

WIKIDATA ET LES DONNÉES OUVERTES

Présentation pour la semaine de
la Science Ouverte



WIKIDATA : KÉSACO?

- base centralisée de données structurées (base de connaissances)
- libre et gratuite; licence [CC0](#), « aucun droit réservé »
- peut être lue et modifiée par des humains et des machines
- l'accès aux données se fait en utilisant des formats standards
- n'est pas un entrepôt de données de recherche, mais peut-être alimenté par des données provenant des projets de recherche

WIKIDATA : ÉLÉMENTS ET PROPRIÉTÉS

Chaque élément et chaque propriété est identifié par un URI (uniform resource identifier)

- **Éléments** (Q suivi d'un numéro) : plus de 100 millions
 - Ex: Q513 = « Everest » = « plus haute montagne de la Terre au-dessus du niveau de la mer »
 - <https://www.wikidata.org/wiki/Q513>
- **Propriétés** (P suivi d'un numéro) : 500 propriétés
 - <https://www.wikidata.org/wiki/Special:ListProperties>
 - Ex: P40 = « enfant » = « lien familial direct indiquant que l'objet du lien est un fils ou une fille du sujet »
 - <https://www.wikidata.org/wiki/Property:P40>

ANATOMIE D'UNE NOTICE STRUCTURÉE WIKIDATA

- Le triplet-

Sujet - Prédicat - Objet

Q - P - Q/str/date/coordonnées/nombre

Everest (Q513)

Sujet

nature de l'élément (P31)

Prédicat

montagne (Q8502)

Objet

Everest (Q513)

Sujet

partie de (P361)

Prédicat

Himalaya (Q5451)

Objet

Everest (Q513)

Sujet

coordonnées géographiques (P625)

Prédicat

27°59'17"N, 86°55'30"E

Objet

ANATOMIE D'UNE NOTICE : SITE DE WIKIDATA

archives municipales de Montpellier (Q64155968)

service d'archives communales de Montpellier
Commune de Montpellier - Archives de Montpellier

Plus de langues

Configurer

Langue	Libellé	Description
français	archives municipales de Montpellier	service d'archives communales de Montpellier
anglais	Archives municipales de Montpellier	Aucune description fournie
anglais américain	Pas de libellé défini	Aucune description fournie
espagnol	archives municipales de Montpellier	Aucune description fournie

Toutes les langues saisies

Déclarations

nature de l'élément

archives communales

1 référence

prononciation (fichier son)  0:02  modifier

LL-Q150 (fra)-T. Le Berre-archives municipales de Montpellier.wav
2,4 s ; 227 kio

langue de l'œuvre, du nom ou du terme français



1 référence

+ ajouter une valeur

pays  France  modifier




1 référence

+ ajouter une valeur

localisation administrative  Montpellier  modifier

1 référence

+ ajouter une valeur

coordonnées géographiques    modifier

43°36'24.52"N, 3°53'38.08"E

1 référence

+ ajouter une valeur

adresse  287 rue Poséidon Hôtel de ville, 1 place Georges Frêche, 34267 Montpellier Cedex 2 (français)  modifier

1 référence

+ ajouter une valeur

site officiel  <http://www.montpellier.fr/491-archives-municipales.htm>  modifier

1 référence

+ ajouter une valeur

WIKIDATA ET DONNÉES DE LA RECHERCHE

- **Référentiel** (de personnes, d'organisme, de lieux, de mots clés, etc)
 - Alignement de ses propres données avec celle de Wikidata
 - Enrichir ses propres données avec de nouvelles données issues de Wikidata
 - Enrichir Wikidata avec des données issues de ses propres projets de recherche (données ouvertes sous une licence CC0)

EXEMPLE DE MISE EN CORRESPONDANCE

Wikidata

Nom	Identifiant	Date de naissance	Lieu de naissance
Louis XVI	Q7732	23 août 1754	château de Versailles

mon projet de recherche

Nom	Identifiant	Lieu de mort
Louis Auguste de France	12170342†	Paris

ACCÈS AUX DONNÉES DE WIKIDATA: SPARQL

- Anatomie d'une requête

```
1 #Enfants de J-S Bach et de Maria Barbara Bach
2 SELECT ?enfant ?enfantLabel
3 WHERE
4 {
5   ?enfant wdt:P22 wd:Q1339. # ?enfant a pour père J-S Bach
6   ?enfant wdt:P25 wd:Q57487. #?enfant a pour mère Maria Barbara Bach
7
8   SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE]". }
9 }
```

Service de requête : <https://query.wikidata.org/>

- Partir d'une requête déjà existante à partir du dossier 'Exemples'
- #Commentaires à rajouter pour reconnaître le type de requête

EN PRATIQUE

- requête : **Toutes les rues de Rennes portant des noms de femmes.**

```
SELECT ?rue ?rueLabel ?personne ?personneLabel
WHERE
{
  # ?rue "a pour nature" "rue".
  # ?rue "a pour localisation administrative" "Rennes".
  # ?rue "est nommé en référence à" "?personne".
  # ?personne "a pour nature" "etre humain".
  # ?personne "a pour sexe" "féminin".

  SERVICE wikibase:label { bd:serviceParam wikibase:language "fr" }
}
```

Hint : partir toujours d'une notice existante (la plus complète) pour comprendre la structure : **rue Colette**

EN PRATIQUE

- requête : **Architectes français originaire d'Auvergne.**

```
SELECT ?element ?elementLabel ?villeLabel ?departementLabel ?coordonnees
```

```
WHERE
```

```
{
```

```
VALUES ?departement {wd:Q12694 wd:Q3113 wd:Q12572 wd:Q3259} # les lieux recherchés sont les 4 départements d'Auvergne
```

```
# ?element "a pour nature" "être humain".
```

```
# ?element "a pour occupation" "architecte".
```

```
# ?element "a fait l'école à" "Ecole Nationale de Beaux-Arts".
```

```
# ?element "est né dans" ?ville.
```

```
# ?ville "a pour localisation administrative" ?departement.
```

```
# ?ville "a pour coordonnées" ?coordonnees.
```

```
SERVICE wikibase:label { bd:serviceParam wikibase:language "fr" }
```

```
}
```

Hint : partir toujours d'une notice existante (la plus complète) pour comprendre la structure : **Gaston Bardet**

POUR EN SAVOIR PLUS

- Billets de blog qui décortiquent les requêtes : <https://bms.hypotheses.org/category/boite-a-outils/wikidata>
- Tutoriel sparql : https://www.wikidata.org/wiki/Wikidata:SPARQL_tutorial/fr