

Wikidata : découverte du corpus et cas d'usage dans le monde de la recherche

Semaine DATA-SHS – 12 décembre 2023

Pierre-Yves Beaudouin

Démo



Voir en temps réel la base de données se remplir :

- <http://listen.hatnote.com/#wikidata>
- <https://codepen.io/Krinkle/full/BwEKgW>
- <http://ahst.ro/lunchtime>
- <https://www.wikidata.org/wiki/Special:RecentChanges>

Plan



- Introduction à Wikidata
- Usages internes et externes
- Interface de Wikidata
- Qualité des données
- Wikidata et la Recherche

Introduction à Wikidata

Mouvement Wikimedia



Wikipédia est hébergée par la [Wikimedia Foundation](#), de même que les projets suivants :



Commons

Médiathèque



Wikidata

Base de données



Wiktionnaire

Dictionnaire



Wikisource

Bibliothèque



Wikinews

Site d'actualités



Wikiquote

Citations



Wikivoyage

Guide de voyage



Wikilivres

Manuels



Wikiversité

Cours et exercices



Wikispecies

Répertoire taxonomique

« Imaginez un monde dans lequel chaque être humain puisse partager librement la somme de toutes les connaissances. C'est notre engagement. »

Wikidata, c'est quoi ?

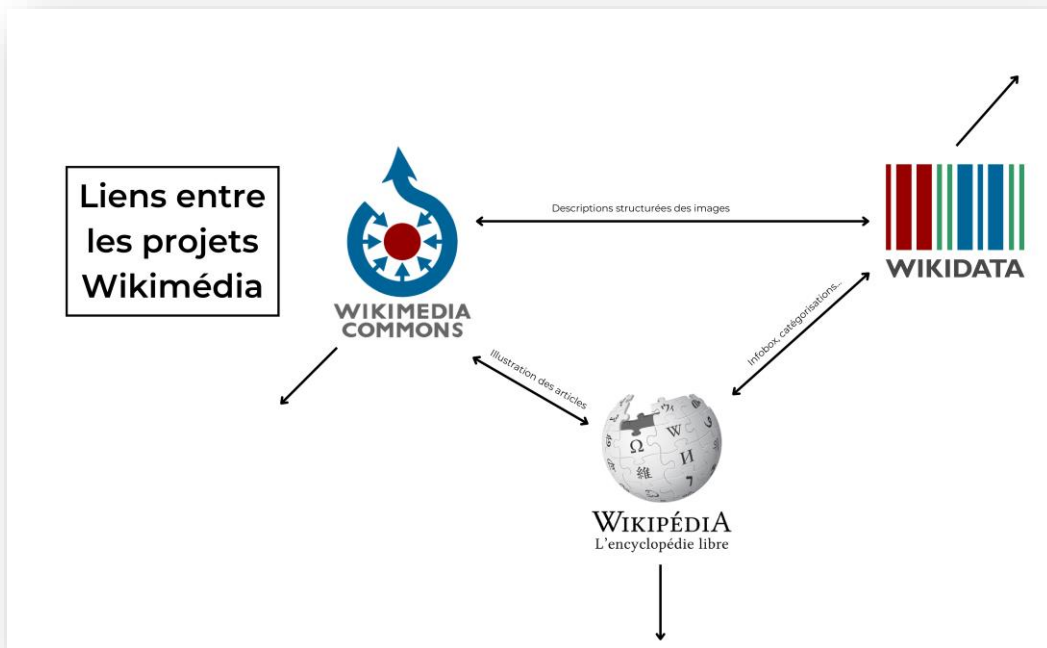


« Amorcée en 2006, Wikidata est un projet de la fondation Wikimedia ouvert à la contribution en 2012. Il vise à [constituer une base de connaissances](#) sur le monde qui puisse être [lue et éditée aussi bien par des humains que par des ordinateurs](#), les données étant organisées et catégorisées de manière standardisée pour pouvoir être manipulées dans divers contextes. Ainsi, Wikidata réunit sous [licence libre](#) Creative Commons Transfert dans le Domaine Public (CC0 1.0) un vaste ensemble de données pouvant être [copiées, modifiées, partagées, améliorées ou exploitées sans autorisation](#) » (Doutreix et Bouchereau, 2021)

CC0

- Instrument développé par **Creative Commons**
- **La mention CC0 n'est pas une licence**, mais une mention que la personne qui détient un droit de propriété littéraire et artistique, tel qu'un droit d'auteur, **renonce à ses droits patrimoniaux et moraux d'auteur**.
- Elle autorise toute personne à réutiliser librement ses travaux, les améliorer, les modifier, quel que soit le but et sans **aucune restriction** de droit, **sauf celles imposées par la loi**.

Vous utilisez déjà Wikidata sans le savoir



Objectifs

Wikidata répond à deux besoins :

1. **Interne au mouvement Wikimedia** : centraliser des données qui étaient avant Wikidata dupliquées sur chaque wiki (liens interlangues, résultats sportifs, données statistiques, etc.).
2. **Externe** : faciliter la réutilisation des contenus par d'autres acteurs que les géants du web.

Usages internes

Liens interlangues

The screenshot shows the French Wikipedia page for "Chat" (Felis catus). The page title is "Chat" and the scientific name is "*Felis catus*". The article is in French. A search bar at the top right contains the text "Rechercher sur Wikipédia" and a "Rechercher" button. On the right side of the article, there is a language selection menu with the text "260 langues" and a dropdown arrow. The menu is open, showing a search bar "Rechercher une langue" and a list of suggested languages under the heading "Langues suggérées". The suggested languages are: Alemannisch, Arpetan, Brezhoneg, Català, Corsu, Deutsch, English, Euskara, Interlingua, Italiano, Nederlands, Occitan, Picard, and Português. At the bottom of the menu, there is a button "+ Ajouter des langues" and a gear icon for settings.

WIKIPÉDIA
L'encyclopédie libre

Rechercher sur Wikipédia **Rechercher**

Chat

260 langues

Sommaire masquer

Début

Dénomination

- > Anatomie
- > Comportements
- > Santé
- > Obligations légales en Europe
- > Histoire
- > Culture populaire et arts
- > Aspects économiques
- > Notes et références
- > Annexes

Felis catus

Article [Discussion](#)

★ Vous lisez un « bon article »

🌐 Pour les articles homologues dans les autres langues

👉 « Matou » redirige ici

Le **Chat domestique** (***Felis catus***, **Félin**). Issu de la domestication du chat sauvage, il est considéré comme une **sous-espèce** de *Felis silvestris*.

Le Chat domestique est l'un des cinquante de **races** différentes. Dans certains pays, le chat entre dans le cadre de la législation du **furet**. Essentiellement territoriaux, les chats ont dit

Rechercher une langue

Langues suggérées

Alemannisch	Euskara
Arpetan	Interlingua
Brezhoneg	Italiano
★ Català	Nederlands
Corsu	Occitan
Deutsch	Picard
★ English	Português





















+ **Ajouter des langues**

Données statistiques

Clasificaciones finales [\[editar \]](#)

Las clasificaciones finalizaron de la siguiente forma:

Clasificación general (*Maillot Jaune*) [\[editar \]](#)


Clasificación general			
Ciclista	Equipo	Tiempo	
1.º  Jonas Vingegaard 🏆	Jumbo-Visma	82 h 05 min 42 s	
2.º  Tadej Pogačar 📄	UAE Team Emirates	+ 7 min 29 s	
3.º  Adam Yates	UAE Team Emirates	+ 10 min 56 s	
4.º  Simon Yates	Team Jayco AlUla	+ 12 min 23 s	
5.º  Carlos Rodríguez	Ineos Grenadiers	+ 13 min 17 s	
6.º  Pello Bilbao	Bahrain Victorious	+ 13 min 27 s	
7.º  Jai Hindley	Bora-Hansgrohe	+ 14 min 44 s	
8.º  Felix Gall	AG2R Citroën Team	+ 16 min 09 s	
9.º  David Gaudu	Groupama-FDJ	+ 23 min 08 s	
10.º  Guillaume Martin	Cofidis	+ 26 min 30 s	
11.º  Thibaut Pinot	Groupama-FDJ	+ 28 min 03 s	
12.º  Sepp Kuss	Jumbo-Visma	+ 37 min 32 s	
13.º  Thomas Pidcock	Ineos Grenadiers	+ 47 min 52 s	
14.º  Rafał Majka	UAE Team Emirates	+ 56 min 36 s	
15.º  Jonathan Castroviejo	Ineos Grenadiers	+ 56 min 37 s	
16.º  Chris Harper	Team Jayco AlUla	+ 57 min 29 s	
17.º  Ben O'Connor	AG2R Citroën Team	+ 1 h 04 min 59 s	
18.º  Wilco Kelderman	Jumbo-Visma	+ 1 h 06 min 46 s	
19.º  Mikel Landa	Bahrain Victorious	+ 1 h 12 min 41 s	
20.º  Valentin Madouas	Groupama-FDJ	+ 1 h 14 min 10 s	

Fuente: [ProCyclingStats](#)

Tour de France 2023

Données diverses

Présidents

Portrait	Identité 	Période		Durée 
		Début 	Fin 	
	Émile Rivière ⁶ (1835 - 1922)	1904	1905	1 an
	Lionel Bonnemère ⁶ (1843 - 1905)	1905	1905	moins d'un an
	Adrien de Mortillet ⁶ (1853 - 1931)	1906	1907	1 an
	P. H. M. Ballet (d) ⁶ (XIX ^e siècle - 1920)	1907	1908	1 an
	Théodore Baudon ⁶ (1848 - 1913)	1908	1909	1 an

**Présidence de la
Société
préhistorique
française**

Infoboîte Wikipédia

Falmarès

Une illustration sous licence libre serait la bienvenue

Biographie

Naissance 19 décembre 2001 (21 ans)
Conakry

Pseudonyme Falmarès

Nationalité guinéenne

Domiciles Guinée (jusqu'en 2016), France (depuis 2017)

Activité Poète

Autres informations

Site web falmarès.com

modifier - modifier le code - modifier Wikidata

Conseil national du numérique



Histoire

Fondation 29 avril 2011

Cadre

Sigle CNum

Type Commission consultative indépendante

Objet social Étudier les questions relatives au numérique

Domaine d'activité Numérique

Siège Halle Freyssinet (75013, France)

Pays France

Coordonnées 48° 49′ 58″ N, 2° 22′ 16″ E

Organisation

Membres 21 (depuis février 2021), 30 (2020)

Président Gilles Babinet, Françoise Mercadal-Delasalles

Secrétaire général Jean Cattan (d) (depuis 2021)

Budget 86 000 € (2019)

Site web www.cnummerique.fr

Jatiya Sangsad Bhaban

জাতীয় সংসদ ভবন



Présentation

Type Bâtiment parlementaire

Fondation 1982

Styles Mouvement moderne, brutalisme

Architecte Louis Kahn

Ouverture 28 janvier 1982

Hauteur 155 m

Surface 800 000 m²

Site web (bn) www.parliament.gov.bd

Localisation

Localisation Dacca

Coordonnées 23° 45′ 44″ N, 90° 22′ 43″ E



modifier - modifier le code - modifier Wikidata

Liens externes



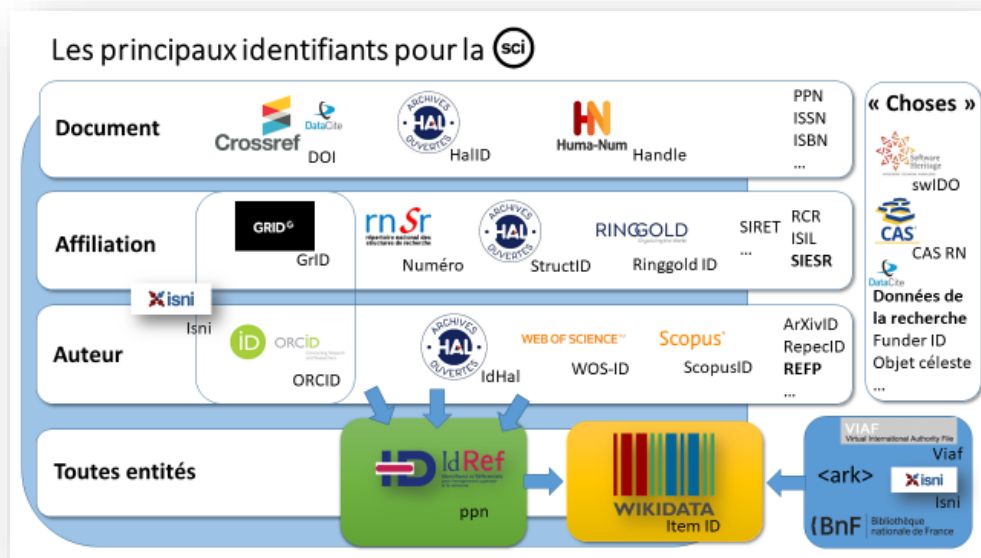
Liens externes [modifier | modifier le code]

- [Site officiel](#) [archive]
- Ressources relatives à la recherche ✎ : [Cairn](#) [archive] · [Collège de France](#) [archive] · [Diffusion des savoirs de l'École normale supérieure](#) [archive] · [Digital Bibliography & Library Project](#) [archive] · [Google Scholar](#) [archive] · [Mathematics Genealogy Project](#) [archive] · [ORCID](#) [archive] · [Research Papers in Economics](#) [archive] · [Scopus](#) [archive] · [Semantic Scholar](#) [archive]
- Ressource relative au spectacle ✎ : [Les Archives du spectacle](#) [archive]
- Ressource relative à plusieurs domaines ✎ : [Radio France](#) [archive]
- Ressource relative à l'audiovisuel ✎ : [IMDb](#) [archive]
- Notices dans des dictionnaires ou encyclopédies généralistes ✎ : [Dictionnaire universel des créatrices](#) [archive] · [Nationalencyklopedin](#) [archive] · [Munzinger](#) [archive] · [Trecani](#) [archive] · [Universalis](#) [archive] · [Visuotinė lietuvių enciklopedija](#) [archive]
- Notices d'autorité ✎ : [VIAF](#) [archive] · [ISNI](#) [archive] · [BnF](#) [archive] (données) · [IdRef](#) [archive] · [LCCN](#) [archive] · [GND](#) [archive] · [Italie](#) [archive] · [Japon](#) [archive] · [CiNii](#) [archive] · [Belgique](#) [archive] · [Pays-Bas](#) [archive] · [Pologne](#) [archive] · [Israël](#) [archive] · [NUKAT](#) [archive] · [Catalogne](#) [archive] · [Norvège](#) [archive] · [Tchéquie](#) [archive] · [Portugal](#) [archive] · [WorldCat](#) [archive]

Identifiants ouverts

« Wikidata [...] est devenue progressivement le point de convergence mondial des identifiants ouverts »

— Comité pour la science ouverte, *Des identifiants ouverts pour la science ouverte : note d'orientation*, 2019



Usages externes

Sites internet



Moteurs de recherche

À propos

Esther Duflo, née le 25 octobre 1972 à Paris, est une économiste franco-américaine. Professeure au Massachusetts Institute of Technology et au Collège de France, elle reçoit le Prix de la Banque de Suède en 2019 conjointement avec son époux Abhijit Banerjee et Michael Kremer. [Wikipédia](#)

Date/Lieu de naissance : 25 octobre 1972 (Âge: 51 ans), Paris

Enseignement : [Massachusetts Institute of Technology](#), PLUS

Époux : [Abhijit Banerjee](#) (m. 2015)

Distinctions : [Prix de la Banque de Suède en sciences économiques en mémoire d'Alfred Nobel](#), PLUS

Parents : [Violaine Duflo](#)

Conseillers pédagogiques : [Abhijit Banerjee](#), [Esther Duflo](#), [Michael Kremer](#), [Angrist](#)

Google

Esther Duflo

Économiste franco-américaine

Esther Duflo, née le 25 octobre 1972 à Paris, est une économiste franco-américaine. [Wikipedia \(FR\)](#)



Wikipedia



IMDb

DuckDuckGo

Esther Duflo

Économiste franco-américaine

Esther Duflo, née le 25 octobre 1972 à Paris, est une économiste franco-américaine. Professeure au Massachusetts Institute of Technology (MIT) et au Collège de France, ell...

[Lire plus sur Wikipedia](#)

Date et lieu de naissance : 25 oct. 1972 (51 ans) à Paris, France

Professions : économiste, professeur d'université

Époux : [Abhijit Banerjee](#) (m. 2015)

Parents : [Michel Duflo](#)

Site Web : econ-www.mit.edu/faculty/eduflo



Wikipedia.org • Comment contribuer ?

Qwant

Moteurs de recherche

Esther Duflo — Wikipédia

https://fr.wikipedia.org/wiki/Esther_Duflo

Esther Duflo, née le 25 octobre 1972 à Paris, est une économiste franco-américaine. Professeure au Massachusetts Institute of Technology (MIT) et au Collège de France, elle reçoit le Prix de la Banque de Suède en 2019 conjointement avec son époux Abhijit Banerjee et Michael Kremer.... [Afficher plus](#)

Biographie

Enfance et études

Esther Duflo naît le 25 octobre 1972 à [Paris](#) dans une famille [protestante](#) . Son père est le [mathématicien](#) ... [Afficher plus](#)

Travaux

Son domaine de recherche porte sur les questions [microéconomiques](#) dans les pays en développement, en particulier la santé, l'éducation, l'accès au crédit, en... [Afficher plus](#)



Esther Duflo, née le 25 octobre 1972 à Paris, est une économiste franco-américaine. Professeure au Massachusetts Institute of Technology et au Collège de France, elle reçoit le Prix de la Ban... [+](#)
[Wikipédia](#)

economics.mit.edu

Wikipédia IMDb

Bing

The screenshot shows the Mir@bel website interface. At the top, there is a navigation bar with the Mir@bel logo, a menu with 'Le réseau', 'Soutien', and 'Listes', a search bar with 'Recherche' and a magnifying glass icon, and a 'Recherche avancée' link with a question mark. A user profile icon is in the top right corner.

Below the navigation bar, a breadcrumb trail reads 'Accueil / Revues / Revue d'économie politique'. The main title 'Revue d'économie politique' is displayed in large blue font, with a list icon and arrows to its right.

Under the title, there is a 'Thématique' section with two tabs: 'Économie, gestion' (selected) and 'Science politique, droit et vie publique'.

A table with the following columns is shown:

Titre	+	ISSN	ISSN-E	Années	Éditeurs	Action
Revue d'économie politique		0373-2630	2105-2883	1887 – ...	Daloz, Éditions Sirey, Imprimerie L. Larose	

Below the table, a thumbnail image of the journal cover is shown on the left. To its right, the following details are listed:

- Titre actuel** Revue d'économie politique
- Site web** http://www.daloz-revues.fr/revues/Revue_d_economie_politique-119.htm
- Périodicité** bimestriel
- Langues** français
- Éditeurs** Daloz (1845 à ...)
Éditions Sirey (1789 à 1965)
Imprimerie L. Larose

At the bottom, there is an 'Autres liens' section with icons and links to 'Cairn international', 'EZB', 'HAL', 'MIAR', 'Scopus', 'Wikipedia (français)', and 'Worldcat'.

ScanR

The screenshot displays the ScanR website interface. At the top left is the ScanR logo with the text "Ministère de l'Enseignement Supérieur et de la Recherche". To its right are navigation links: "A propos", "Recherche", and "Focus". On the top right, there is a "Partager" button and language options "FR" and "EN". The main content area features a grid of logos for various reference systems:

- RÉFÉRENTIELS**: A blue square button.
- doi**: The Digital Object Identifier logo.
- GRID**: The Global Research Identifier Database logo.
- Insee**: The French National Institute of Statistics and Economic Studies logo, with the tagline "Mesurer pour comprendre".
- rnSr**: The Répertoire National des Structures de Recherche logo.
- IdRef**: The Identifiants et Référentiels pour l'enseignement supérieur et la recherche logo.
- ORCID**: The Open Researcher and Contributor Identifier logo.
- WIKIDATA**: The Wikimedia project logo.
- Cooperative Patent Classification**: The logo for the European Patent Office and United States Patent and Trademark Office.

Navigation arrows are visible at the bottom left and right of the grid.

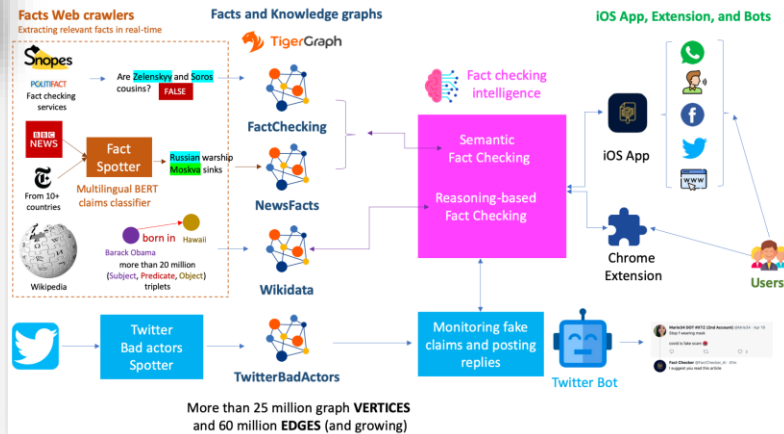
ScanR (suite)

The screenshot displays the ScanR website interface. At the top left is the ScanR logo with the text "scanR" and "Centre de Recherche en Économie et Statistique". To its right are navigation links: "A propos", "Recherche", and "Focus". On the top right, there is a "Partager" button and language options "FR | EN". A dark blue header bar contains the text "Ecole d'économie de Paris" and a "Naviguer par section" dropdown menu. The main content area is split into two columns. The left column shows a Wikipedia article snippet for the Paris School of Economics, featuring its logo (a pink circle with "pse" in white) and the text: "PARIS SCHOOL OF ECONOMICS ECOLE D'ÉCONOMIE DE PARIS". Below the logo, the text reads: "PSE-École d'économie de Paris, aussi appelée PSE, acronyme de sa traduction anglaise Paris School of Economics, est un réseau thématique de recherche avancée créé en 2006 sous la forme juridique d'une fondation de coopération scientifique. Elle a pour vocation de former des économistes, et de contribuer à la recherche en économie en France." A "Voir l'article" link is at the bottom right of this section. The right column shows a video player with a blue background, the hashtag "#ECOPOURTOUS", and the text "CONTENU EN VUE-T-IL LES ÉCONOMIQUES ?". A red play button is centered over the video. The video player header includes the name "Étienne Madinier | Co..." and a menu icon.

Assistants vocaux, IA



© MarkChristopherWebster, CC-BY-SA-4.0



Production de statistiques



Biais de genre



Who's Who in France



Encyclopædia Britannica



Oxford Dictionary of National Biography



Encyclopædia Universalis



Divers

- [Invités de la matinale de France Inter](#) (méthodologie)
- [Invités des émissions de Frédéric Taddeï](#) (méthodologie)

Grace a [SPARQL \(Q54871\)](#), voici une liste de quelques possibilités offertes lorsque tous ces épisodes sont insérés dans Wikidata:

- Parité homme/femme des émissions TV de Frédéric Taddéi: [requête](#) [résultat](#)
- Age des invités des émissions de Frédéric Taddéi (ici, la date de l'émission est fondamentale, car on prend l'age relatif à l'émission): [requête](#) [resultat](#) [age moyen par émission](#)
- Profession des invités d'Interdit d'interdire (requête non optimisée): [requête](#) [résultat](#) [bubble chart](#)
- Liste des personnes les plus invitées par Frédéric Taddéi toutes émissions confondues (et le winner is: [Emmanuel Todd](#)⁽⁷¹³⁷⁷⁵⁾): [requête](#) [résultat](#)

Présentation de l'interface

En-tête

Esther Duflo (Q434509)

Élément [Discussion](#)










économiste franco-américaine

 [modifier](#)

▾ [Plus de langues](#)

Langue	Libellé	Description	Également connu comme
français	Esther Duflo	économiste franco-américaine	
allemand	Esther Duflo	französische Ökonomin	Esther Caroline Duflo Esther C. Duflo
anglais	Esther Duflo	French-American economist	Esther Caroline Duflo Esther C. Duflo E. C. Duflo Duflo E. C.
espagnol	Esther Duflo	economista francesa	

Déclarations

sexe ou genre	 féminin Q6581072	 modifier
	▶ 3 références	
		+ ajouter une valeur
pays de citoyenneté	 France Q142	 modifier
	 date de début 1972	
	rang dans la série 1	
	▼ 0 référence	+ ajouter une référence
	 États-Unis Q30	 modifier
	date de début 2012	
	rang dans la série 2	
	▶ 1 référence	+ ajouter une valeur
nom dans la langue maternelle de la personne	 Esther Duflo (français)	 modifier
	▼ 0 référence	+ ajouter une référence
		+ ajouter une valeur

Identifiants

Identifiants

identifiant VIAF

74926739

 modifier

▶ 3 références

+ ajouter une valeur

identifiant ISNI

0000 0000 7851 4053

 modifier

▶ 1 référence

+ ajouter une valeur

identifiant BIBSYS d'une
personne

2046010

 modifier

▶ 1 référence

+ ajouter une valeur

identifiant CANTIC

981058529119406706



 modifier


▼ 0 référence


+ ajouter une référence


+ ajouter une valeur


Liens interwikis


Wikipédia (60 entrées)  


Wikilivres (0 entrée) 


Wikinews (1 entrée) 
 [Категория:Эстер Дюфло](#)


Wikiquote (1 entrée) 
 [Esther Duflo](#)

Wikisource (0 entrée) 

Wikiversité (0 entrée) 

Wikivoyage (0 entrée) 

Wiktionnaire (0 entrée) 

Autres sites (1 entrée) 
 [Category:Esther Duflo](#)

Wikipédia (60 entrées)  

[Дюфло, Эстер](#)

[إستير دوفلو](#)

[استير دوفلو](#)

[ইস্টার ডাফ্লো](#)

[Ester Dúflo](#)

[Эстер Дюфло](#)

[Эстэр Дзюфло](#)

[Естер Дюфло](#)

[এস্টের দুফ্লো](#)

[Esther Duflo](#)

[ئسنەر دوفلو](#)

[Esther Duflová](#)

[Esther Duflo](#)

[Esther Duflo](#)

[Esther Duflo](#)

[Esther Duflo](#)

[Esther Duflo](#)

[Esther Duflo](#)

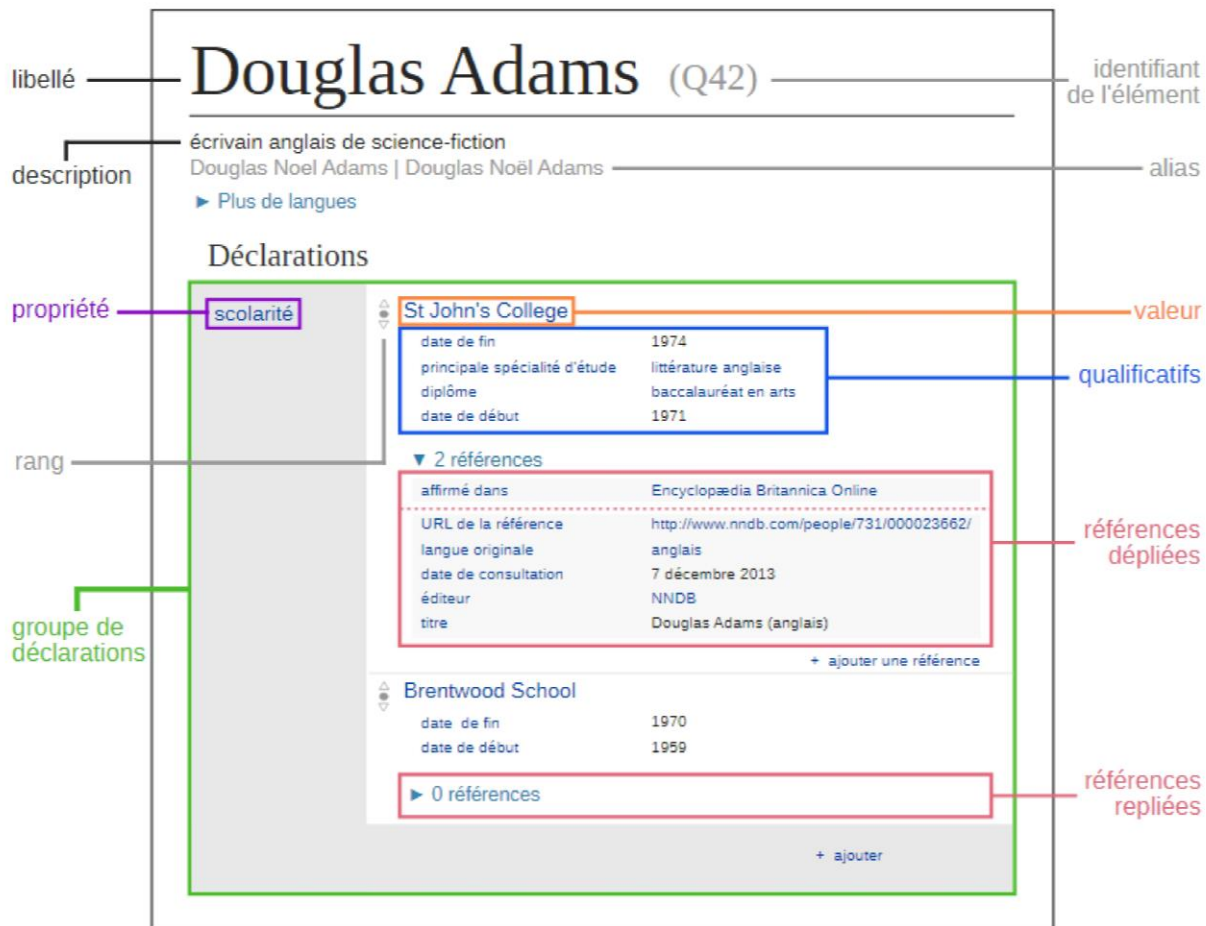
[Esther Duflo](#)

[Esther Duflo](#)

[استر دوفلو](#)

[Esther Duflo](#)

[Esther Duflo](#)

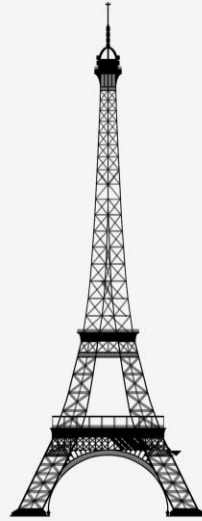


SUJET : TOUR EIFFEL (Q243)

Propriété
→ matériau (P186)

Valeur
acier
(Q11427)

Valeur
fer forgé
(Q860451)



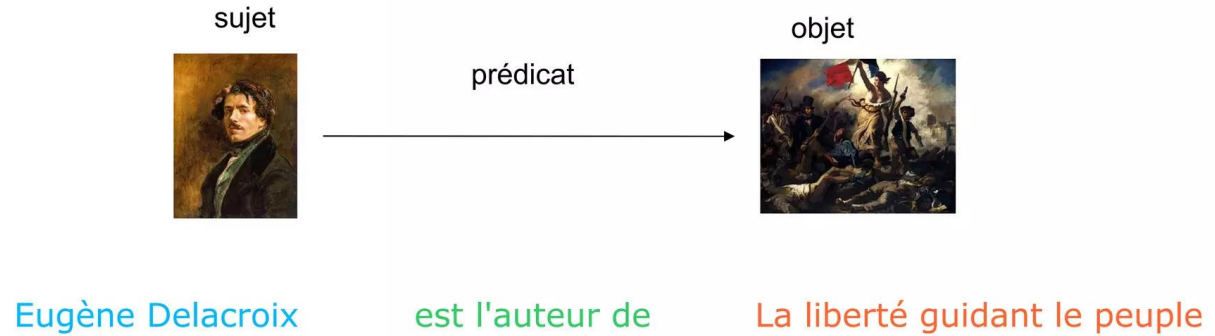
Propriété
→ hauteur (P2048)

Valeur
324 mètres
(au sommet)

Valeur
300 mètres
(architecture)

[MOOC]
Wikidata

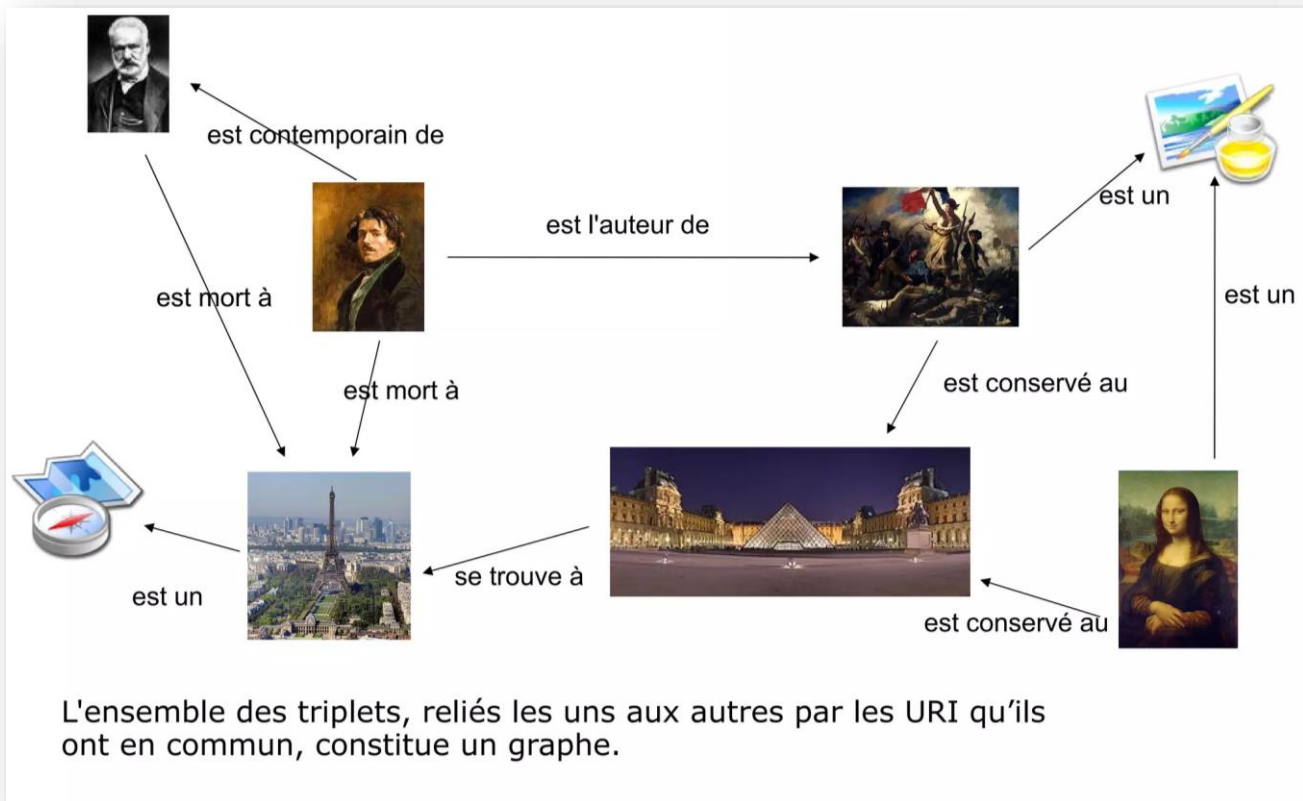
RDF



Le principe de RDF est d'exprimer l'information sous forme de phrases simples (sujet verbe complément, ou sujet prédicat objet).

Ces phrases sont appelées "triplets".

Graphe de connaissance



L'ensemble des triplets, reliés les uns aux autres par les URI qu'ils ont en commun, constitue un graphe.

Qualité des données



Création de nouvelles propriétés

Proposition

Toute personne qui contribue à Wikidata peut proposer une nouvelle propriété sur la page dédiée : https://www.wikidata.org/wiki/Wikidata:Property_proposal/fr

Discussion

La nouvelle propriété fait l'objet de discussions.

Création

La nouvelle propriété est créée sur Wikidata après approbation, ou rejetée

Suivi statistique



Statistiques

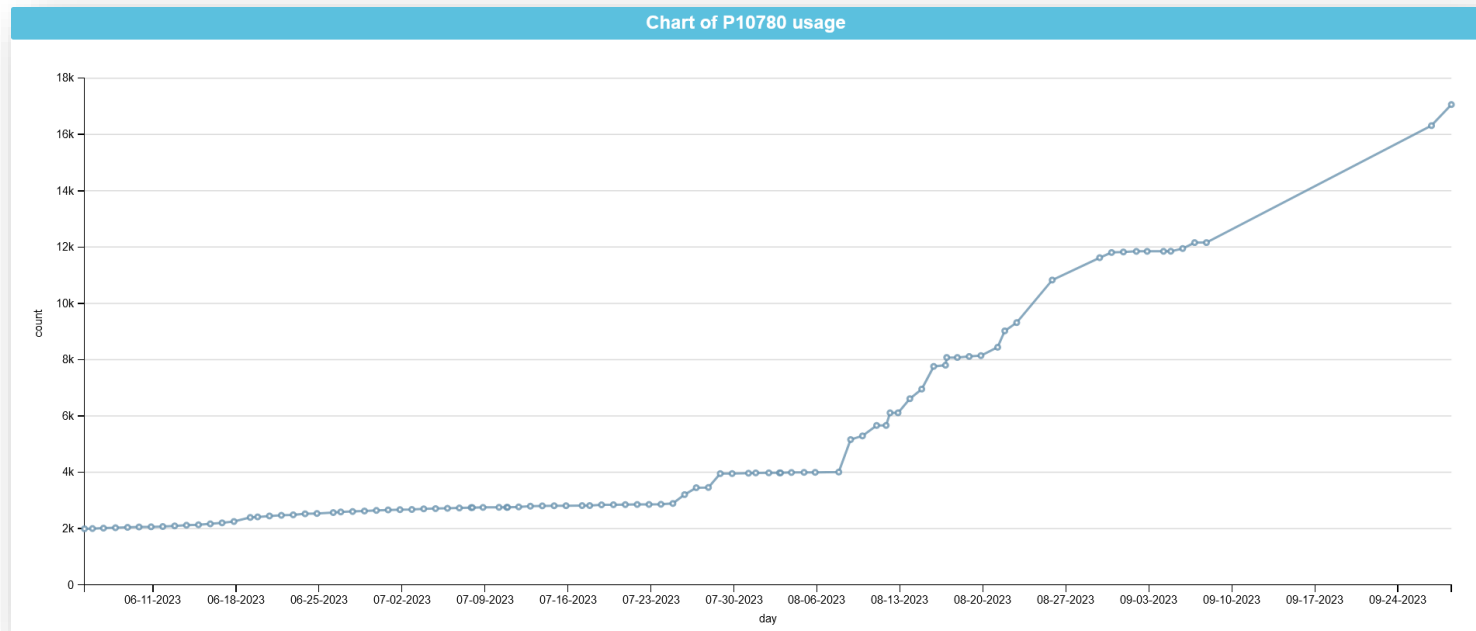
Identifiant The Conversation (P8207)

Utilisations actuelles	Total	2 912	
	Déclaration principale	2 872 ↗	98,6% d'utilisations
	Qualificatifs	4 ↗	0,1% d'utilisations
	Référence	36 ↗	1,2% d'utilisations

identifiant Radio France (P10780)

Utilisations actuelles	Total	18 130	
	Déclaration principale	18 094 sur 25 484 (complet à 71%) ↗	99,8% d'utilisations
	Qualificatifs	6 ↗	<0,1% d'utilisations
	Référence	30 ↗	0,2% d'utilisations

Historique d'utilisation id Radio France



Statistiques

Types statistics [\[modifier \]](#)

Accepted ⇅	Type	Usage ⇅
Oui	être humain (Q5)	16872
Oui	groupe de musique (Q215380)	568
Oui	groupe de rock (Q5741069)	58
Oui	duo musical (Q9212979)	51
Oui	firme (Q4830453)	48
Oui	société anonyme avec appel public à l'épargne (Q891723)	35
Oui	entreprise (Q6881511)	35

LSE Thesis Project Dashboard

Top groupings (Minimum 1 items)		Top Properties (used at least 0 times per grouping)											
Name	Count	nature de l'élément (P31)	titre (P1476)	auteur ou autrice (P50)	auteur ou autrice (chaîne) (P2093)	organisme de soutenance (P4101)	sujet ou thème principal (P921)	langue de l'œuvre, du nom ou du terme (P407)	date de publication (P577)	DOI (P356)	œuvre intégrale disponible sur (P953)	nombre de pages (P1104)	statut des droits d'auteur (P6216)
London School of Economics (Q174570)	4242	100.0% (4242)	100.0% (4242)	100.0% (4242)	0% (0)	100.0% (4242)	16.55% (702)	100.0% (4242)	100.0% (4242)	9.9% (420)	99.93% (4239)	52.99% (2248)	100.0% (4242)
No grouping	0	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)
Totals (all items)	4242	100.0% (4242)	100.0% (4242)	100.0% (4242)	0% (0)	100.0% (4242)	16.55% (702)	100.0% (4242)	100.0% (4242)	9.9% (420)	99.93% (4239)	52.99% (2248)	100.0% (4242)

Wikidata Thesis Toolkit



- London School of Economics
- Stanford Libraries
- Nouvelle Zélande

https://www.wikidata.org/wiki/Wikidata:WikiProject_Wikidata_Thesis_Toolkit

Contraintes



Contraintes



quand le système signale les erreurs

Définition ?

Règle qui définit automatiquement la cohérence des données de Wikidata.

















Exemple ?

Si vous renseignez la propriété P22 ("père") sur l'élément d'un prénom

Fonctionnement ?

Si vous enregistrez une valeur non-conforme, Wikidata signale que le triplet pose problème.

Contraintes

contrainte de propriété Wikidata 	 contrainte de type		 modifier
	relation	nature de l'élément	
	classe	être humain	
	▼ 0 référence		 ajouter une référence
	 contrainte de valeur distincte sur Wikidata		 modifier
statut de la contrainte	contrainte obligatoire		
▼ 0 référence		 ajouter une référence	
 contrainte d'utilisation unique par élément		 modifier	
statut de la contrainte	contrainte obligatoire		
▼ 0 référence		 ajouter une référence	
 contrainte de format		 modifier	
expression régulière	[1-9]d+		
statut de la contrainte	contrainte obligatoire		
▼ 0 référence		 ajouter une référence	
 contrainte d'autre propriété à avoir		 modifier	
propriété	employé(e) par		
statut de la contrainte	suggestion		
▼ 0 référence		 ajouter une référence	

Expression régulière

The screenshot shows a user interface for editing a regular expression. On the left, there is a label 'expression régulière' in a light blue box. To its right, the regular expression '[1-9]\d+' is displayed in purple text. A small icon with three dots and arrows is positioned to the left of the expression. Below the expression, there is a dropdown menu showing '0 référence'. On the right side of the interface, there are three buttons: a yellow pencil icon labeled 'modifier', a green plus icon labeled 'ajouter une référence', and another green plus icon labeled 'ajouter une valeur'.

- **1-9** : un chiffre compris entre 1 et 9
- **d+** : répéter une ou plusieurs fois le caractère qui précède

identifiant Radio France



agnes-spiquel-courdille



modifier

Problèmes potentiels



contrainte d'utilisation unique par élément [Aide](#) [Discuter](#)

Cette propriété contient généralement une seule valeur.

[+ ajouter une référence](#)



agnes-spiquel



modifier

langue de l'œuvre, du nom
ou du terme français Q150

▼ 0 référence

[+ ajouter une référence](#)

[+ ajouter une valeur](#)

identifiant The Conversation



94923



modifier

▼ 0 référence

+ ajouter une référence



AAA



modifier

Problèmes



contrainte d'utilisation unique par élément [Aide](#) [Discuter](#)

Cette propriété contient généralement une seule valeur.

contrainte de format [Aide](#) [Discuter](#)

La valeur de la propriété **identifiant The Conversation** (AAA) doit correspondre à l'expression rationnelle

`[1-9]\d+ .`

+ ajouter une référence

+ ajouter une valeur

identifiant Un
Bibliothèque S

modifier

+ ajouter une référence

Identifiants

identifiant Radio France



lysiane-sellam



 modifier

Problèmes



contrainte de langue de libellé complexe [Aide](#) [Discuter](#)

Les entités qui déclarent une propriété **identifiant Radio France** doivent aussi comporter au moins un libellé en français.

[+ ajouter une référence](#)

[+ ajouter une valeur](#)

[+ ajouter une déclaration](#)

Rapport de contraintes





Type « être humain (Q5) » : l'élément doit contenir la propriété « nature de l'élément (P31) » avec l'une des classes « être humain (Q5) » ou leurs sous-classes (définies avec [sous-classe de \(P279\)](#)). ([Aide](#))

Des exceptions sont possibles car des valeurs rares peuvent exister. Les [exceptions peuvent être spécifiées](#) en utilisant [sauf \(P2303\)](#).

Liste des violations pour cette contrainte : [Database reports/Constraint violations/P8207#Type Q5, SPARQL](#) ↗



Valeur distincte : la valeur de cette propriété devrait être différente de celles de tous les autres éléments. ([Aide](#))

Liste des violations pour cette contrainte : [Database reports/Constraint violations/P8207#Unique value, rapport actualisé toutes les heures, SPARQL \(every item\)](#) ↗, [SPARQL \(by value\)](#) ↗



Une seule valeur : cette propriété ne contient généralement qu'une seule valeur. ([Aide](#))

Liste des violations pour cette contrainte : [Database reports/Constraint violations/P8207#Single value, rapport actualisé toutes les heures, SPARQL](#) ↗



Format « [1-9]\d+ » : les valeurs doivent avoir une présentation utilisant cette forme ([syntaxe PCRE](#)). ([Aide](#))

Liste des violations pour cette contrainte : [Database reports/Constraint violations/P8207#Format, rapport actualisé toutes les heures, SPARQL](#) ↗



Élément « employé(e) par (P108) » : Les éléments avec cette propriété devraient également avoir la propriété « employé(e) par (P108) ». ([Aide](#))

Des exceptions sont possibles car des valeurs rares peuvent exister. Les [exceptions peuvent être spécifiées](#) en utilisant [sauf \(P2303\)](#).

Liste des violations pour cette contrainte : [Database reports/Constraint violations/P8207#Item P108, search](#) ↗, [SPARQL](#) ↗



Élément « scolarité (P69) » : Les éléments avec cette propriété devraient également avoir la propriété « scolarité (P69) ». ([Aide](#))

Des exceptions sont possibles car des valeurs rares peuvent exister. Les [exceptions peuvent être spécifiées](#) en utilisant [sauf \(P2303\)](#).

Liste des violations pour cette contrainte : [Database reports/Constraint violations/P8207#Item P69, search](#) ↗, [SPARQL](#) ↗



Élément « occupation (P106) » : Les éléments avec cette propriété devraient également avoir la propriété « occupation (P106) ». ([Aide](#))

Des exceptions sont possibles car des valeurs rares peuvent exister. Les [exceptions peuvent être spécifiées](#) en utilisant [sauf \(P2303\)](#).

Liste des violations pour cette contrainte : [Database reports/Constraint violations/P8207#Item P106, search](#) ↗, [SPARQL](#) ↗

Exemple de rapport de contrainte

Compte rendu de violations de contraintes pour **identifiant The Conversation** ([Discussion](#), [utilisation, éléments](#) ↗, [changements](#) ↗, [propriétés liées](#) ↗): identifier for contributors to the online news outlet The Conversation

Date du rapport : 23 octobre 2023, 11:10 (UTC) — Nombre d'éléments analysés : 2 859

Le compte rendu est généré d'après les règles énoncées sur la [page de la propriété P8207](#).

La mise à jour écrase la page. Certains éléments peuvent être corrigés depuis la dernière mise à jour : vérifiez le [suivi des liens](#).

Quand le [dump incrémentiel](#) et le bot fonctionnent comme prévu, les éléments corrigés avant 07:00 UTC n'apparaissent plus dans la mise à jour suivante.

Ce compte rendu peut contenir des faux positifs.

[Vue d'ensemble](#)

[Liste des propriétés](#)

[Propositions de propriété](#)

[Propriétés fréquentes](#)

[Special:ConstraintReport](#)

Exemple de rapport de contrainte

"Type être humain (Q5)" violations [[modifier](#)]

Violations count: 0

Types statistics [[modifier](#)]

Accepted ▾	Type ▾	Usage ▾
Oui	être humain (Q5)	2859

Exemple de rapport de contrainte

"Single value" violations [\[modifier \]](#)

Violations count: 11

- [Yoshio Sugimoto \(Q11522129\)](#): [527](#), [291](#)
- [Edward Barbier \(Q30068992\)](#): [11706](#), [423126](#)
- [Nattavudh Powdthavee \(Q30072259\)](#): [114923](#), [1032882](#)
- [Nancy Pachana \(Q38545346\)](#): [460147](#), [343917](#)
- [Simon Springer \(Q47087173\)](#): [453681](#), [771566](#)
- [Sarah Brough \(Q59670618\)](#): [163908](#), [389687](#)
- [Donald Rothwell \(Q60668397\)](#): [9843](#), [1705](#)
- [Bronwen Neil \(Q72553241\)](#): [417711](#), [322246](#)
- [Alberto Giubilini \(Q87750746\)](#): [355945](#), [278708](#)
- [Julie Redfern \(Q88769827\)](#): [19509](#), [1244929](#)
- [Ben J Marais \(Q90300421\)](#): [1018354](#), [15193](#)

"Type être humain (Q5), orchestre (Q42998), institution sociale (Q178706), groupe d'humains (Q16334295), personnage imaginaire (Q115537581)" violations [[modifier](#)]

Violations count: 9

- polythéisme (Q9163): croyance religieuse (Q2728698)
- mangaka (Q191633): profession (Q28640)
- TF1 (Q214683): chaîne de télévision (Q2001305)
- odeur (Q485537):
- Association Générale des Institutions de Retraite des Cadres (Q2826868):
- C8 (Q3011095): chaîne de télévision (Q2001305)
- usine AZF de Toulouse (Q3083316): usine (Q83405), démoli ou détruit (Q56556915)
- bijou (Q3575260):
- 13'Organisé (Q100271345): album (Q482994)

Oui	service d'hébergement d'images (Q2721136)	1
Non	croyance religieuse (Q2728698)	1
Oui	comité d'éthique médicale (Q27913684)	1
Oui	groupe de hackers (Q2801262)	1
Oui	application de rencontre en ligne (Q28134845)	1
Oui	agence française de sécurité sanitaire (Q2826640)	1
Oui	archives départementales (Q2860456)	1
Non	profession (Q28640)	1
Oui	Autorité administrative indépendante (Q2872764)	1

Outils



Outils

- Mix'n'match
- QuickStatements
- OpenRefine



Cet utilisateur adore **Wikidata**.



Cet utilisateur adore **OpenRefine**



This user uses **QuickStatements**.



This user uses **Mix'n'match**.

Wikidata et la Recherche

Mapping Diversity



- Rue Pierre-et-Marie-Curie à Paris (OpenStreetMap)
- Marie Curie
- élément Q7186 (Wikidata)
- propriété P21 « sexe ou genre »
- <https://mappingdiversity.eu>

Top Wealth and its historical origins

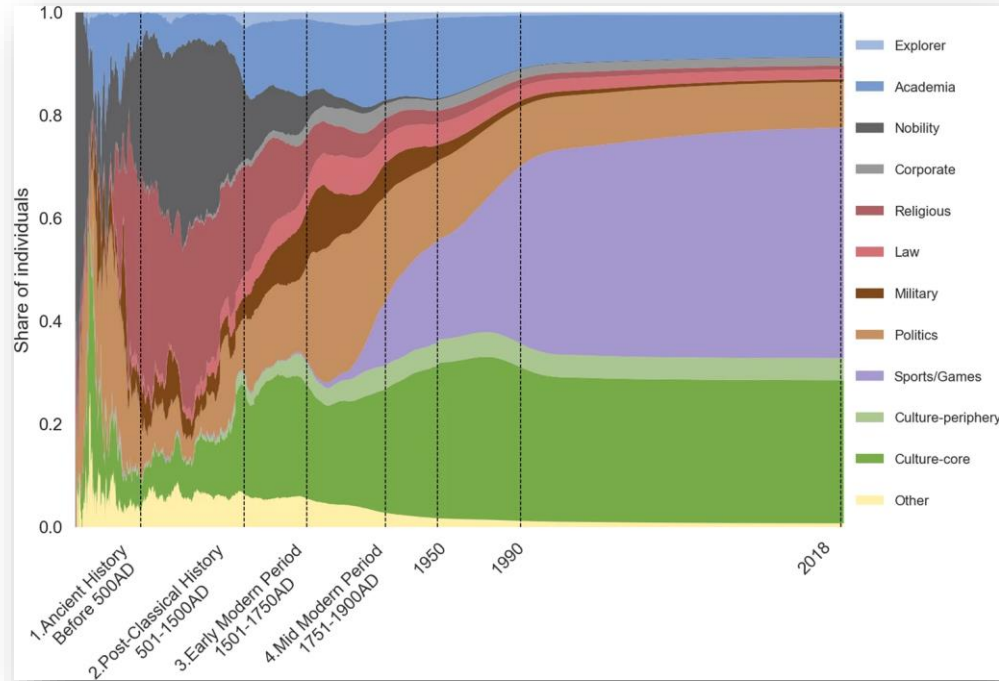
63

	<i>n</i>
Entries on the <i>Manager Magazin</i> rich list for 2019	1,032
Individuals identified in Wikidata	439 (from 394 entries)
Individuals identified in Wikidata with at least one family tie	193
Ancestors on Rudolf Martin's list for 1913 connected to at least one of the individuals on <i>Manager Magazin</i> list 2019 via family ties in Wikidata	76

Top Wealth and Its Historical Origins: Identifying Entrenched Fortunes by Linking Rich Lists over 100 Years (Daria Tisch et Emma Ischinsky 2023)

A cross-verified database of notable people

64

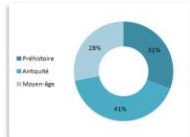


Morgane Laouenan,
Palaash Bhargava,
Jean-Benoît
Eyméoud, Olivier
Gergaud, Guillaume
Plique, Etienne
Wasmer (2002)

thésaurus PACTOLS

PACTOLS, thésaurus pour l'archéologie

35 ans d'indexation des 40 bibliothèques de la Fédération et ressources sur l'Antiquité (Frantiq, GDS du CNRS) dans le Catalogue collectif indexé.



Lexique évolutif

L'archéologie de la Préhistoire jusqu'au XXe s.
Les sciences de l'Antiquité
Europe et monde méditerranéen

60000 concepts multilingues

6 collections + 1 liste éponymes
Peuples
Anthroponymes
Chronologie
Oeuvres
Lieux
Sujets

Technique

<https://pactols.frantiq.fr>
<https://github.com/frantiq/PACTOLS>

Licence libre ODbL v1.0

ISO 25964 sur l'interopérabilité des thésaurus
SKOS
Import/exports - webservice REST

Identifiants pérennes à chaque concept :
Ark

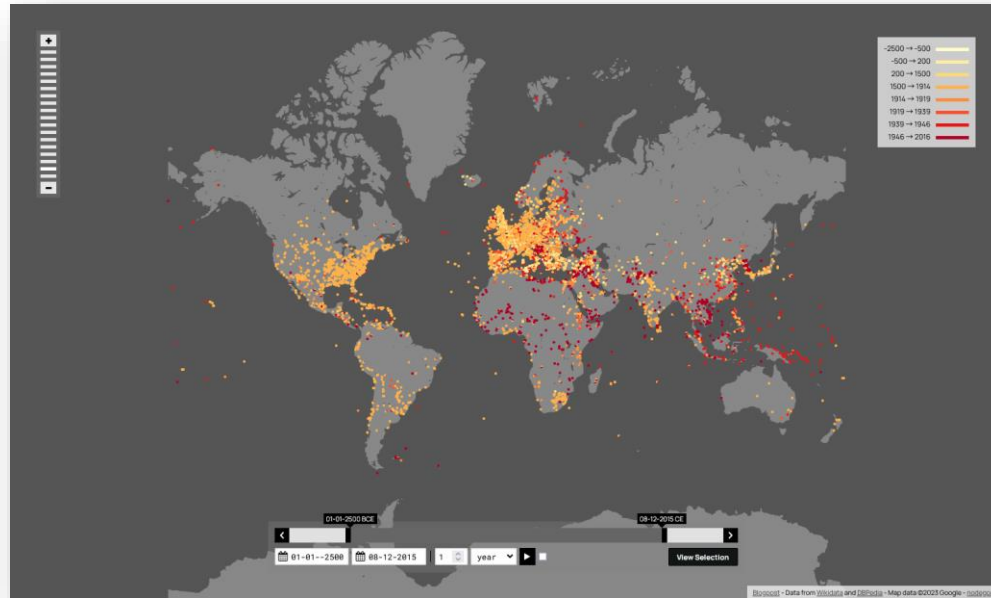


projet Arkéo MOM

Aligner pour mieux diffuser, l'expérience du thésaurus PACTOLS pour l'archéologie avec Wikidata (Christelle Molinié, Blandine-Nouvel, Miled Rousset 2021)

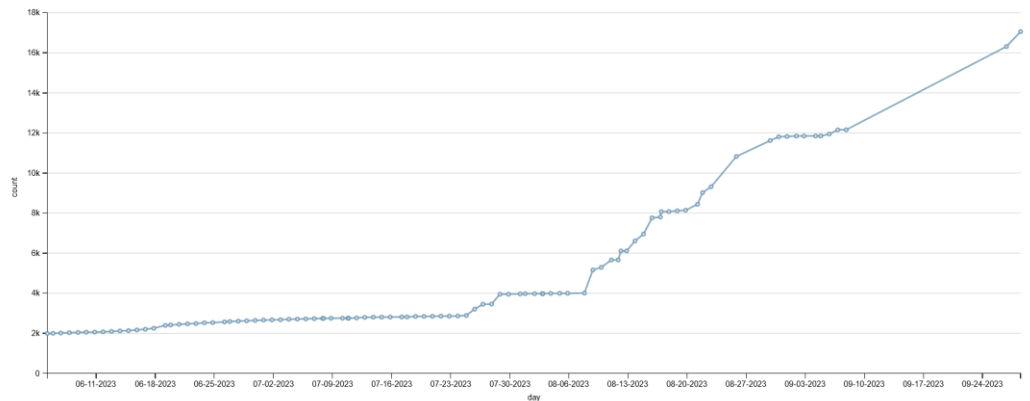
A Wikidata/DBpedia Geography of Violence

<http://battles.nodogoat.net/viewer.p/23/385/scenario/1/geo/fullscreen>



Wikipédia et les femmes scientifique (WIKIF)

67



Importer des
identifiants Radio
France : exploration,
limites et question
(Delphine Montagne
2023)

Pour en savoir plus



WIKIMÉDIA FRANCE

VOUS PRÉSENTE LE

[MOOC]
Wikidata

formations.wikimedia.fr

The graphic features a background of fine, wavy grey lines. A red rectangular border frames the central text. The text 'WIKIMÉDIA FRANCE' is in a bold, black, sans-serif font. Below it, 'VOUS PRÉSENTE LE' is in a smaller, black, sans-serif font. The word '[MOOC]' is in a very large, bold, black, sans-serif font, with the letters enclosed in square brackets. Below '[MOOC]', the word 'Wikidata' is written in a black, serif font. At the bottom, a solid red horizontal bar contains the URL 'formations.wikimedia.fr' in a white, sans-serif font.



weburfist.univ-bordeaux.fr



[@UrfistBordeaux](https://twitter.com/UrfistBordeaux)



urfist@u-bordeaux.fr

Restons en contact :
pierre-yves.beaudouin@u-bordeaux.fr

 **URFIST**
bordeaux

 université
de **BORDEAUX**