

Qu'est-ce qu'une donnée de qualité ?

- Définition
- Qu'est-ce qu'une métadonnée ?
- Optimiser la pérennité des stockages

Définition

Une donnée de qualité est avant tout une donnée documentée (à l'aide d'un schéma de métadonnées) et pérennisée (grâce aux outils de stockage). Ce sont ces critères qui en feront une donnée pertinente et réutilisable dans le web sémantique.

Qu'est-ce qu'une métadonnée ?

Une métadonnée est une information rattachée à la donnée qui permet de la décrire (littéralement *une donnée sur la donnée*) : elle renseigne le contenu intellectuel, le contexte de production, les caractéristiques techniques des fichiers et des données ainsi que les propriétés et droits d'usage. On en trouve différents types : les métadonnées de description, qui donnent une idée précise du contenu de la ressource ; les métadonnées de gestion qui permettent d'accéder à la ressource (métadonnées d'identification, de provenance et de contexte) ; et enfin les métadonnées de préservation qui garantissent un accès pérenne au document et sa compréhension dans le long terme (métadonnées techniques, de structure, de droit).

Quelques liens :

- <https://www.opendatasoft.com/fr/blog/metadonnees-ce-qu'il-faut-savoir-avant-de-publier-vos-donnees/> Le blog de la société OpenDataSoft reprend de façon plus détaillée ces éléments, et aborde la question des schémas de métadonnées. Il s'agit d'aider les producteurs à assurer une description optimale de leurs données avant ouverture.
- <https://dorum.fr/metadonnees-standards-formats/> La fiche synthétique et les ressources de DoRANum sur la question.

Optimiser la pérennité des stockages

La question posée, en plus de celle des infrastructures, est celle du format des données qui doit être compatible avec les technologies en place. Un format de données est une convention utilisée pour représenter des données, des informations représentant un texte, une page, une image, un son, un fichier exécutable, etc. Il peut être ouvert (sa spécification est alors publiquement accessible) ou fermé (lorsque sa spécification est secrète). Éventuellement, un format peut aussi être normalisé (par une institution publique ou internationale, à l'exemple de ISO ou du W3C) lorsqu'il n'est pas propriétaire (soit s'il a été élaboré par une entreprise dans un but essentiellement commercial). Exemples de formats de métadonnées : CSV, XML, JSON, RDF...

- <https://dorum.fr/stockage-archivage/> La fiche synthétique et les ressources de DoRANum sur la question.