

Découverte de Wikidata, SPARQL et du Wikidata Query Service (WDQS)

Envel Le Hir – CC BY-SA 4.0

www.lehir.net

25 juillet 2017

France, Rennes, Les Champs Libres, 4C, Magenta

Programme

- Tour de table
- Présentation de Wikidata
- Présentation de SPARQL
- Utilisation du Wikidata Query Service

Wikidata

- Projet-frère de l'encyclopédie Wikipédia, lancé en octobre 2012
- Base de connaissance
 - base de données sur plus de 29 millions de concepts en juillet 2017
 - libre : licence Creative Commons « *Transfert dans le Domaine Public* » CC0 (\approx domaine public)
 - collaborative
 - multilingue

Représentation d'un concept

The diagram illustrates the structure of a Wikidata concept page for Douglas Adams (Q42). It is annotated with labels and lines pointing to specific elements:

- libellé**: Points to the main label "Douglas Adams (Q42)".
- description**: Points to the description "écrivain anglais de science-fiction" and the aliases "Douglas Noel Adams | Douglas Noël Adams".
- propriété**: Points to the "scolarité" property label.
- rang**: Points to the "2 références" section header.
- groupe de déclarations**: Points to the entire "Déclarations" section.
- identifiant de l'élément**: Points to the "(Q42)" part of the label.
- alias**: Points to the text "Douglas Noel Adams | Douglas Noël Adams".
- valeur**: Points to the "St John's College" label.
- qualificatifs**: Points to the table of qualifications for St John's College.
- références dépliées**: Points to the expanded reference for Encyclopædia Britannica Online.
- références repliées**: Points to the collapsed reference for Brentwood School.

Déclarations

scolarité

St John's College

date de fin	1974
principale spécialité d'étude	littérature anglaise
diplôme	baccalauréat en arts
date de début	1971

▼ 2 références

affirmé dans	Encyclopædia Britannica Online
URL de la référence	http://www.nndb.com/people/731/000023662/
langue originale	anglais
date de consultation	7 décembre 2013
éditeur	NNDB
titre	Douglas Adams (anglais)

+ ajouter une référence

Brentwood School

date de fin	1970
date de début	1959

► 0 références

+ ajouter

Ontologies

- Wikidata est organisé en projets
- Exemple : [projet France](#) et ses sous-projets
Communes, Monuments historiques, Élus
- Permet de trouver les propriétés à utiliser pour chaque type de concept

SPARQL

- « SPARQL Protocol and RDF Query Language »
 - standard du W3C : norme SPARQL 1.1 (2012)
- Langage de requête de base de données graphe RDF
- Triplet RDF : sujet → prédicat → objet
- Exemples :
 - Les Misérables* → date de publication → 1862
 - Les Misérables* → auteur → Victor Hugo
 - Victor Hugo → prénom → Victor
 - Victor Hugo → nom de famille → Hugo
 - Victor Hugo → date de naissance → 26 février 1802

Wikidata Query Service (WDQS)

- Adresse : <https://query.wikidata.org/>
- Vérifiez la langue de l'interface en haut à droite
- Permet d'interroger Wikidata grâce à SPARQL
 - afficher les données
 - les visualiser sous différentes formes
 - les télécharger
 - les partager

Première requête

- Liste des musées
- **SELECT** ?musee
WHERE {
 ?musee wdt:P31 wd:Q33506 .
}

Exhaustivité

- Liste (complète) des musées
- SELECT ?musee

WHERE {

 ?musee **wdt:P31/wdt:P279*** wd:Q33506 .

}

Libellés et descriptions

- Liste des musées avec leurs libellés
- Utiliser l'auto-complétion
- `SELECT ?musee ?museeLabel ?museeDescription
WHERE {
 ?musee wdt:P31/wdt:P279* wd:Q33506 .
 SERVICE wikibase:label { bd:serviceParam
wikibase:language "fr" . }
}`

Ajouter une condition simple

- Liste des musées **en France**
- ```
SELECT ?musee ?museeLabel ?museeDescription
WHERE {
 ?musee wdt:P31/wdt:P279* wd:Q33506 .
 ?musee wdt:P17 wd:Q142 .
 SERVICE wikibase:label { bd:serviceParam
 wikibase:language "fr" . }
}
```

# Récupérer les valeurs d'une propriété

- Localisation administrative des musées
- `SELECT ?musee ?museeLabel ?localisationLabel  
WHERE {  
 ?musee wdt:P31/wdt:P279* wd:Q33506 .  
 ?musee wdt:P17 wd:Q142 .  
 ?musee wdt:P131 ?localisation .  
 SERVICE wikibase:label { bd:serviceParam  
 wikibase:language "fr" . }  
}`

# Récupérer les valeurs d'une propriété, même lorsqu'elle n'est pas renseignée

- ```
SELECT ?musee ?museeLabel ?localisationLabel
WHERE {
  ?musee wdt:P31/wdt:P279* wd:Q33506 .
  ?musee wdt:P17 wd:Q142 .
  OPTIONAL { ?musee wdt:P131 ?localisation . }
  SERVICE wikibase:label { bd:serviceParam
    wikibase:language "fr" . }
}
```

Afficher les résultats avec des images

- Choix de la visualisation
- `SELECT ?musee ?museeLabel ?image`
`WHERE {`
 `?musee wdt:P31/wdt:P279* wd:Q33506 .`
 `?musee wdt:P17 wd:Q142 .`
 `OPTIONAL { ?musee wdt:P18 ?image . }`
 `SERVICE wikibase:label { bd:serviceParam`
 `wikibase:language "fr" . }`
`}`

Afficher les résultats sur une carte

- ```
SELECT ?musee ?museeLabel ?image ?coord
WHERE {
 ?musee wdt:P31/wdt:P279* wd:Q33506 .
 ?musee wdt:P17 wd:Q142 .
 OPTIONAL { ?musee wdt:P18 ?image . }
 ?musee wdt:P625 ?coord .
 SERVICE wikibase:label { bd:serviceParam
 wikibase:language "fr" . }
}
```

# Croiser les données

- Liste des musées avec leur région
- **SELECT ?musee ?museeLabel ?regionLabel**  
**WHERE {**  
    ?musee wdt:P31/wdt:P279\* wd:Q33506 .  
    **?musee wdt:P131+ ?region .**  
    **?region wdt:P31 wd:Q36784 .**  
    SERVICE wikibase:label { bd:serviceParam  
wikibase:language "fr" . }  
**}**



# Trier les résultats

- Trier les résultats par région
- ```
SELECT ?musee ?museeLabel ?regionLabel
WHERE {
  ?musee wdt:P31/wdt:P279* wd:Q33506 .
  ?musee wdt:P131+ ?region .
  ?region wdt:P31 wd:Q36784 .
  SERVICE wikibase:label { bd:serviceParam wikibase:language
"fr" . }
}
ORDER BY ?regionLabel
```

Limiter le nombre de résultats

- Limiter à 100 résultats
- ```
SELECT ?musee ?museeLabel ?regionLabel
WHERE {
 ?musee wdt:P31/wdt:P279* wd:Q33506 .
 ?musee wdt:P131+ ?region .
 ?region wdt:P31 wd:Q36784 .
 SERVICE wikibase:label { bd:serviceParam wikibase:language
"fr" . }
}
LIMIT 100
```

# Affecter une valeur à une variable

- Utile pour travailler sur un petit ensemble de données
- Ici, on ne récupère que le Musée de Bretagne
- `SELECT ?musee ?museeLabel ?regionLabel`

`WHERE {`

`BIND(wd:Q3329701 AS ?musee) .`

`?musee wdt:P31/wdt:P279* wd:Q33506 .`

`?musee wdt:P131+ ?region .`

`?region wdt:P31 wd:Q36784 .`

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

`}`

# Filtrer les résultats

- Musées ouverts à partir de 1987
- SELECT ?musee ?museeLabel ?date

WHERE {

?musee wdt:P31/wdt:P279\* wd:Q33506 .

?musee wdt:P17 wd:Q142 .

**?musee wdt:P571 ?date .**

**FILTER(?date >= "1987-01-01"^^xsd:dateTime) .**

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

}

ORDER BY DESC(?date)

# Enlever des résultats

- Liste des musées mais pas à Paris
- MINUS  $\approx$  FILTER NOT EXISTS
- SELECT ?musee ?museeLabel

WHERE {

  ?musee wdt:P31/wdt:P279\* wd:Q33506 .

  ?musee wdt:P17 wd:Q142 .

**FILTER NOT EXISTS { ?musee wdt:P131+ wd:Q90 . }**

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

}

# Compter le nombre de résultats

- Nombre de musées ayant leur région renseignée

- **SELECT (COUNT(\*) AS ?count)**

WHERE {

?musee wdt:P31/wdt:P279\* wd:Q33506 .

?musee wdt:P131+ ?region .

?region wdt:P31 wd:Q36784 .

}

# Aggréger

- Nombre d'oeuvres par musée
- ```
SELECT ?musee ?museeLabel (COUNT(*) AS ?count)
WHERE {
  ?musee wdt:P31/wdt:P279* wd:Q33506 .
  ?musee wdt:P17 wd:Q142 .
  ?oeuvre wdt:P276 ?musee .
  SERVICE wikibase:label { bd:serviceParam wikibase:language "fr" . }
}
GROUP BY ?musee ?museeLabel
ORDER BY DESC(?count)
```

Aggréger

- Nombre de musées par région
- **SELECT ?region ?regionLabel (COUNT(*) AS ?count)**
WHERE {
 ?musee wdt:P31/wdt:P279* wd:Q33506 .
 ?musee wdt:P131+ ?region .
 ?region wdt:P31 wd:Q36784 .
 SERVICE wikibase:label { bd:serviceParam wikibase:language "fr" . }
}
GROUP BY ?region ?regionLabel
ORDER BY DESC(?count)

Aggréger

- Superficie des musées par région
- **SELECT ?region ?regionLabel (MIN(?superficie) AS ?min) (AVG(?superficie) AS ?moy) (MAX(?superficie) AS ?max)**

WHERE {

?musee wdt:P31/wdt:P279* wd:Q33506 .

?musee wdt:P131+ ?region .

?region wdt:P31 wd:Q36784 .

?musee wdt:P2046 ?superficie .

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

}

GROUP BY ?region ?regionLabel

Questions ?