

ANNEXES RAPPELS et DEFINITIONS

Règlement du guichet de financement

« Référentiels de la culture et découvrabilité des contenus »

Edition 2023

Les métadonnées sont des données sur une donnée. Par exemple, un document multimédia comme une image ou une vidéo est considéré comme étant une donnée qui va elle-même comporter des métadonnées telles que l’auteur de l’image, le titre, la date de création de l’image etc.

Standards de métadonnées : il devient intéressant de se mettre d’accord de la façon la plus large et la plus admise possible sur la définition des métadonnées. Ainsi, on peut s’entendre sur le fait que le nom du créateur d’une image est une information importante. On peut s’entendre également sur le fait d’écrire toujours le prénom, le nom et la date de naissance par exemple. Cela revient donc à standardiser la bonne compréhension par tous de ces métadonnées et à les rendre ainsi les plus interopérables possibles. Dublin Core est par exemple un standard de métadonnées assez connu et repris internationalement.

Les référentiels : Il s’agit de données hautement partageables et stables. Lorsque ces métadonnées s’appuient sur des référentiels on peut dire que la qualité des données est au rendez-vous. En effet, si les standards de métadonnées permettent de s’entendre sur le type de métadonnées et la forme qu’elles prendront, l’utilisation de référentiels consiste à les renseigner « ensemble » ou au moins à les partager en s’appuyant donc sur les mêmes corpus que l’on pourrait par exemple co-produire et maintenir de façon collégiale ce qui en plus de gagner en efficacité, permet de gagner en interopérabilité sémantique et technique. Les standards de métadonnées correspondent donc aux règles qui concernent « la forme » alors que les référentiels servent « le contenu ».

Dans le document, la **sémantisation des données** est entendue comme étant l’opération technique permettant de formaliser et structurer les métadonnées sous forme de triplet RDF (avec un sujet – un objet – un prédicat) dans des formats technologiques compatibles avec les standards du W3C pour une diffusion et interopérabilité sur le Web des données liées. Des données sémantisées sont donc des données qui disposent de métadonnées RDF. On parle de sémantisation car c’est grâce à cette structuration de l’information que la machine distingue par exemple que la donnée « avocat » concerne le fruit et non le « métier d’avocat ».

Les alignements : on entend ici par alignement l’opération technique consistant à créer des correspondances entre des entités exprimées dans des triplets RDF. Ce liage de données permet de créer des passerelles entre différents référentiels au moyen de propriétés dont

la sémantique est « comprise » par la machine. Par exemple : le concept « Avocat » dans mon référentiel a un lien qui indique que ce concept est une sous-classe (relation de type terme spécifique) du concept « arbre fruitier » du référentiel de wikipédia. Les alignements permettent aussi de créer des passerelles linguistiques et ainsi de développer le multilinguisme. Par exemple, le concept « automobile » de mon référentiel est le même (relation équivalence exacte – « sameAs ») que le concept « car » en langue anglaise de wikipédia. Les métadonnées des deux concepts « automobile » sont liées techniquement par cette nouvelle métadonnée de liage (de type Equivalence exacte).

Ces liens sont typés conformément à la norme ISO 25964 (**ISO 25964** *Information and documentation - Thesauri and interoperability with other vocabularies*).

Le Graphe des connaissances culturelles ou **Graphe Culture** consiste à interconnecter au moyen des alignements les grands référentiels des domaines de la culture.

Ces alignements permettent *in fine* de générer de la connaissance. En effet, un moteur de recherche qui s'appuie sur des données RDF liées entre elles va non seulement vous permettre d'avoir des résultats de recherche pertinents car la question est « comprise » mais il va également vous pousser des résultats provenant de sources différentes dont vous n'aviez pas connaissance. C'est ainsi que le moteur de recherche Google fonctionne lorsque vous lui demandez des informations sur des « personnes connues ».

Une question ?

Pour toute question, vous êtes invités à vous adresser à :

Service du Numérique

Département du numérique pour la transformation des politiques culturelles et de l'administration des données

contact-depnum.snum@[culture.gouv.fr](mailto:contact-depnum.snum@culture.gouv.fr)