



**BorderCloud**

Web 3.0 Hosting & Factory

# Input/Output Data Science à l'université Paris Saclay

## Comment utiliser le Web des données ?

Karima Rafes, BorderCloud

01/02/2016



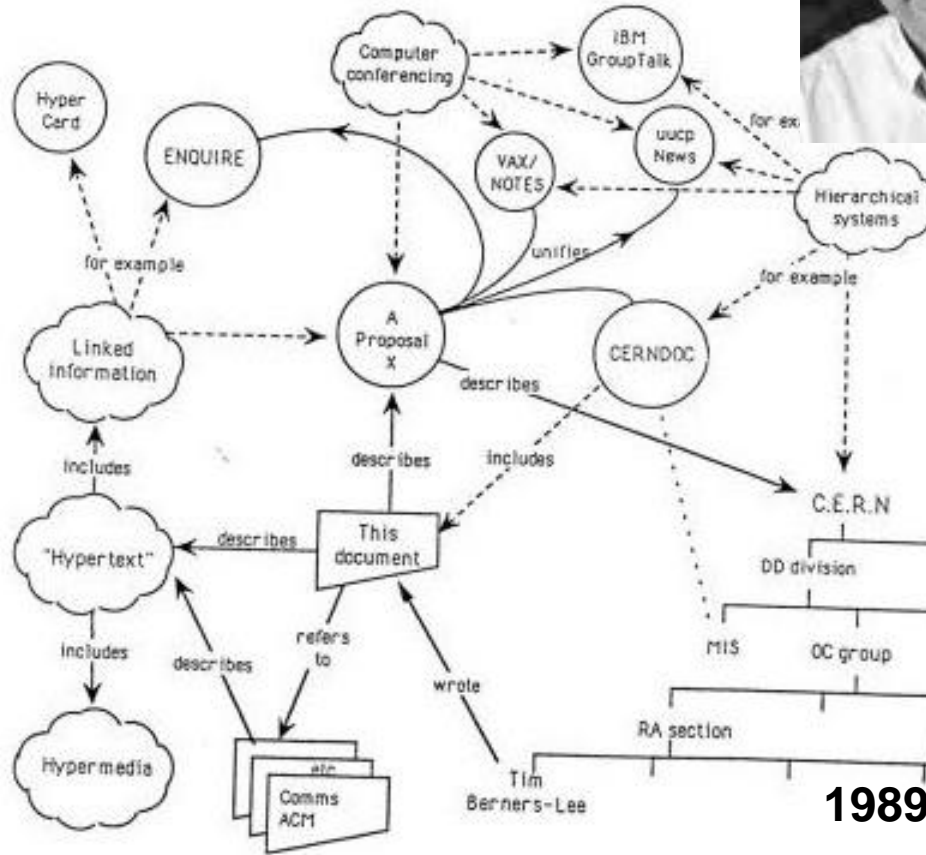
# Sommaire

1. Historique du mouvement
  2. Définition
  3. Constituer un réseau global d'informations
  4. Comment accéder aux données de Wikidata via son propre langage de programmation
  5. Ecrire sa première requête SPARQL avec Wikidata
- Conclusion



$$1 + 1 = 3$$

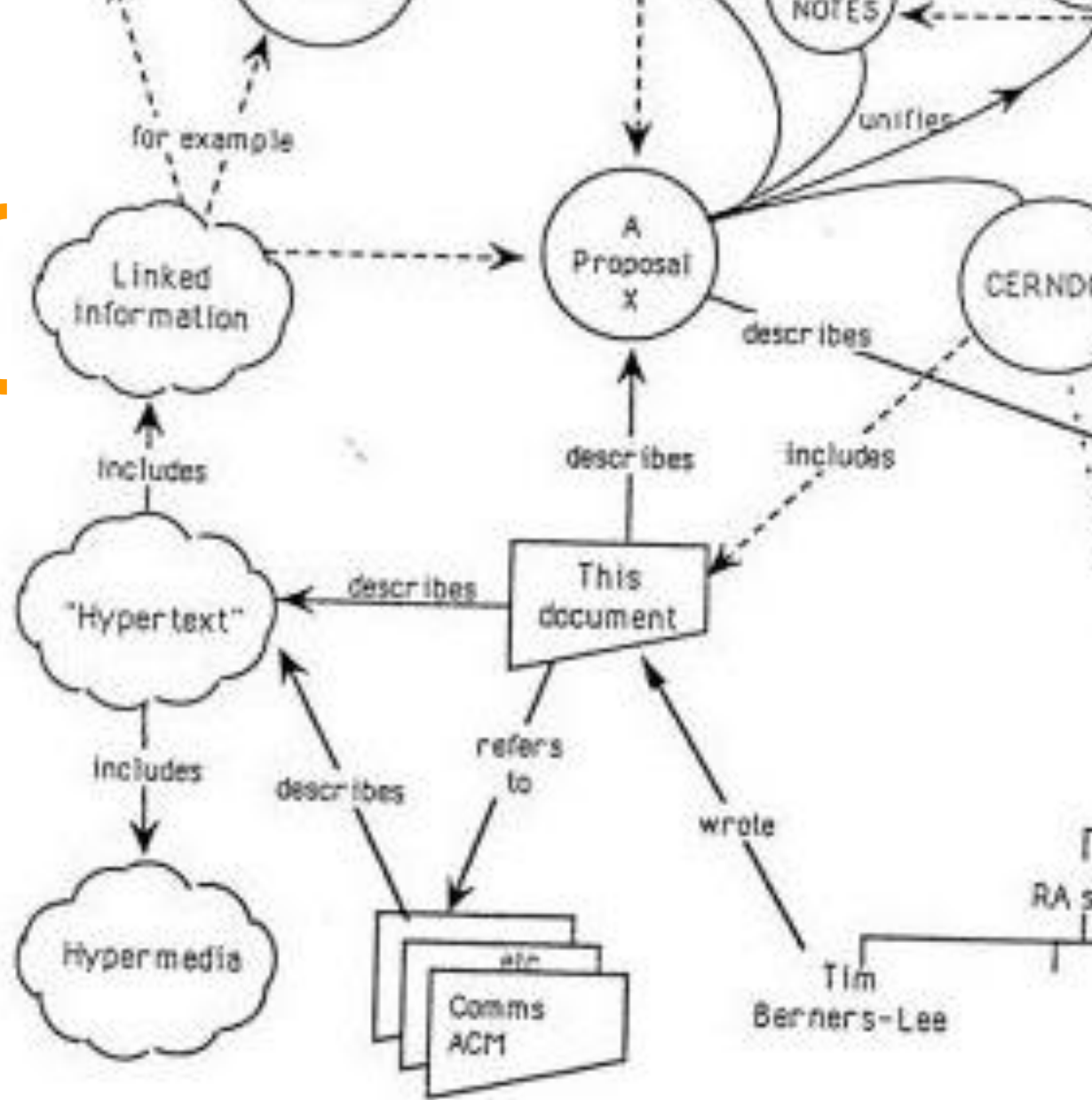
1. Historique du mouvement



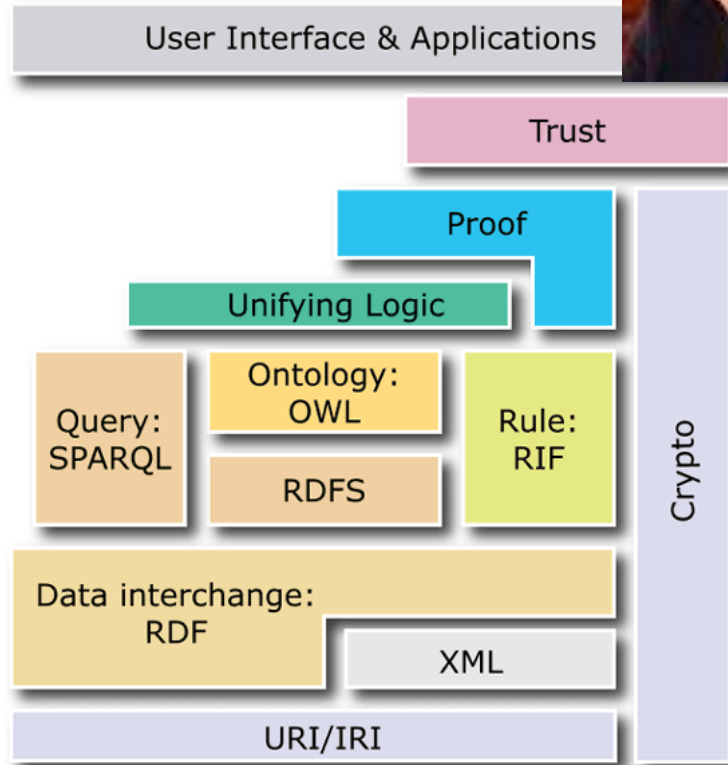
Si on reliait tous les documents sur Internet !?

**Et le Web fut ?**

**Le concept "informations liées" apparaît dès les origines.**



**Dès ses origines, le Web n'avait pas la finalité de relier seulement des documents mais de relier les informations.**



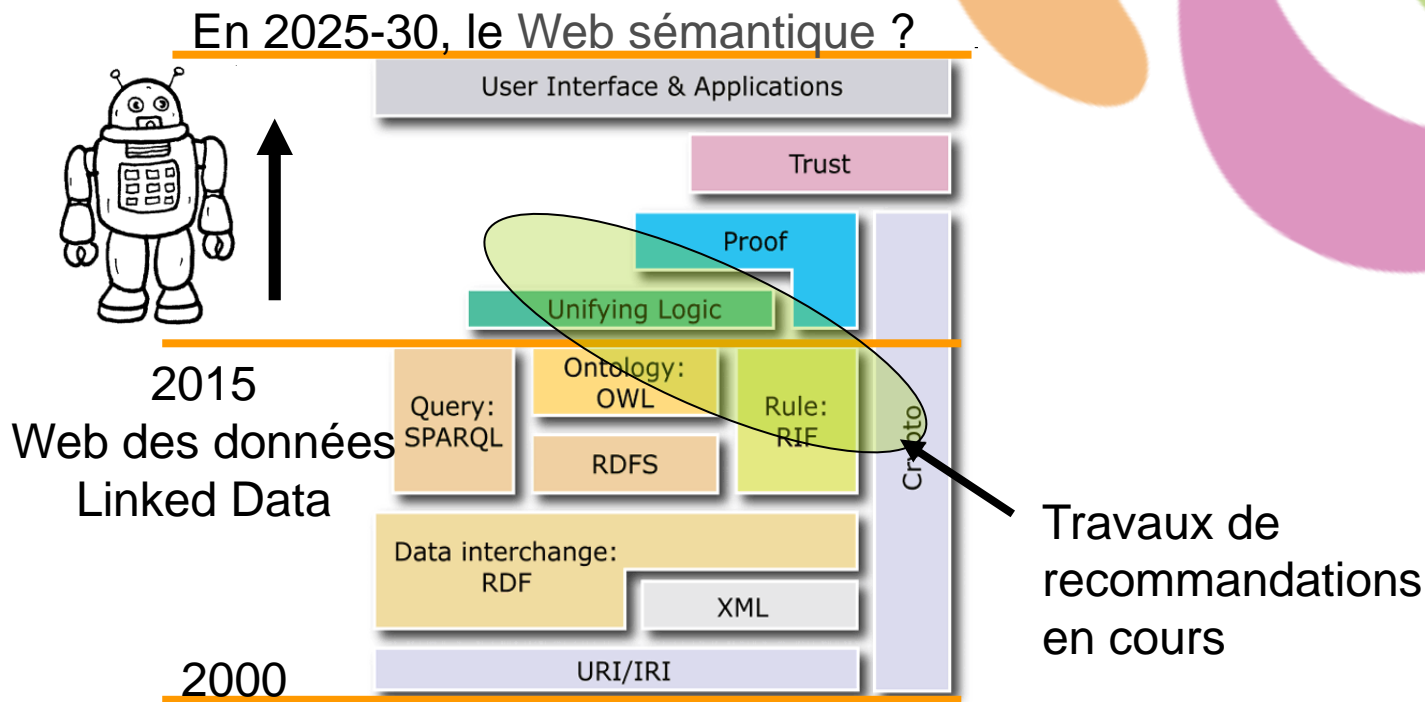
**En 1999, le concept de Web sémantique est créé/structuré pour décrire la finalité du Web mais le concept fut mal compris et souvent détourné.**

# Objectif du Web sémantique



Pouvoir créer de nouvelles connaissances  
avec l'aide de machines.

# Le Web des données... en route vers le Web sémantique



**En 2006, on crée un point d'étape au Web sémantique que l'on nomme le Web des données ou Linked Data. Ce concept offre un objectif plus simple à comprendre/expliciter et moins utopique car il est de + en + facile de faire des démonstrations.**



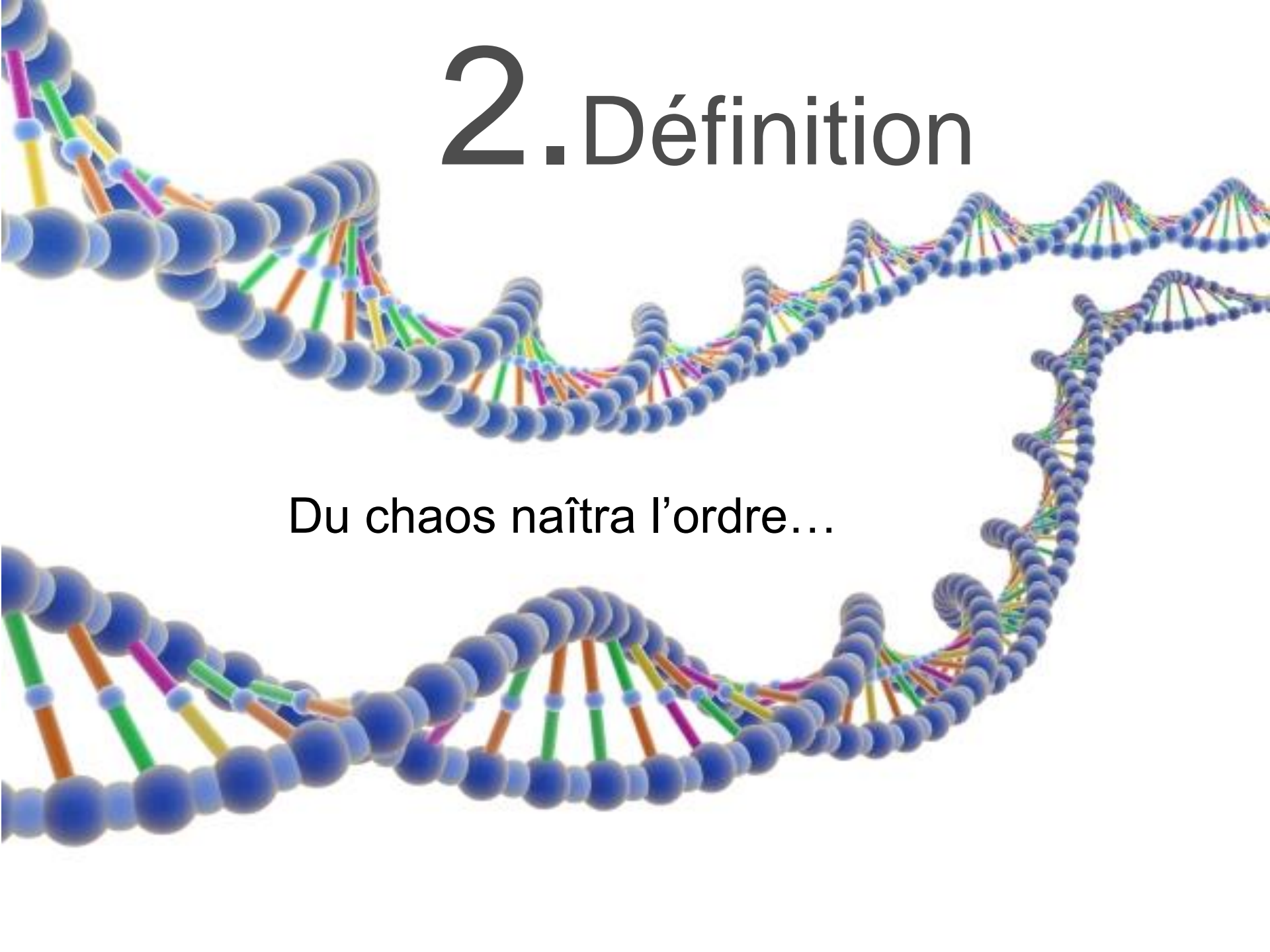
# Objectif du Web des données



Pouvoir accéder aux informations disponibles  
avec l'aide de machines.

# 2. Définition

Du chaos naître l'ordre...



# Définition :

## le Web de données ou Linked Data

Initiative visant à favoriser la publication de données :

- 🌀 **Interopérables** à l'aide d'un modèle commun pour représenter une information
- 🌀 **Structurées** à l'aide d'ontologies
- 🌀 **Accessibles** à travers le Web
- 🌀 **Reliées** entre elles pour constituer un réseau global d'informations



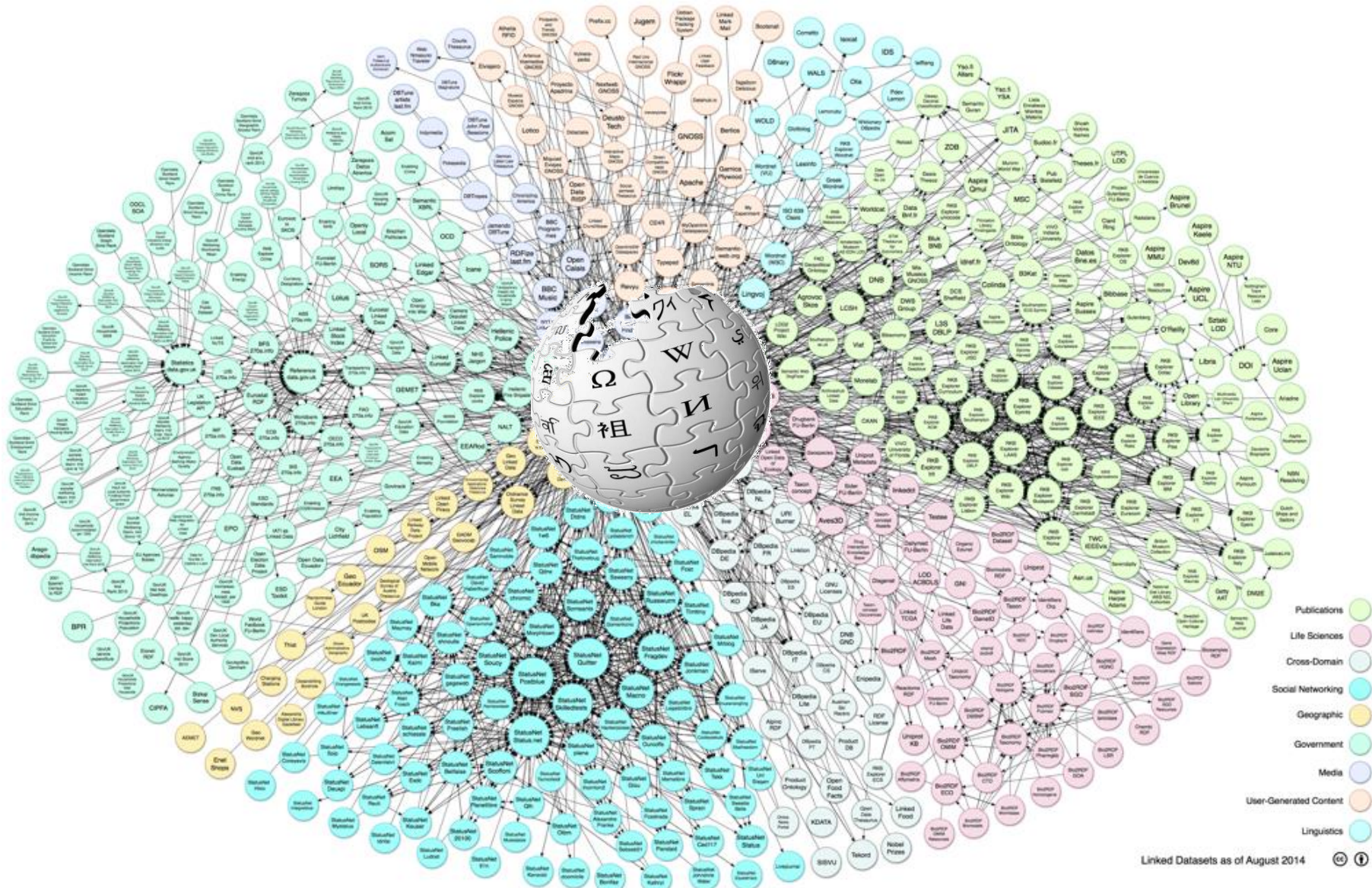
3. Constituer un réseau global d'informations en agissant localement

# Projets locaux

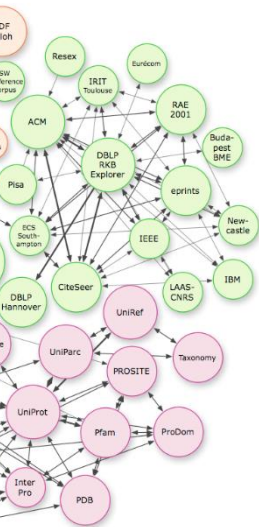


- 🌀 Bases de connaissances
    - Grégorius, INRIA, YAGO,...
  - 🌀 Annuaires des scientifiques, du personnel de l'université et des projets
    - LRI, Adonis,...
  - 🌀 Bases de traçabilité de la production des résultats de la recherche
    - DAAP,...
- ➔ Besoin d'un référentiel commun pour éviter de réinventer la roue.  
Eviter de décrire le monde à nouveau.

# Le Web des données a-t-il déjà ce type de référentiel ?



# Ontologies dynamiques et partagées



As of March 2009



# Demo 1.1 : Les Wikipediens décrivent le monde

**WIKIPEDIA**  
The Free Encyclopedia

[Main page](#)  
[Contents](#)  
[Featured content](#)  
[Current events](#)  
[Random article](#)  
[Donate to Wikipedia](#)  
[Wikipedia store](#)

**Interaction**  
[Help](#)  
[About Wikipedia](#)  
[Community portal](#)  
[Recent changes](#)  
[Contact page](#)

**Tools**  
[What links here](#)  
[Related changes](#)  
[Upload file](#)  
[Special pages](#)  
[Permanent link](#)  
[Page information](#)  
[Wikidata item](#)

## Dalmatian (dog) (Q17504 ttl rdf/xml json sparql)

From Wikipedia, the free encyclopedia

The **Dalmatian** is a breed of large dog,<sup>[3]</sup> noted for its unique black or liver spotted coat and mainly used as a carriage dog in its early days. Its roots trace back to Croatia and its historical region of Dalmatia.<sup>[4][5]</sup> Today, it is a popular family pet, and many dog enthusiasts enter Dalmatians into kennel club competitions.

### Contents [hide]

- Characteristics
  - Body
  - Coat
  - Uses
  - Litter size
- Health
  - Deafness
  - Hyperuricemia
    - Dalmatian-Pointer Backcross Project
    - The Dalmatian Heritage Project
- History
  - Duties
- In popular culture

### Dalmatian



Black-spotted Dalmatian

**Other names**  
Carriage Dog  
Spotted Coach Dog  
Firehouse Dog  
Plum Pudding Dog

**Nicknames**  
Dal

**Country of origin**  
Croatia<sup>[1]</sup>

### Traits [hide]

**Height**  
Male 23–24 inches (58–61 cm)<sup>[2]</sup>



# Demo 1.2 : Wikidata extrait les données structurées pour les mettre à disposition des systèmes en CC0



- Main page
- Community portal
- Project chat
- Create a new item
- Item by title
- Recent changes
- Random item
- Random Primary Sources item
- Help
- Donate
- Print/export
  - Create a book
  - Download as PDF
  - Printable version
- Tools
  - What links here
  - Related changes
  - Special pages
  - Permanent link
  - Page information
  - Concept URI
  - Cite this page
  - Primary Sources list

Item [Discussion](#)

[Read](#)

[View history](#)



[More](#) ▾

## Dalmatian (Q17504)

Dog Breed

[\[edit\]](#)

No aliases defined

▾ In more languages

Language	Label	Description	Also known as
English	Dalmatian	Dog Breed	
French	dalmatien	race de chiens	

### Statements

subclass of	<a href="#">dog</a> <a href="#">[edit]</a>
	▾ 0 references <a href="#">[add reference]</a>
-----	
<a href="#">[add]</a>	
instance of	<a href="#">dog breed</a> <a href="#">[edit]</a>
	▶ 1 reference
-----	
<a href="#">[add]</a>	

# Wikidata en clair

Wikidata constitue une base de données :

- 🌀 D'identifiants uniques sur le Web
  - 🌀 Avec une ontologie dynamique variant au grès des découvertes et de l'actualité.
  - 🌀 Licence CC0 : réutilisable gratuitement par toute la société (entreprises incluses)
  - 🌀 Modification ouverte à tous
- ➔ Des millions de concepts sont déjà dans ce référentiel

# Atelier 1 : Wikidata et Paris Saclay

- 🌀 Ce référentiel peut simplifier la fabrication de nombreux services.
- 🌀 Le premier d'entre eux est de réutiliser la force de Wikipedia pour partager des données utiles au sein de l'université.
- 🌀 Par exemple : retrouver les données, les appareils de mesures, les scientifiques, les équipes, les projets de l'université, etc.

➔ Demo

# Demo : taper un mot dans Google et ouvrir la page Wikipédia pour retrouver les données de l'université à propos de ce sujet.



WIKIPÉDIA  
L'encyclopédie libre

- Accueil
- Portails thématiques
- Article au hasard
- Contact
- Contribuer
- Débuter sur Wikipédia
- Aide
- Communauté
- Modifications récentes
- Faire un don

- Imprimer / exporter
- Créer un livre
- Télécharger comme PDF
- Version imprimable

- Outils
- Pages liées
- Suivi des pages liées
- Importer un fichier
- Pages spéciales
- Adresse permanente
- Information sur la page
- Élément Wikidata
- Citer cette page
- Vérif. ébauches
- Vérif. homonymies

Article Discussion Research Scientist Device Lab Lire Modifier Modifier le code Plus

Rechercher

## Web sémantique (Q54837 ttl rdf/xml json sparql)

### Datasets in relation

- [DBPedia](#) : This is a copy of DBPedia 2014. Only the french and english instances have been loaded for a total of ~600m triples. (source)
- [Wikidata](#) : Wikidata is a free linked database for the structured data of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wikisource and others. The content is available under a free license, exported using standard formats, and can be interlinked to other open data sets on the linked data web. (source)

### Scientists in relation

- [Denis HUMBERT](#) : (email)
- [Karima RAFES](#) : (email)

### Inria teams in relation

Automated reasoning	Data exchange	Intelligent agents	Knowledge representation and reasoning
Markup languages	Query languages	Resource Description Framework (RDF)	Semantic networks
Social tagging	Vagueness and fuzzy logic	Web Ontology Language (OWL)	Web and social media search
Web search engines			

*Ne doit pas être confondu avec Sémantique des langages de programmation.*

Le **Web sémantique**, ou **toile sémantique**<sup>1</sup>, est une extension du Web standardisée par le World Wide Web Consortium (W3C)<sup>2</sup>. Ces standards encouragent l'utilisation de formats de données et de protocoles d'échange normés sur le Web, avec comme format de base le Resource Description Framework (RDF).

Selon le W3C, le Web sémantique fournit un modèle qui permet aux données d'être



Logo du W3C pour le Web sémantique

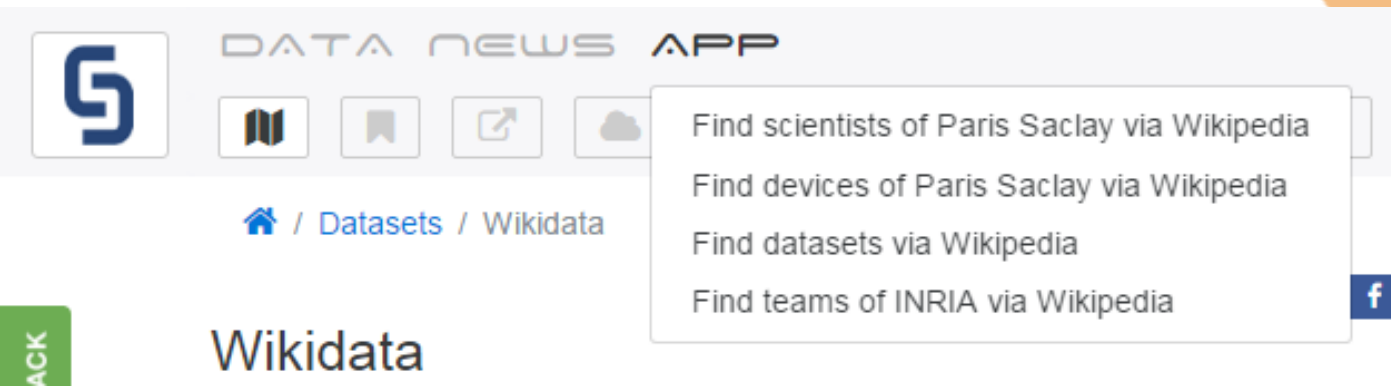


# Atelier 1 : Wikidata et Paris Saclay



1. Aller sur Wikipedia
2. Se connecter ou s'inscrire si ce n'est pas déjà fait.
3. Aller sur le site <https://io.datascience-paris-saclay.fr>
4. Dans le menu APP : ouvrir le gadget qui vous intéresse.
5. Suivre les instructions pour installer le Gadget « Trouver les datasets de Paris Saclay »

# Atelier 1 : Wikidata et Paris Saclay



The screenshot shows the Wikidata interface. At the top, there is a navigation bar with the Wikidata logo (a blue 'S' inside a square) and the text "DATA NEWS APP". Below this are several icons: a book, a speech bubble, a share icon, and a cloud icon. A search bar is present, and a dropdown menu is open, displaying the following suggestions:

- Find scientists of Paris Saclay via Wikipedia
- Find devices of Paris Saclay via Wikipedia
- Find datasets via Wikipedia
- Find teams of INRIA via Wikipedia

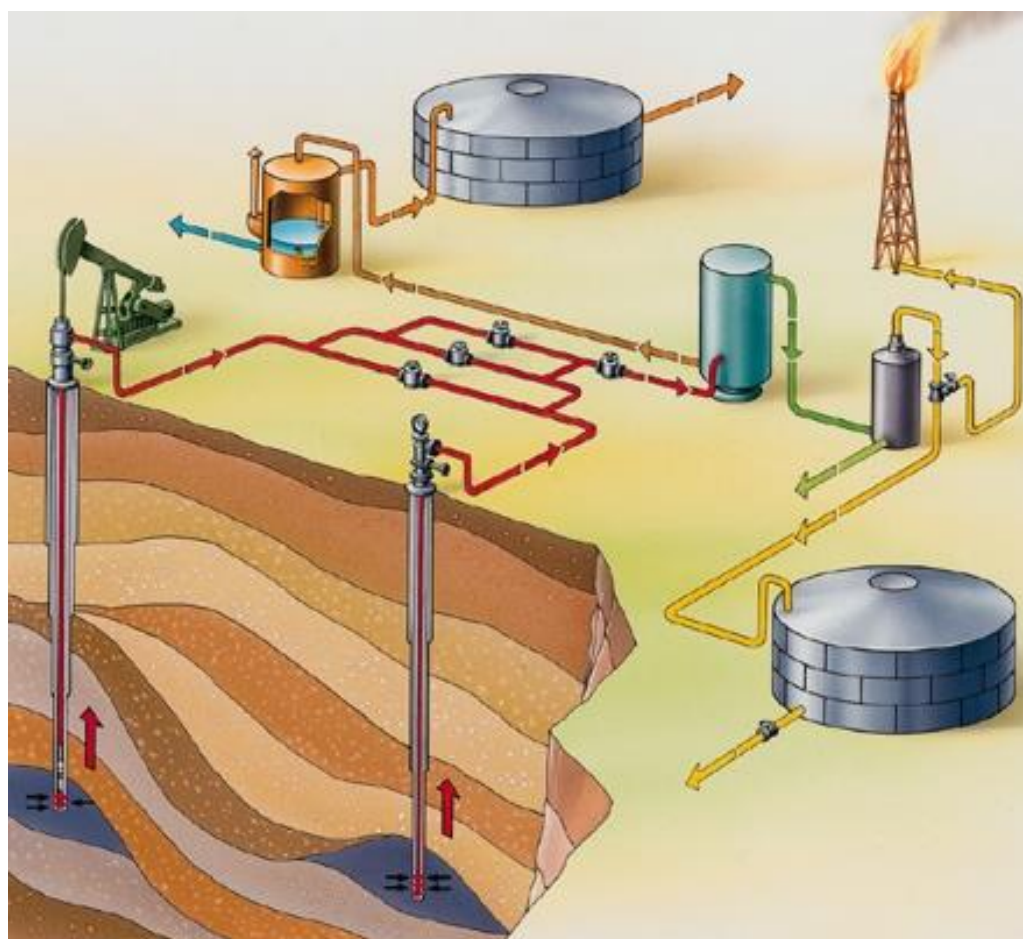
Below the search bar, the breadcrumb navigation reads "Home / Datasets / Wikidata". The main heading "Wikidata" is visible. On the left side, there is a green button labeled "ACK". On the right side, there is a blue Facebook icon.



User page [Discussion](#) [Read](#) [Edit](#) [View history](#)

## User:Karima Rafes/global.js

[< User:Karima Rafes](#)

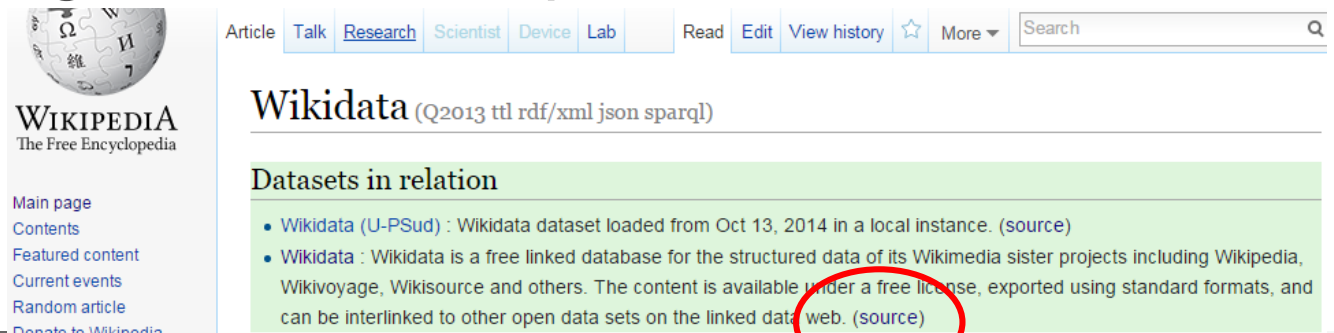


4. Comment accéder aux données de Wikidata ?

# Atelier 2 :

## Distributions de Wikidata

1. Aller sur le site <https://io.datascience-paris-saclay.fr>
2. Utiliser le champ recherche et taper Wikidata  
Ou (si vous avez enregistré le gadget de ParisSaclay)
1. Passer par Google
2. Ouvrir la page Wikipedia de Wikidata
3. Ouvrir l'onglet recherche pour accéder aux sources.



Article [Talk](#) [Research](#) [Scientist](#) [Device](#) [Lab](#) [Read](#) [Edit](#) [View history](#) [☆](#) [More](#)

### Wikidata (Q2013 ttl rdf/xml json sparql)

#### Datasets in relation

- [Wikidata \(U-PSud\)](#) : Wikidata dataset loaded from Oct 13, 2014 in a local instance. (source)
- [Wikidata](#) : Wikidata is a free linked database for the structured data of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wikisource and others. The content is available under a free license, exported using standard formats, and can be interlinked to other open data sets on the linked data web. (source)



# Wikidata

[general knowledge](#) [Semantic Web](#) [Wikipedia](#) [Wikidata](#) [Wikivoyage](#) [Wikisource](#)

### Description

Wikidata is a free linked database for the structured data of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wikisource and others.

The content is available under a free license, exported using standard formats, and can be interlinked to other open data sets on the Internet data base.

<https://www.wikidata.org> [karima.rafes@gmail.com](mailto:karima.rafes@gmail.com)

### Location

Références de Wikidata.  
Domaine scientifique (bleu clair)  
Tag décrivant le contenu des données (bleu foncé)

Description

### Distribution

SPARQL endpoint [CC0](#) [srx](#) [json](#) [oqv](#) 2018-01-31 [Details](#) [Q](#) [Go](#)

RDF Exports [CC0](#) [nt](#) [gz](#) [json](#) 2018-01-31 [Details](#) [Go](#)

Distributions  
C'est-à-dire les moyens d'y accéder

### Publishers by Occupation

[science](#) [computer science](#) [scientist](#) [profession](#)

This query retrieves the number of scientists publishing in Wikidata.

5

Exemples illustrant les données



Le nb d'étoiles correspond à la capacité de réutiliser ces données sans avoir besoin de les traiter

Détail d'accès  
Chercher des exemples d'utilisation

Editer une nouvelle requête

Ouvrir le formulaire officiel

3★ DATA	SPARQL endpoint	CC0	srx	json	csv	2016-01-31	Details	🔍	📄	🔗 Go
2★ WWW	RDF Exports	CC0	nt	gz	json	2016-01-31	Details			🔗 Go

Type d'accès :  
WWW site Web  
Torrent via P2P  
File : fichier  
DATA via une API

Pour accéder au site Web ou aux fichiers

De nombreux exemples existent sur le site officiel  
<https://query.wikidata.org>

The image shows a screenshot of the Wikidata Query Service interface. A modal window titled "SPARQL Query Examples" is open, displaying a list of query examples. The window has a search bar at the top with the placeholder text "Type to filter" and a search icon. Below the search bar, there is a list of query examples, each with a title, an eye icon, and a pencil icon. The examples are:

- US presidents and spouses
- List of presidents with causes of death
- Causes of death ranking
- Politicians who died of cancer (of any type)
- People born before year 1880 with no death date
- Largest cities with female mayor
- List of countries ordered by the number of their cities with female mayor
- How many states this US state borders
- Whose birthday is today?
- Who discovered the most asteroids?
- Who discovered the most planets?
- American universities founded before the states they reside in were created
- What is the relation between Terrell Buckley and Miami Dolphins?

The background shows the Wikidata Query Service interface with a "Beta" badge and a "Execute" button.

# Recherche multicritère : sélectionner un exemple

**Domains** ☰

- 1 aerospace engi...
- 1 branch of scie...
- 4 computer science
- 1 cosmology
- 2 ecology
- 1 economics
- 3 economy
- 1 environmental ...
- 2 geography

**Tag** ☰

- 1 asteroid
- 1 automobile
- 1 digital library
- 1 European Union
- 2 female
- 1 file format
- 1 France

Search

+ ☰

32 results (0.07 seconds)

## Aires protégées par taille

écologie aire protégéeProtected areas sorted by size, whatever the unit (km<sup>2</sup>, mi<sup>2</sup>, ha, ac)

## Asteroids discoverers by countries

cosmology asteroid

Asteroids discoverers by countries

## Automobile manufacturers by country

economy automobile

Piechart of automobile manufacturers by country.

## Cemetery of Père-Lachaise of Paris - year of death

Wikidata Wikidata

Cimetière du Père-Lachaise - année de décès

Personnalités enterrées au cimetière du Père-Lachaise (hors columbarium) par année de décès

# Un exemple est en 4 parties

## Asteroids discoverers by countries

cosmology asteroid

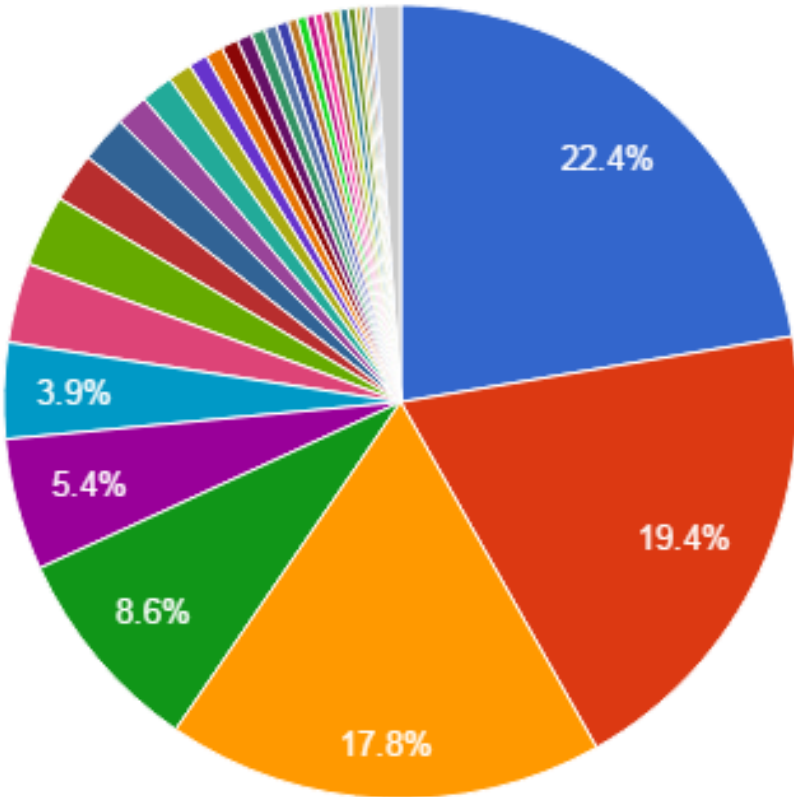
Asteroids discoverers by countries

Références de Wikidata.  
Domaine scientifique (bleu clair)  
Tag décrivant le contenu des données (bleu foncé)

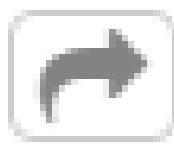
Description

Outils

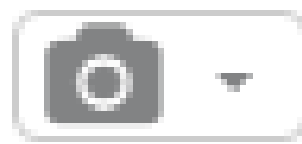
Table ou graphique



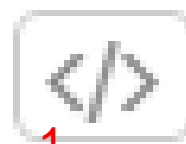
- United States of America
  - Netherlands
  - Japan
  - Belgium
  - Germany
  - Soviet Union
  - Italy
  - Czech Republic
  - Australia
  - Russia
  - France
  - Croatia
  - Ukraine
  - Canada
- ▲ 1/3 ▼



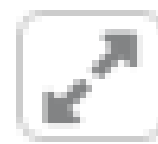
Partager



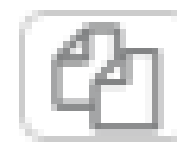
Capturer



Réutiliser



Plein écran



Faire une nouvelle requête en partant de cet exemple

Reuse these data in your code

Query, endpoint and code for reusing the same data

SPARQL Javascript HTML Matlab Python R PHP Wiki

Choisir son langage préféré

```

from SPARQLWrapper import SPARQLWrapper, JSON

sparql = SPARQLWrapper("https://query.wikidata.org/bigdata/namespace/wdq/sparql")
sparql.setQuery("""
    PREFIX bd: <http://www.bigdata.com/rdf#>
    PREFIX wikibase: <http://wikiba.se/ontology#>
    PREFIX wd: <http://www.wikidata.org/entity/>
    PREFIX wdt: <http://www.wikidata.org/prop/direct/>

    select ?countryLabel (COUNT(?asteroid) as ?nb)
    where {
        ?asteroid wdt:P31 wd:Q3863 ;
            wdt:P61 ?discoverer .
        ?discoverer wdt:P27 ?country .

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

    GROUP BY ?countryLabel
    ORDER BY DESC(?nb)""")
sparql.setReturnFormat(JSON)
results = sparql.query().convert()

for result in results["results"]["bindings"]:
    print(result["label"]["value"])

```

Copy

3 Copier dans son code en vérifiant que les dépendances fonctionnent

Méthode pour récupérer les données d'un exemple sans avoir besoin de le coder

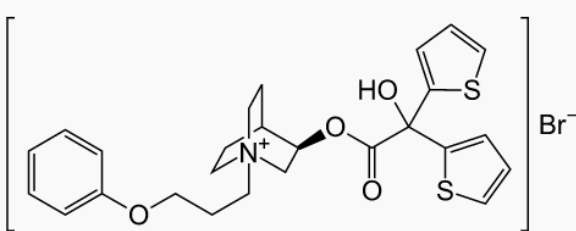
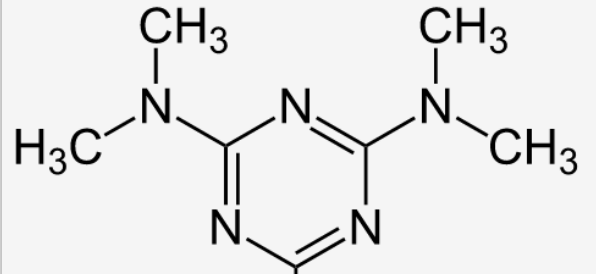
Howto use SPARQL with Python ? 4

Lire la doc si nécessaire

# LinkedWiki extension

pour afficher vos données dans votre Wiki

Wikidata : pharmaceutical drug

Chemical formula	Date	Structure
$C_{26}H_{30}NO_4S_2$		
$C_9H_{13}N_6$		

Sputnik 1 [\[w\]](#)

1957



English Create account Log in

Discussion Read Edit Edit source View history Search

WIKI LOVES AFRICA Wiki Loves Africa: share African cultural fashion and adornment pictures with the world!

ets

MediaWiki extensions manual

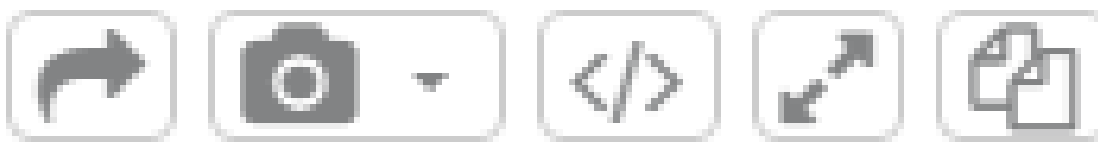
**LinkedWiki**  
Release status: stable

 **LinkedWiki**

**Implementation** Parser extension, Special page, Data extraction

**Description** Print array or widget with a query Sparql on your Semantic MediaWiki, including a SPARQL Endpoint(optional)

**Author(s)** Karima Rafes [\[w\]](#) (Karima Rafes<sup>talk</sup>)



1 Faire une nouvelle requête en partant de cet exemple

## Copy of Asteroids discoverers by countries

Title

Domain

Tag

Short description

Dataset

Query  English

A travers le service, vous pouvez partir d'un exemple pour en proposer un nouveau.

➔ Seul contrainte, savoir écrire une requête SPARQL



```
SELECT ?variable
WHERE {
    BIND("Hello World" as ?variable)
}
```

5. Ecrire sa première requête SPARQL avec Wikidata ?

# L'ambition de SPARQL



- 🌀 Offrir une interopérabilité au niveau des données structurées ou non à travers l'Internet
  - L'utilisation des services Web ou de service REST demande une connaissance de l'API pour accéder aux données mais pas avec SPARQL
  - Un agent peut découvrir les données et les interroger à condition de comprendre leurs ontologies
- 🌀 Les données disponibles via SPARQL sont ce que l'on nomme le Web des données.

# Objectifs

Apprendre à interroger Wikidata.

- Base du langage SPARQL
- Sélection les instances d'un concept de Wikidata
- Sélectionner des propriétés

## SPARQL endpoint

CC0

srx

json

csv

2016-01-31

Details



Go

Please note that the service is currently in beta

3★DATA

mode, which means that details of the data set or the service provided can change without prior warning.

SPARQL endpoint :

<https://query.wikidata.org/bigdata/namespace/wdq/sparql>

Copy

La seule chose à préciser pour accéder à un endpoint SPARQL public. C'est son adresse sur le Web.

Un agent ou un humain peut à partir de cette information découvrir l'ontologie qu'il utilise et découvrir les données qu'ils possèdent.

# Copy of Asteroids discoverers by countries

Title

Domain

Tag

Short description

Dataset

Query 17 } .  
18  
19 GROUP BY ?countryLabel  
20 ORDER BY DESC(?nb)"} English

Press **Ctrl + Space** to activate auto completion.

Chart

Options

Style

Description de l'exemple

Sélection du Endpoint SPARQL

Requête SPARQL

Tester une requête, documentation

Zone de configuration du graphique

# IRI unique

## (Internationalized Resource Identifier)



- Chaque chose doit être représentée par un IRI.
- Un IRI doit être :
  - Unique sur le Web
  - Valide : un IRI dans un navigateur doit afficher un document
  - Lisible : (si possible) dans la langue de l'être humain qui l'a créé
  - Lié : si vous décrivez une chose qui est déjà décrite sur le web avec un IRI, il faut lier les IRIs.
- Un IRI pointe sur un document qui doit également être lisible pour les humains et les machines. Un IRI affiche
  - une page HTML pour un être humain,
  - un document RDF pour une machine.

# IRIs



*Write full IRIs:*

```
<http://this.is.a/full/URI/written#out>
```

*Abbreviate IRIs with prefixes:*

```
PREFIX foo: <http://this.is.a/URI/prefix#>
```

```
... foo:bar ...
```

```
⇒ http://this.is.a/URI/prefix#bar
```

*Shortcuts:*

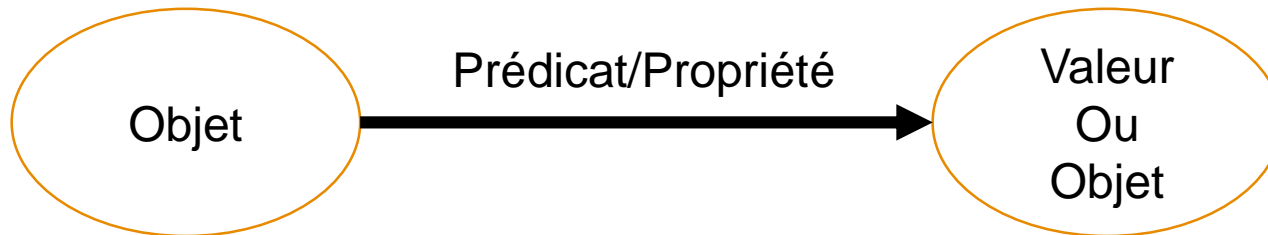
```
a ⇒ rdf:type
```

# RDF

(Resource Description Framework)

## Triplets

- Modèle de graphe destiné à **décrire de façon formelle les ressources Web** et leurs métadonnées, de façon à permettre le traitement automatique (par des machines) de telles descriptions.
- Un document structuré en RDF est constitué d'un ensemble de triplets.**



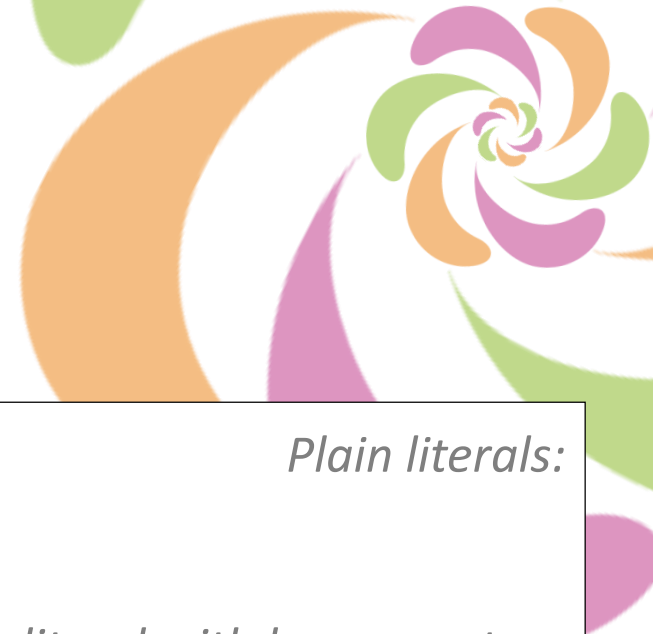
[https://en.wikipedia.org/wiki/Voyager\\_1](https://en.wikipedia.org/wiki/Voyager_1)

<https://www.wikidata.org/wiki/Property:P18> (Image)

<https://commons.wikimedia.org/wiki/File:Voyager.jpg>



# Littéraux



*Plain literals:*

`"a plain literal"`

*Plain literal with language tag:*

`"bonjour"@fr`

*Typed literal:*

`"13"^^xsd:integer`

*Shortcuts:*

`true`

⇒ `"true"^^xsd:boolean`

`3`

⇒ `"3"^^xsd:integer`

`4.2`

⇒ `"4.2"^^xsd:decimal`

`30/05/2011`

⇒ `"2011-05-30"^^xsd:date`

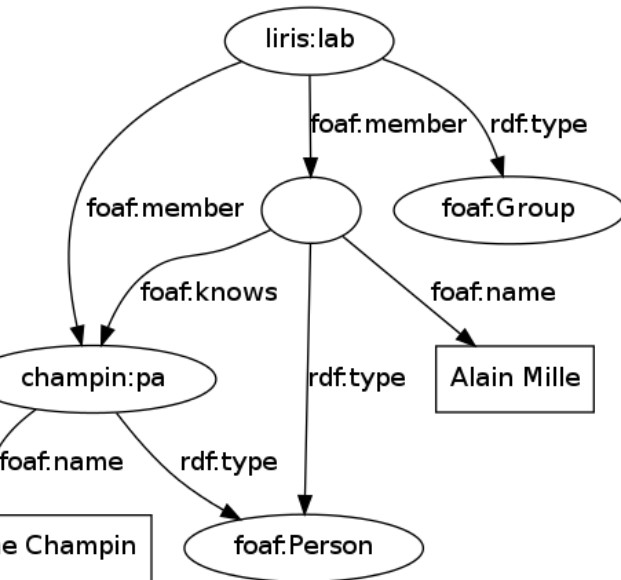
`30/05/2011 12:00`

⇒ `"2011-05-30T12:00:00+02:00"^^xsd:datetime`

# Les structures de représentation des faits avec des ontologies



A travers RDF :



- Les données peuvent prendre n'importe quelle structure. On parle aussi d'ontologie.
- Plusieurs structures/ontologies peuvent exister simultanément dans la même BDD

```
1 PREFIX bd: <http://www.bigdata.com/rdf#>
2 PREFIX wikibase: <http://wikiba.se/ontology#>
3 PREFIX wd: <http://www.wikidata.org/entity/>
4 PREFIX wdt: <http://www.wikidata.org/prop/direct/>
5
6 select  ?countryLabel (COUNT(?asteroid) as ?nb)
7 where {
8     ?asteroid wdt:P31 wd:Q3863 ;
9     ?asteroid wdt:P31{nature de l'élément} wd:Q3863{astéroïde} ;
10    ?asteroid wdt:P61 ?discoverer .
11    wdt:P61{découvreur ou inventeur} ?discoverer .
12    ?discoverer wdt:P27 ?country .
13    ?discoverer wdt:P27{pays de nationalité} ?country .
14
15    SERVICE wikibase:label {
16      bd:serviceParam wikibase:language "en" .
17    } .
18
19 GROUP BY ?countryLabel
20 ORDER BY DESC(?nb)
```

Français

Langue d'affichage des labels de Wikidata et de la recherche par autocomplétion

Lignes traduisant les ID de Wikidata avec leurs liens

Press **ctrl + space** to activate auto completion.

Test query

Doc

Upload the results

# Commentaires



*Comments:*

```
# Comments start with a '#'  
# continue to the end of the line
```

En fonction des logiciels des erreurs peuvent se produire.

# Préfixes communs



prefix...	...stands for
rdf:	<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>
rdfs:	<a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#</a>
owl:	<a href="http://www.w3.org/2002/07/owl#">http://www.w3.org/2002/07/owl#</a>
xsd:	<a href="http://www.w3.org/2001/XMLSchema#">http://www.w3.org/2001/XMLSchema#</a>
dc:	<a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/</a>
foaf:	<a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/</a>

 Il vous en manque un ? Allez sur <http://prefix.cc>

OU

```
bd:serviceParam wikibase:language "en" .
```



Prefix bd doesn't exist.

Add : bd

# Une requête SELECT à la loupe

Déclarer les préfixes

(*facultatif*)

Pas d'espace entre

Le nom du préfixe et

Les deux points

**PREFIX** **foo**: <...>

**PREFIX** **bar**: <...>

...

**SELECT** ...

**FROM** <...>

**FROM NAMED** <...>

**WHERE** {

...

}

**GROUP BY** ... (1.1)

**HAVING** ... (1.1)

**ORDER BY** ...

**LIMIT** ...

**OFFSET** ...

**BINDINGS** ... (1.1)

Définir

le jeu de données

(*facultatif*)

Clause  
de définition  
des résultats

Conditions  
qui devront  
être respectées

Modificateurs  
de résultats

(*facultatif*)

# Variables

*Variables:*

```
?var1 ?anotherVar ?and_one_more
```

- 🌀 Ne commence pas par un chiffre
- 🌀 Sensible à la casse
- 🌀 Pas d'espace
- 🌀 Signifiant car il sert de nom de colonne
  - Introduire le mot-clé AS pour changer le nom des colonnes

# Les triplets



```
ex:myWidget ex:partNumber "XY24Z1" .
```

*Match an exact RDF triple:*

```
?person foaf:name "Lee Feigenbaum" .
```

*Match one variable:*

```
conf:SemTech2009 ?property ?value .
```

*Match multiple variables:*

(1 sujet, 2 prédicat, 3 objet)


1. Le sujet représente la ressource à décrire ;
2. Le prédicat représente un type de propriété applicable à cette ressource ;
3. L'objet représente une donnée ou une autre ressource : c'est la valeur de la propriété.



# Atelier 3 :

## ex 1 : Afficher 10 triplets



 Vous allez via le formulaire :  
Sélectionner le endpoint de Wikidata

 Vous allez afficher dans un tableau 10 triplets



# Zoom sur l'essentiel



## Endpoint :

`https://query.wikidata.org/bigdata/namespace/wdq/sparql`

## Query :

```
SELECT *
WHERE {
    ?subject ?property ?literalOrObject .
}
LIMIT 10
```

## Graphique :

`google.vizualisation.Table`

# Atelier 3 :

## ex 2 : chercher les instances d'un concept

- 🌀 Dans la requête précédente vous allez supprimer la variable propriété
- 🌀 Puis appuyer sur CTRL+ESPACE
- 🌀 Sélectionner  
Add a property

```
1 select *
2 where {
3     ?object | ?valueOrObject .
4 }
5 LIMIT 10
```

Press **Ctrl + space** to activate

- Add a property (example for
- Add a item (example item of
- Add the label service of Wi
- Add the langMatches filter

# Atelier 3 :

## ex 2 : chercher les instances d'un concept

- 🌀 ?p: vient d'apparaître
- 🌀 Il vous suffit de commencer à taper :
  - Nature de (français)
  - Instance (anglais)
- 🌀 Puis avec CTRL+SPACE, une liste de propriétés apparaît
- 🌀 sélectionner « Nature de l'élément »

```
?object ?p:natu|?valueOrObject .
```

```
nature de l'élément (nature de, expression d
```

```
abondance naturelle (abondance des isotopes
```

# Atelier 3 :

## ex 2 : chercher les instances d'un concept

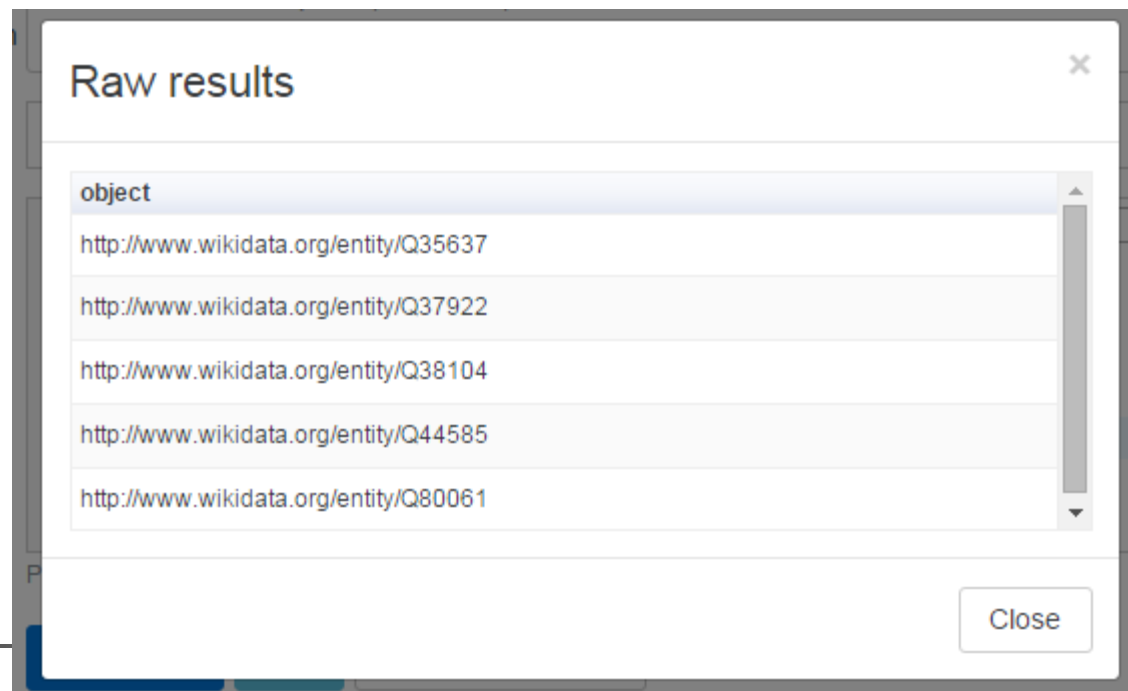
- 🌀 Maintenant, supprimer la variable `literalOrObject`
- 🌀 Puis appuyer sur CTRL+ESPACE
- 🌀 Sélectionner « Add an item »
- 🌀 Commencer à taper un concept dans la langue que vous préférez (en sélectionnant la langue en haut à droite)
- 🌀 Puis appuyer sur le bouton tester la requête et voir les erreurs



# Atelier 3 :

## ex 2 : chercher les instances d'un concept

- 🌀 S'il n'y a pas d'erreurs, appuyer sur tester le graphique pour obtenir un aperçu avant la sauvegarde de votre exemple



# Atelier 3 :

## ex 3 : sélectionner des données

1. Aller sur Wikipedia pour sélectionner un exemple de vos instances
2. Aller ensuite sur sa page Wikidata pour faire la liste des propriétés que vous souhaitez utiliser dans votre requête.
3. Retourner dans le formulaire de la requête
4. Rajouter dans l'entête le texte  
**?subject ?subjectLabel**
5. Dans le corps de la requête ajouter via CTRL+ESPACE : le label service
6. Changer le tag « en » par « fr » si vous souhaitez voir les labels en français ou anglais.

# Zoom sur l'essentiel

```
SELECT ?subject ?subjectLabel
WHERE{
    ?subject wdt:P31 wd:Q7191 .

# ICI
    SERVICE wikibase:label {
        bd:serviceParam wikibase:language "en,fr" .
    }
}
LIMIT 10
```



# Atelier 3 :

## ex 3 : sélectionner des données

- Maintenant, copier le triplet suivant pour chaque variable que vous souhaitez récupérer :

**?subject ?p: ?variable1 .**

- Rechercher via CTRL+ESPACE la propriété que vous aviez trouvé sur la page de Wikidata.

- Puis dans l'entête de la requête écrire :

**?variable1 ou ?variable1Label**

(si c'est un objet)

# Zoom sur l'essentiel

```
SELECT ?subject ?subjectLabel
       ?variable1 ?variable2
WHERE {
    ?subject wdt:P31 wd:Q7191 .
    ?subject wdt:Pxxx ?variable1 .
    ?subject wdt:Pxxx ?variable2 .
    SERVICE wikibase:label {
        bd:serviceParam wikibase:language "en,fr" .
    }
}
LIMIT 10
```

# Graphique

EN CONSTRUCTION...simplification en cours

Choisir son graphique

Lire la documentation de  
son graphique

Tester

Chart

Table

Test chart

Doc & Example

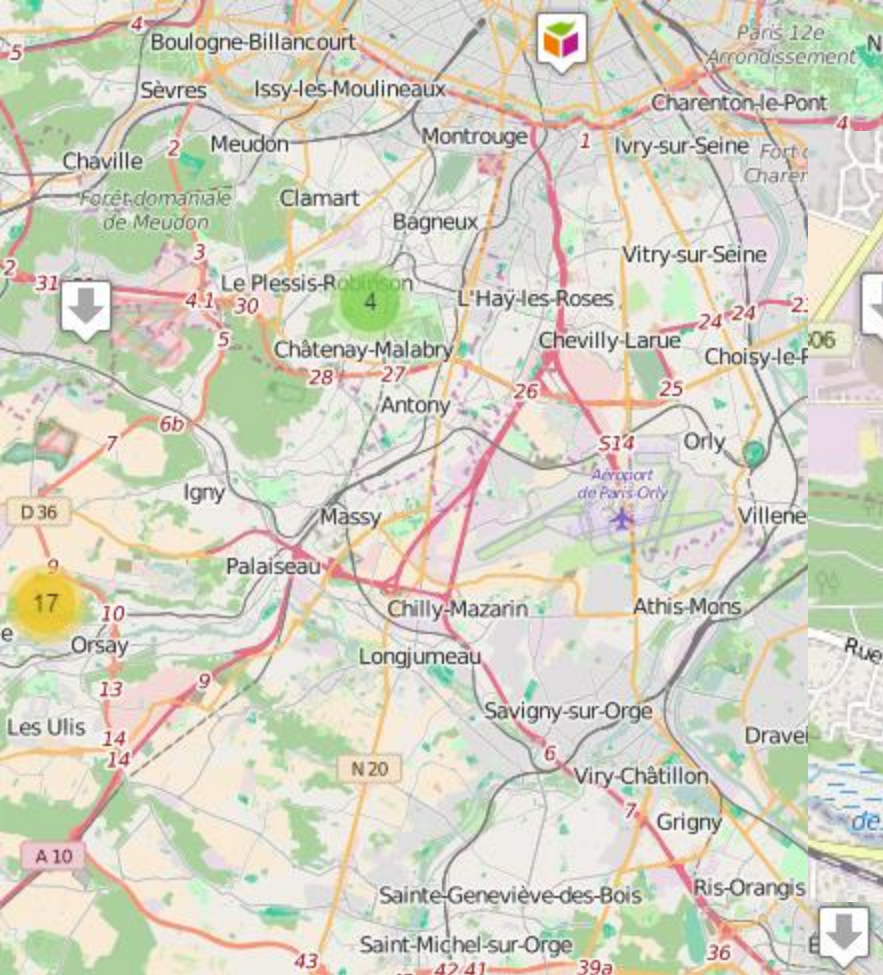
Options

Certains graphiques ont des paramètres

Style

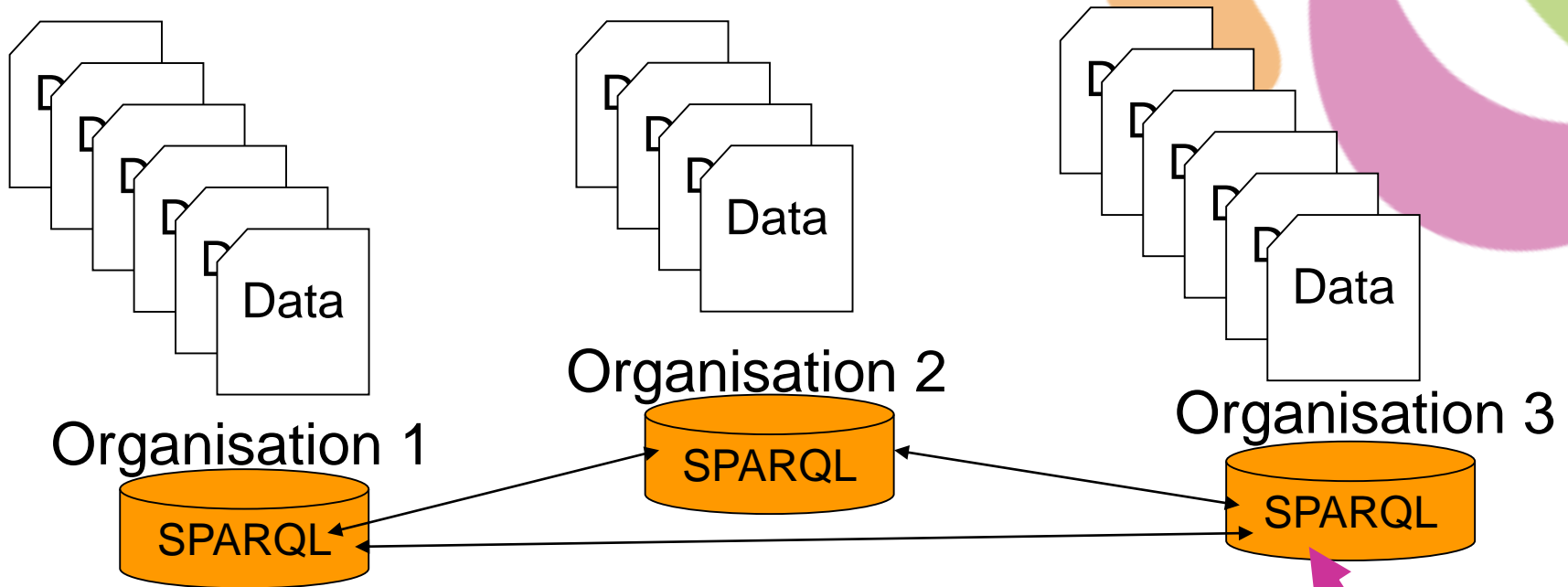
```
width:100%; height:200px;
```

Modifier la taille



5. Wikidata et moi et moi et moi

# Le Linked Data supporte les données décentralisées avec plusieurs structures



La décentralisation des données avec des structures différentes à travers le Web permet aussi de garder le contrôle des données par leurs propriétaires.

Le croisement des données peut se faire du côté du client (respect des données privées) ou des serveurs.



# La jointure des données entre bases de données

- 🌀 Via des requêtes fédérées, on peut croiser les données de Wikidata avec les données d'autres sources.

Par exemple : afficher les petites bibliothèques des services de l'université avec celle décrite dans Wikipedia.

[https://io.datascience-paris-saclay.fr/query/University\\_library\\_in\\_France](https://io.datascience-paris-saclay.fr/query/University_library_in_France)

# Conclusion

- 🌀 Vous n'avez vu que 5% de SPARQL
  - 🌀 Cependant, vous savez chercher un exemple ou un dataset dans le service
  - 🌀 Tester un exemple de requête
  - 🌀 Chercher la documentation SPARQL
  - 🌀 Découvrir l'ontologie de Wikidata
  - 🌀 Sauvegarder une requête
- ➔ Organisation de formations à IODS dans les établissements de Paris Saclay qui le souhaitent.



# BorderCloud

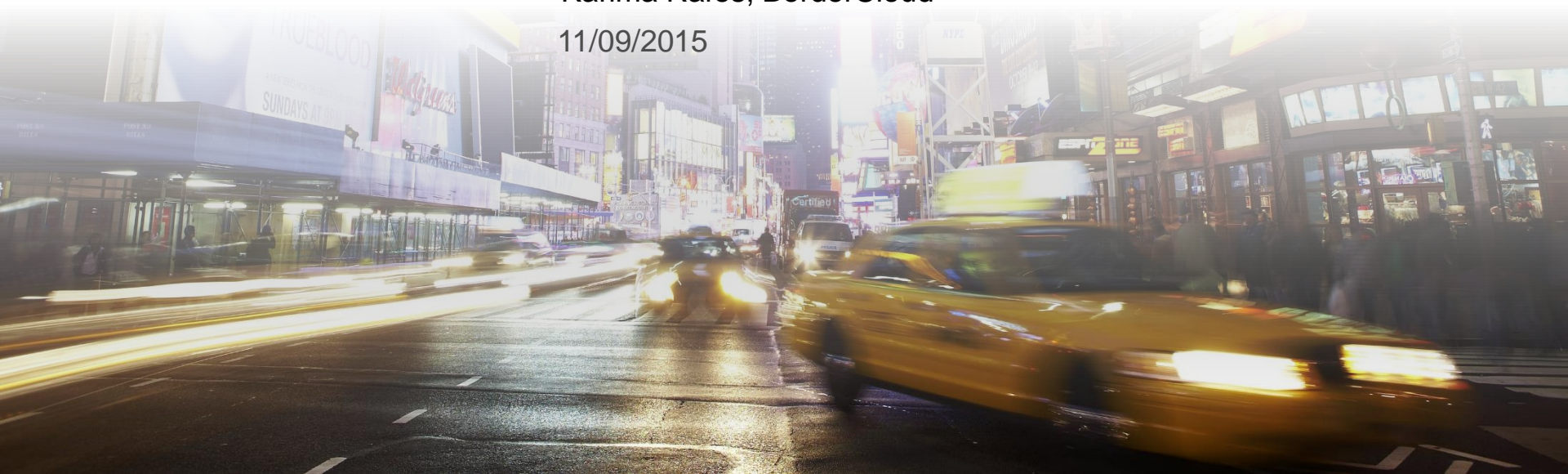
Web 3.0 Hosting & Factory

## Services de formation

### TP OnLine

Karima Rafes, BorderCloud

11/09/2015





# Service pour donner des formations interactives

🌀 Catalogue sur <http://bordercloud.com>

🌀 TPO sur <http://tpo.bordercloud.com>

🌀 Auto-complétion de Wikidata, prefixes, etc.

```
1 PREFIX wikibase: <http://wikiba.se/ontology#>
2 PREFIX wd: <http://www.wikidata.org/entity/>
3 PREFIX wdt: <http://www.wikidata.org/prop/direct/>
4 PREFIX p: <http://www.wikidata.org/prop/>
5 PREFIX q: <http://www.wikidata.org/prop/qualifier/>
6 PREFIX v: <http://www.wikidata.org/prop/statement/>
7
8 SELECT ?countryLabel (count(*) as ?count) WHERE {
9   ?city wdt:P31/wdt:P279* wd:Q515 .
10  ?city p:P6 ?statement .
11  ?statement v:P6 ?mayor .
12  ?mayor wdt:P21 ?female .
13  FILTER NOT EXISTS {
14    ?city wdt:P17 ?country .
15    # If available, get the "ru" label of the country, u
16    SERVICE wikibase:label {
17      bd:serviceParam wikibase:language "en" .
18    }
19  }
20  PREFIX bd doesn't exist. Add : bd
```

```
1 PREFIX wikibase: <http://wikiba.se/ontology#>
2 PREFIX wd: <http://www.wikidata.org/entity/>
3 PREFIX wdt: <http://www.wikidata.org/prop/direct/>
4 PREFIX p: <http://www.wikidata.org/prop/>
5 PREFIX q: <http://www.wikidata.org/prop/qualifier/>
6 PREFIX v: <http://www.wikidata.org/prop/statement/>
7
8 SELECT ?countryLabel (count(*) as ?count) WHERE {
9   ?city wdt:P31/wdt:P279* wd:Q515 .
10  ?city p:P6 ?statement .
11  ?statement v:P6 ?mayor .
12  ?mayor wdt:P21 wd:Q6581072 .
13  FILTER NOT EXISTS { ?statement q:P582 ?x }
14  ?city wdt:P17 ?country
15  # If available, get the "ru" label of the country, u
16  SERVICE wikibase:label {
17    bd:serviceParam wikibase:language "en" .
18  }
19  PREFIX bd doesn't exist. Add : bd
```

# Service pour partager vos découvertes

Demo <http://tpo.bordercloud.com/K.php?i=1abdb0c0e9df986a715e76bb5ea0977f>

List of countries ordered by the number of their cities with female mayor



## Endpoint:

<https://query.wikidata.org/bigdata/namespace/wdq/sparql>

Copy

## Query:

```
PREFIX wikibase: <http://wikiba.se/ontology#>
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX p: <http://www.wikidata.org/prop/>
PREFIX q: <http://www.wikidata.org/prop/qualifier/>
PREFIX v: <http://www.wikidata.org/prop/statement/>

SELECT ?countryLabel (count(*) as ?count) WHERE {
  ?city wdt:P31/instance of wd:Q515 . # find instances of subclasses of city
  ?city p:P6 ?statement . # with a P6 (head of government) statement
  ?statement v:P6 ?mayor . # ... that has the value ?mayor
  ?mayor wdt:P21 wd:Q6581072 . # ... where the ?mayor has P21 (sex or gender) female
  FILTER NOT EXISTS { ?statement q:P582 ?x } # ... but the statement has no P582 (end date) qualifier
  ?city wdt:P17 ?country # Also find the country of the city

  # If available, get the "ru" label of the country, use "en" as fallback:
  SERVICE wikibase:label {
    bd:serviceParam wikibase:language "en" .
  }
} GROUP BY ?country ?countryLabel ORDER BY DESC(?count) LIMIT 100
```

Copy

## Result:





# BorderCloud

**Web 3.0 Hosting & Factory**

---

<http://www.bordercloud.com>

[contact@bordercloud.com](mailto:contact@bordercloud.com)