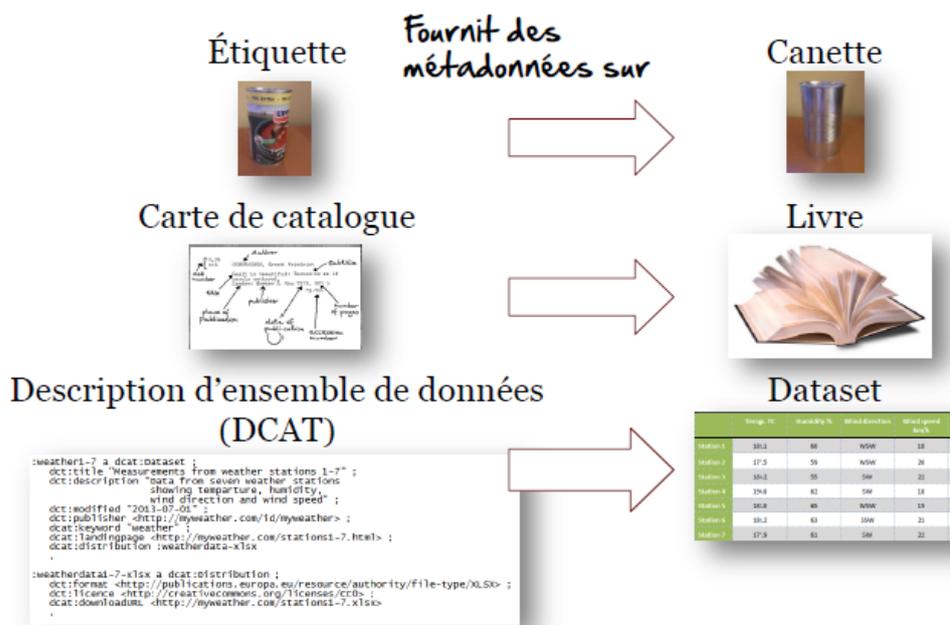
“Les métadonnées sont de l’information structurée qui décrit, explique, localise ou facilite autrement l’obtention, l’utilisation ou la gestion d’une ressource d’information. Les métadonnées sont souvent appelées données des données ou information sur l’information.” -- National Information Standards Organization

Les métadonnées fournissent des informations permettant de comprendre des **données** (documents, images, bases de données), des **concepts** (par exemple, les systèmes de classification) et **des entités du monde réel** (par exemple, les personnes, les organisations, les lieux, peintures, produits).

### 2.2. Types de métadonnées

- **Métadonnées descriptives**, décrivent une ressource à des fins de découverte et d’identification.
- **Métadonnées structurelles**, par ex. des modèles de données et des données de référence.
- **Métadonnées administratives**, renseignent afin d’aider à la gestion d’une ressource.

## Des exemples de métadonnées



## ***Deux approches pour fournir des métadonnées sur le Web***

*XML (Tree/container approach), RDF (approche basée sur le triplets)*

### ***2.3. La gestion des métadonnées***

*Les métadonnées doivent être gérées pour assurer leur...*

- **Disponibilité:** les métadonnées doivent être stockées où elles peuvent être consultées et indexées afin de pouvoir être trouvées.
- **Qualité:** les métadonnées doivent être de qualité constante afin que les utilisateurs sachent qu'ils peuvent y faire confiance.
- **Persistance:** les métadonnées doivent être entretenues au fil du temps.
- **Licence ouverte:** les métadonnées devraient être disponibles sous une licence du domaine public pour permettre leur réutilisation.

*Le cycle de vie de métadonnées est **plus grand** que le cycle de vie des données:*

- Les métadonnées **peuvent être créées avant que les données ne soient créées** ou capturées, par exemple, pour informer sur les données qui seront disponibles dans le future.
- Les métadonnées doivent être **conservées après que les données ont été supprimées**, par exemple, pour informer sur les données qui ont été déclassées ou retirées.

### ***2.4. Schéma de métadonnées***

*« Un système de labellisation, marquage et codage utilisé pour l'enregistrement d'information de catalogue ou pour structurer des notices descriptives. Un schéma de métadonnées établit et définit les éléments de données et les règles d'application pour l'usage de ces éléments de données pour décrire une ressource. »*

**2.5. Le cycle de vie des métadonnées :** *Création, entretien, mise à jour, stockage, publication des métadonnées et manipulation de la suppression des données.*

#### ***Créer vos métadonnées***

*La création de métadonnées peut être prise en charge par des processus (semi)-automatiques.*

- Les propriétés de documents générées par des outils (de bureau), par exemple la date de création d'un document.
- Informations spatiales et temporelles capturées par des caméras, des capteurs ...

- Informations issues du processus de publication, par exemple l'emplacement de fichier ou l'URL

*Cependant, d'autres caractéristiques requièrent une intervention humaine:*

- Quel est l'objet de la ressource (par exemple un lien vers le vocabulaire d'un sujet)?
- Comment peut-on utiliser les ressources (par exemple un lien vers une licence)?
- Où puis-je trouver plus d'information sur cette ressource (par exemple un lien vers un site Web ou de la documentation qui décrit la ressource)?
- Comment de l'information de qualité peut être incluse?

### ***Entretien vos métadonnées***

*Les approches pour l'entretien des métadonnées doivent être appropriées pour le type de données qui sont publiées.*

- Si **les données ne changent pas**, les **métadonnées** peuvent être relativement **stables**. Des changements (conversions en vrac) peuvent avoir lieu hors ligne lorsque cela est nécessaire.
- Si **les données changent fréquemment** (par exemple, les données de capteurs en temps réel), les **métadonnées** doivent être étroitement couplé au flux de données et les **changements** doivent être pratiquement **instantanés**.

### ***Mise à jour de vos métadonnées - planifier pour le changement***

*Les métadonnées opèrent dans un contexte mondial qui fait l'objet de changements!*

- **Organisation** - des départements sont établis, fusionne avec d'autres et des responsabilités sont transmises.
- **L'utilisation des données** - de nouvelles applications émergent autour des données.
- **Les données de référence** - les vocabulaires contrôlés évoluent et se relient.
- **Normes de données et de technologies** – le cycle de vie des technologie se raccourcit tout le temps; à quoi ressemblera le Web de demain?
- **Outils et systèmes** - Évolution du stockage, de la bande passante, de la mobilité ...

*Les métadonnées doivent être gardées à jour dans la mesure du possible, en tenant compte du temps et des budgets à disposition.*

### ***Stocker vos métadonnées - quelles sont les options?***

*En fonction des exigences opérationnelles, les métadonnées peuvent être intégrées avec les données ou stockées séparément des données.*

- Intégrer les métadonnées dans les données (documents bureautiques, MP3, JPG, données RDF) facilite l'échange de données.
- La séparation des métadonnées et des données (par exemple dans une base de données), avec des liens vers des fichiers de données correspondants rend la gestion plus facile.

En fonction de la disponibilité des outils et des exigences de performances et de capacité, les métadonnées peuvent être stockées dans une **base de données relationnelle «classique»** ou un **« store » de triplets en RDF**.

### ***La gestion de la suppression des données***

*Dans de nombreux cas, les métadonnées doivent survivre même après la suppression des données qu'elles décrivent.*

**Le déclasserment** ou **la suppression** des données peut se produire, par ex.:

- Lorsque les données ne sont plus nécessaires
- Lorsque les données ne sont plus valables.
- Lorsque les données sont mauvaises.
- Lorsque les données sont retirées par le propriétaire / éditeur

Dans ce cas, les métadonnées doivent **contenir l'information** que les données ont été **supprimées**, et si elles ont été **archivées**, comment et où une **copie d'archivage** peut être demandée.