

\* 16/06/2023

# Wikidata pour la recherche

Présentation à l'Atelier rennais de  
la donnée



## \* *Résidence Wikimedia*

- Résidence Wikimedia de février 2023 à février 2024 à l'URFIST de Rennes, dans le cadre d'une convention entre l'association Wikimedia France et le MESR. Projet intitulé « Ouvrir la science à l'ouest ».
- **Objectif** : renforcer les liens entre les projets Wikimedia et le monde de la recherche, dans un contexte de promotion de la science ouverte.
- Deux autres résidences Wikimedia au sein des URFIST : l'une à Bordeaux, et l'autre à Toulouse.
- Page [Projet:Wikifier la science](#)

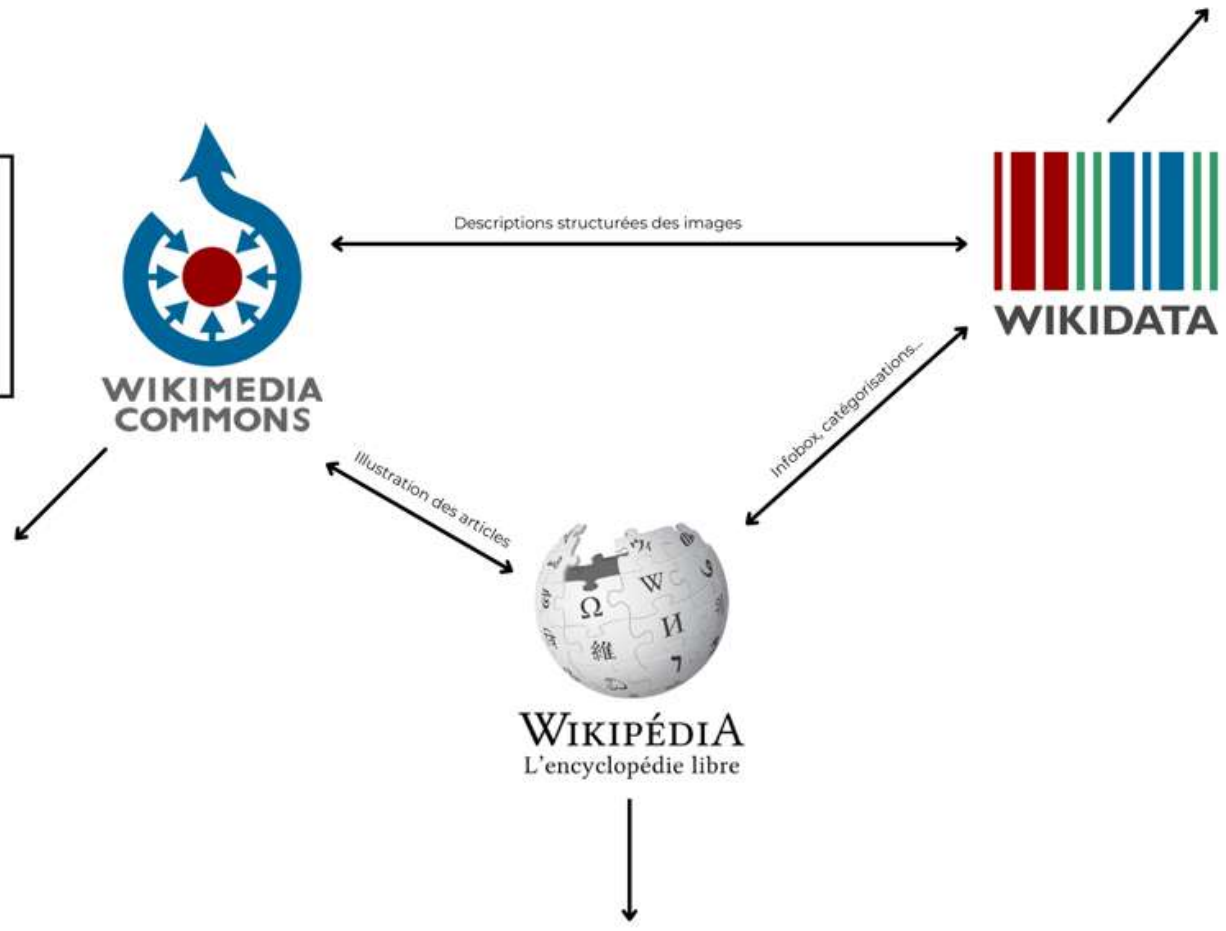
\* Plan

1. *Présentation de Wikidata*
2. *Exemples de données pour la recherche*
3. *Scientométrie et science ouverte*
4. *Requêtes et traitement des données*



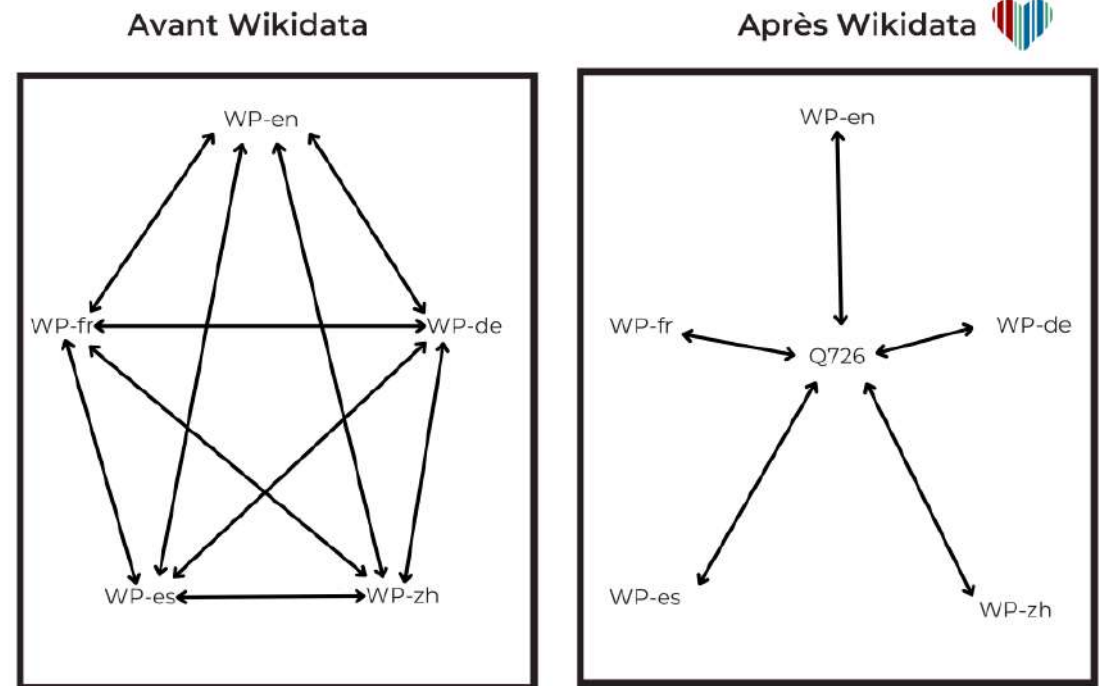


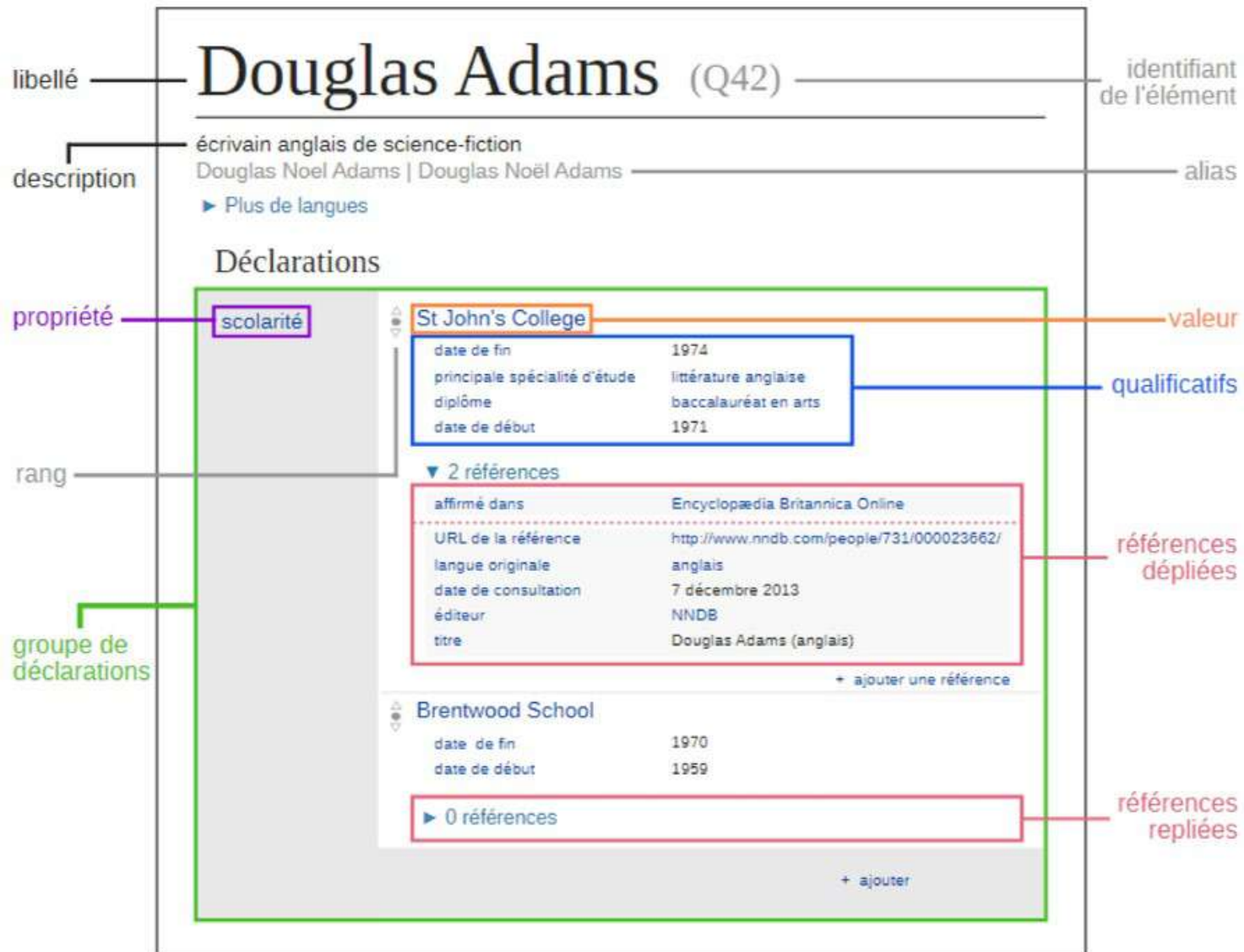
Liens entre  
les projets  
Wikimédia



# Wikidata, c'est quoi ?

- Wikidata est une base de données, reliées entre elles de manière **structurée**.
- Wikidata décrit des **éléments** dans toutes leurs dimensions, via un identifiant unique qui commence par la lettre **Q**.
- Les données sont sous **licence CC0**, donc réutilisables.





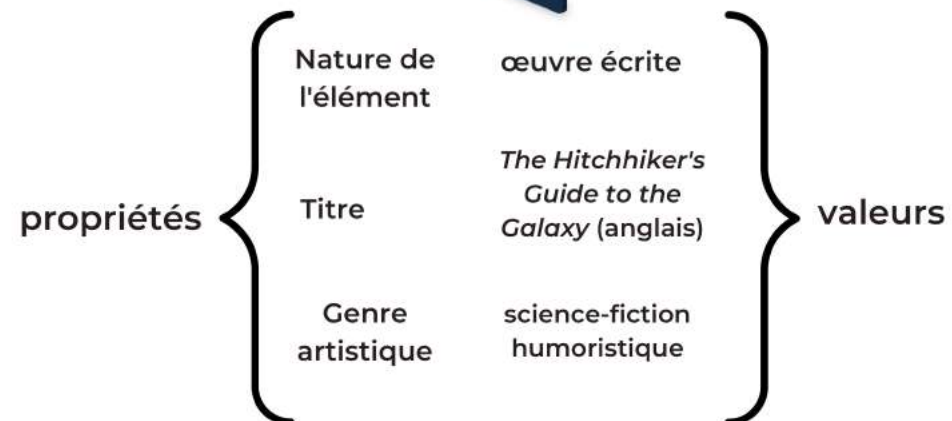


# Déclarations ?

sujet : Q3107329  
(Le guide du routard galactique)

Déclaration = triplet composé de :

- sujet
  - propriété
  - valeur
- 
- Q3107329 a pour nature d'être une œuvre écrite
  - Q3107329 a pour titre en anglais "The Hitchhiker's Guide to the Galaxy"
  - Q3107329 a pour genre artistique "science-fiction humoristique"



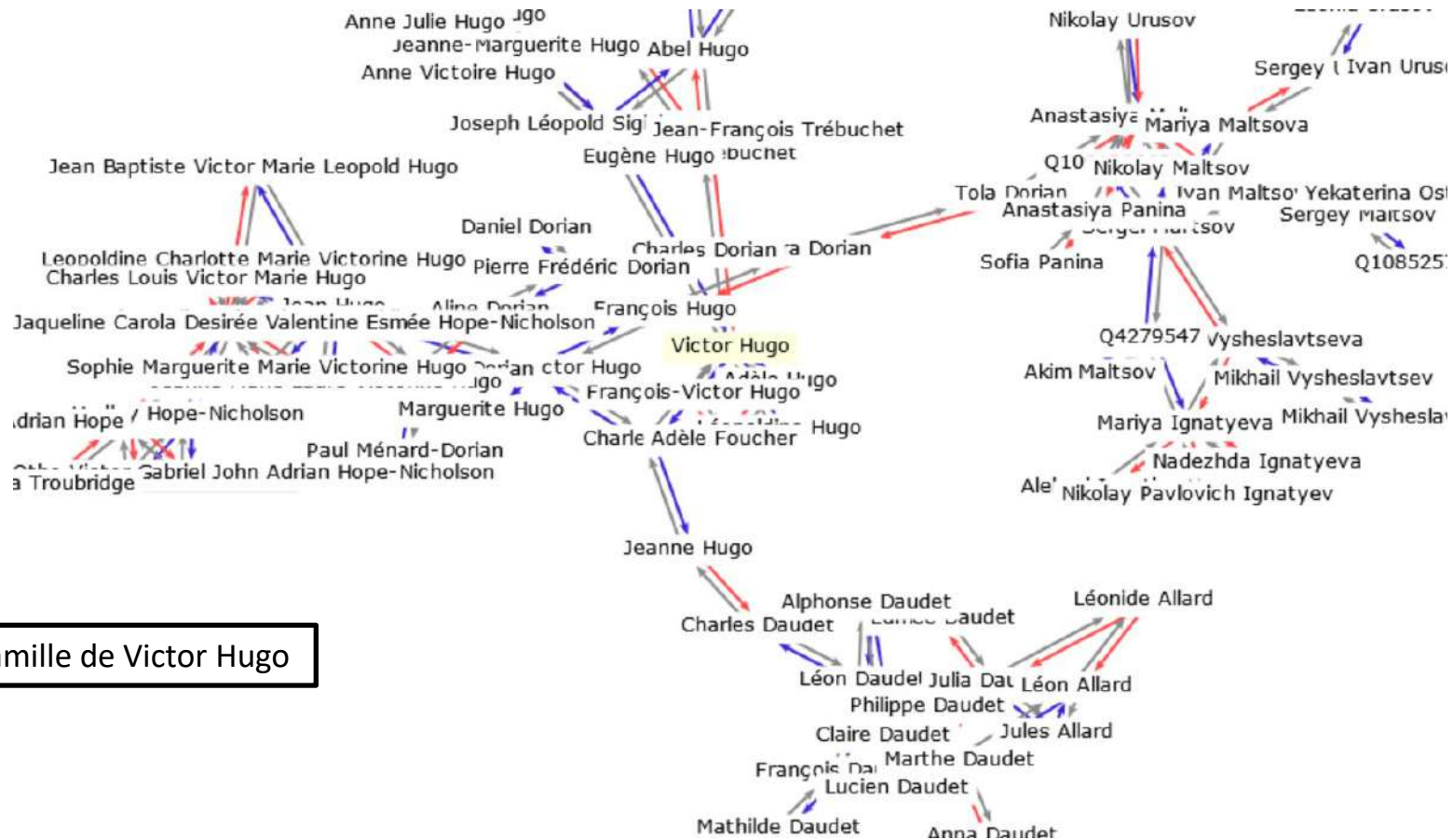




# Généalogie

## Outil GeneaWiki

<https://magnus-toolserver.toolforge.org/ts2/geneawiki/>



Famille de Victor Hugo



# Patrimoine



La porcelaine de Limoges

**Plateforme PciLab** pour l'inventaire du patrimoine culturel immatériel en France

<https://www.pci-lab.fr/>

## La porcelaine de Limoges

Domaine(s) : *Savoir-faire Arts du feu Arts graphiques Pratiques alimentaires*  
Sujet(s) : *Porcelaine Arts de la table Céramique Émail (verre) Entreprise du patrimoine vivant Maître d'art Ornemaniste Dessin technique Peinture Blanc Arts décoratifs Bijou Orfèvrerie*  
Langue(s) : *Français*  
Localisation : *Haute-Vienne Limoges Nouvelle-Aquitaine*  
Temporalité : *XVIIIe siècle XIXe siècle XXe siècle XXIe siècle*  
Médias : *Photographie*

Grâce au kaolin, Limoges devient, dès la fin du XVIIIe, un important centre de production de la porcelaine.

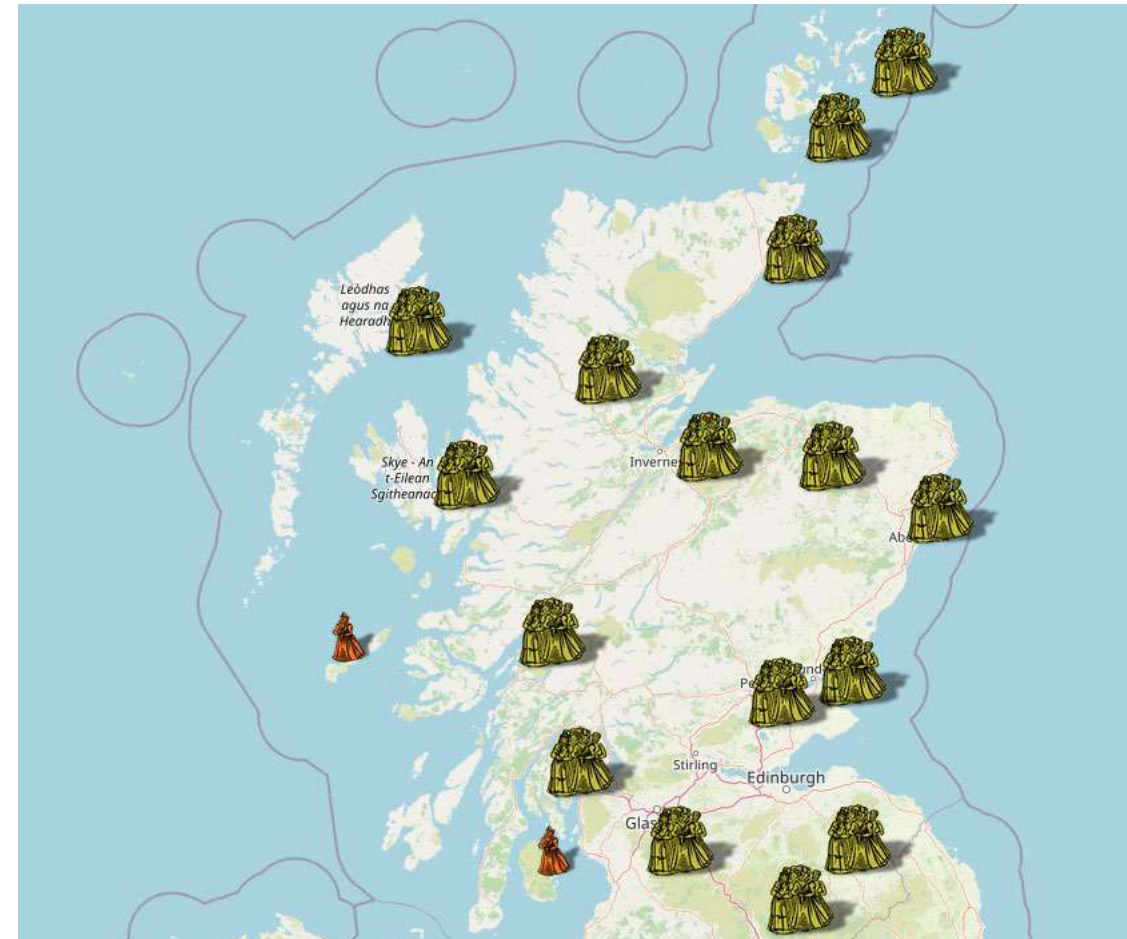
[Lire la suite](#) ▶



## Histoire

**Witches**, projet Wikidata de l'Université d'Edimbourg sur les sorcières accusées (XVIe-XVIIIe).

<https://witches.is.ed.ac.uk/>



# Art



Crotos

Cosmos  
Callisto  
Lab

Tout

Langue Français

impressionnisme

Rechercher

591 Résultats

-40000

1 2 3 4 5 ... 10 Au hasard



Charing Cross Bridge  
Claude Monet - musée d'art d'Indianapolis



La Femme à l'ombrelle — Madame  
Monet et son fils  
Claude Monet - National Gallery of  
Art



View of Mont Blanc, Seen from La Faucille  
Théodore Rousseau - Minneapolis Institute of Art



Arbres et route  
Paul Cézanne - Fondation Bames



Crotos, moteur de recherche et d'affichage d'œuvres d'art s'appuyant sur Wikidata et Wikimedia Commons

<https://zone47.com/crotos/>



# URFIST

rennes

**\* Scientométrie et science ouverte**



# Identifiants centralisés

Identifiants		
Identifiant VIAF	78252174 + 1 référence	modifier + ajouter une valeur
Identifiant ISNI	0000 0003 6554 4542 + 1 référence	modifier
	0000 0003 6554 4534 + 1 référence	modifier
	0000 0001 2140 1274 + 1 référence	modifier + ajouter une valeur
Identifiant VIAF	425 241945 + 1 référence	modifier + ajouter une valeur
Identifiant SAND d'auteur	0000117187 + 1 référence	modifier + ajouter une valeur
Identifiant ISNI d'une personne	90535051 + 0 référence	modifier + ajouter une référence + ajouter une valeur
Identifiant Bibliothèque nationale de France	001170472 + 1 référence	modifier + ajouter une valeur

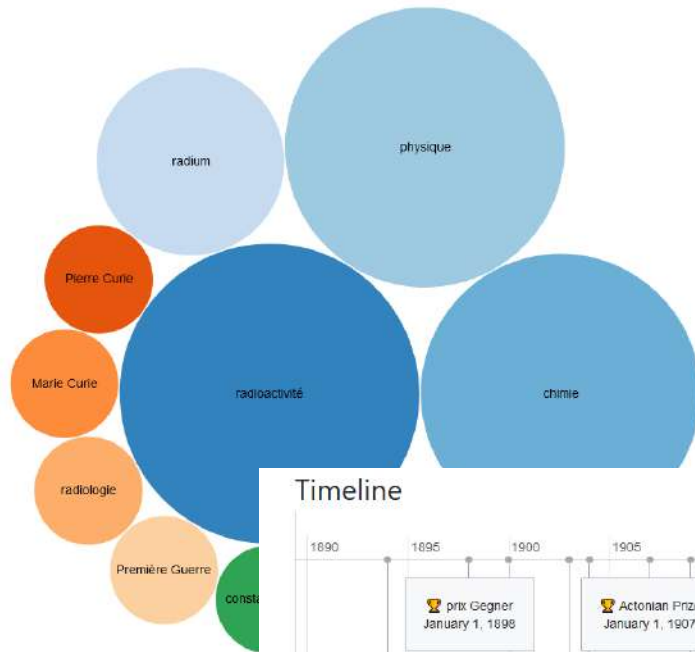
« devenue progressivement le point de convergence mondial des identifiants ouverts(...) Faire que tous les chercheurs français, vivants ou ayant vécu aient (...) **un Q Wikidata**, aligné avec leur(s) identifiant ORCID, IdHAL(...) »

(Comité pour la Science ouverte , 2019)

# Scholia

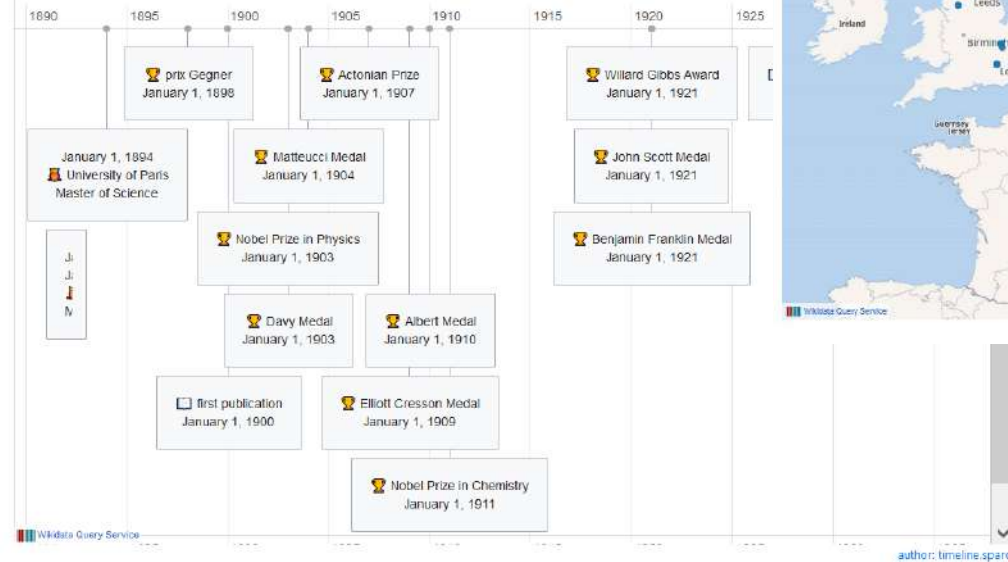
## Topic scores

Topics based on a weighting between fields of work, topics of authored works and topics of citing works.



Marie Curie sur Scholia

## Timeline



## List of publications

Reload

Search:

Date	Work	Type	Pages	Venue	Authors
1931-07-01	The Radioactive Constants as of 1930: Report of the International Radium-Standards Commission	scholarly article	19	Reviews of Modern Physics	André-Louis Debierne, Otto Hahn, Hans Geiger, Egon Schweidler, Marie Curie, Ernest Rutherford
1924-01-01	L'Isotopie et les élém				
1923-01-01	Pierre Curie				
1921-01-01	The Discovery of Rad				
1921-01-01	La Radiologie et la g				

## Co-author map

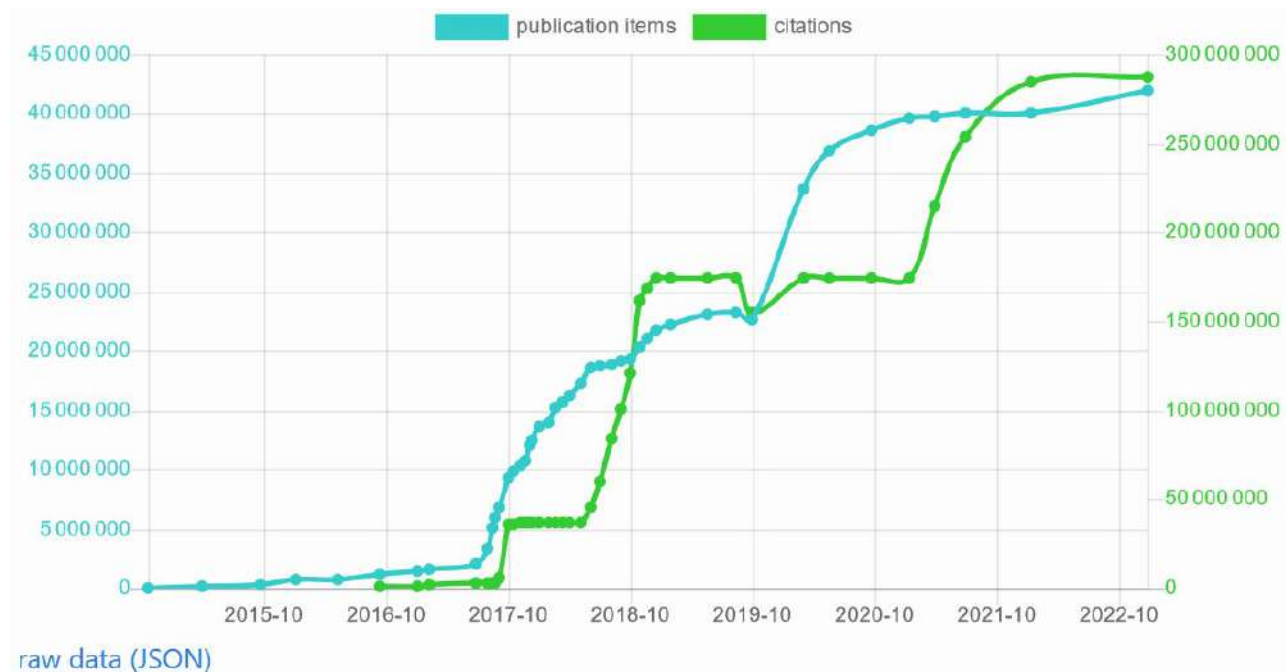


<https://scholia.toolforge.org/>

# WikiCite

Scholia s'inscrit dans le projet WikiCite :

- Initiative pour créer une base de données bibliographiques basée sur Wikidata
- Import massif d'articles scientifiques
- Objectif d'accroissement des citations dans les projets Wikimedia
- <http://wikicite.org/>



# *Guides pour l'ESR*

## **WikidataESR**

Guide réalisé par Julien Gossa visant l'harmonisation des informations sur l'organisation de l'enseignement supérieur français sur Wikidata.

<https://github.com/cpesr/wikidataESR>

## **Wikidata Thesis Toolkit**

Guide réalisé par Helen Williams (London School of Economics Library) et Ruth Elder (University of York Library) afin d'ajouter des métadonnées sur les thèses dans Wikidata.

[https://www.wikidata.org/wiki/Wikidata:WikiProject\\_Wikidata\\_Thesis\\_Toolkit](https://www.wikidata.org/wiki/Wikidata:WikiProject_Wikidata_Thesis_Toolkit)



# SPARQL

SPARQL est un **langage de requête** sémantique de base de données. Le point d'accès SPARQL de Wikidata permet d'interroger ses données.

➤ Interface de requête :

<https://query.wikidata.org/>

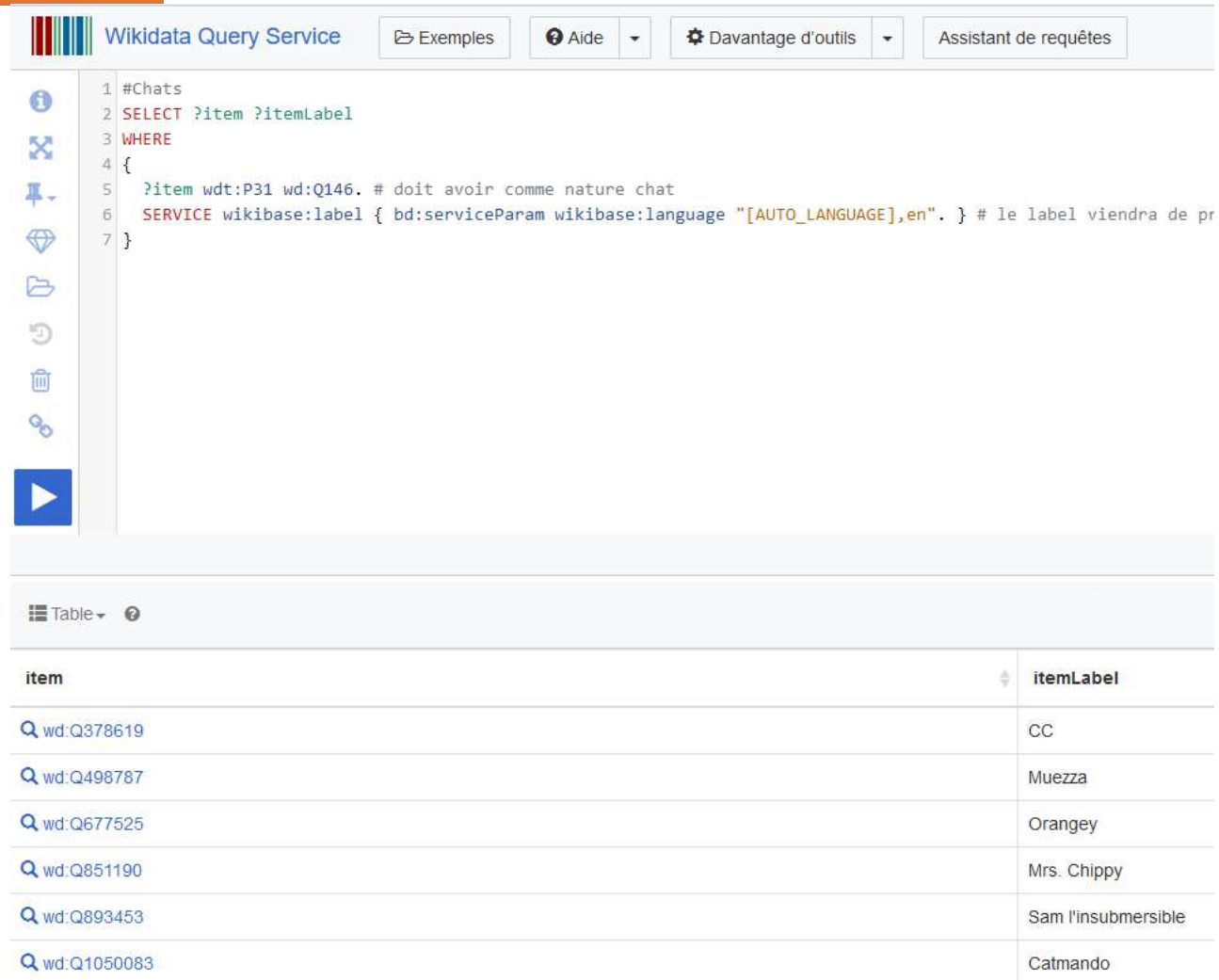
➤ Portail d'aide du service de requêtes :

[https://www.wikidata.org/wiki/Wikidata:SPARQL\\_query\\_service/Wikidata\\_Query\\_Help/fr](https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/Wikidata_Query_Help/fr)

➤ Exemples de requêtes :

[https://www.wikidata.org/wiki/Wikidata:SPARQL\\_query\\_service/queries/examples/fr](https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/queries/examples/fr)

## Requête « Chats »



The screenshot shows the Wikidata Query Service interface. At the top, there's a header with the Wikidata logo, the text "Wikidata Query Service", and buttons for "Exemples", "Aide", "Davantage d'outils", and "Assistant de requêtes". Below the header is a text area containing a SPARQL query:

```
1 #Chats
2 SELECT ?item ?itemLabel
3 WHERE
4 {
5   ?item wdt:P31 wd:Q146. # doit avoir comme nature chat
6   SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". } # le label viendra de pr
7 }
```

Below the query editor is a table with the following data:

item	itemLabel
<a href="#">wd:Q378619</a>	CC
<a href="#">wd:Q498787</a>	Muezza
<a href="#">wd:Q677525</a>	Orangey
<a href="#">wd:Q851190</a>	Mrs. Chippy
<a href="#">wd:Q893453</a>	Sam l'insubmersible
<a href="#">wd:Q1050083</a>	Catmando

# OpenRefine

- Logiciel libre de nettoyage et mise en forme de données.
- Se présentant comme un tableur, il permet entre autres **l'alignement de données avec Wikidata** et **l'import de données de Wikidata**.

OpenRefine test openrefine.xlsx Permalien

Facette / Filtre Défaire / Refaire 1 / 1

Rafraîchir Tout réinitialiser Tout supprimer

**Nom\_revue: judgment** changer  
2 choix Trier par: nom compte  
matched 8  
none 2  
Facette par nombre de choix

**Nom\_revue: best candidate's score** changer réinitialiser  
87 — 101

**10 lignes**  
Voir en: lignes entrées Afficher: 5 10 25 50 100 500 1000 lignes

	Toutes	Nom_revue	ISSN_elec
1.	☆	À tradire Choisir une nouvelle correspondance	2968-3912
2.	☆	Amplitude du droit Choisir une nouvelle correspondance	2826-1305
3.	☆	Motifs <input checked="" type="checkbox"/> Motifs An International Journal of English Studies (100) <input checked="" type="checkbox"/> Motifs (100) <input checked="" type="checkbox"/> Créer un nouvel élément Chercher une correspondance	2726-0399
4.	☆	Apparence(s) Choisir une nouvelle correspondance	1954-3778
5.	☆	Humanités numériques Choisir une nouvelle correspondance	2736-2337
6.	☆	La Bretagne linguistique Choisir une nouvelle correspondance	2727-9383
7.	☆	LISA (revue) Choisir une nouvelle correspondance	1762-6153
8.	☆	Revue internationale de Photolittérature (PHLiT) Choisir une nouvelle correspondance	2606-6785
9.	☆	Recherches en éducation Choisir une nouvelle correspondance	1954-3077
10.	☆	Revue Francophone Santé et Territoires <input checked="" type="checkbox"/> Revue francophone sur la santé et les territoires (RFST) (87) <input checked="" type="checkbox"/> Créer un nouvel élément Chercher une correspondance	2492-3672



# R/Python

## RStudio

- WikidataR : <https://cran.r-project.org/web/packages/WikidataR/index.html>
- Package SPARQL : <https://github.com/cran/SPARQL>

## Python

- Pywikibot : <https://github.com/wikimedia/pywikibot>

# *Modifications groupées*

## **QuickStatements**

Outil d'édition par lots de Wikidata

<https://quickstatements.toolforge.org/#/>

## **Mix'n'match**

Outil servant à l'importation des jeux de données dans Wikidata

<https://mix-n-match.toolforge.org/#/>



Merci pour votre attention.

**Contact :** [juliette.halimi@univ-rennes2.fr](mailto:juliette.halimi@univ-rennes2.fr)

**Pseudo Wikimedia :** Juliette en résidence

# Crédits

Documents utilisés :

- Formation Wikidata 3h par Amélie Charles, CC-BY-SA-4.0
- Formation Wikidata par Pierre-Yves Beaudouin, CC0

Présentation sous licence CC-BY-SA

