



Julien Cojan
Inria – équipe Wimmics






- Extraction de données depuis Wikipédia en français
- Membre du comité d'internationalisation (12 chapitres)
<http://dbpedia.org/Internationalization>

Données

Données extraites de Wikipédia



WIKIPÉDIA
L'encyclopédie libre

- Accueil
- Portails thématiques
- Index alphabétique
- Article au hasard
- Contacteur Wikipédia
- Contribuer
- Imprimer / exporter
- Boîte à outils
- Autres langues
 - Afrikaans
 - Alemannisch
 - Aragonés
 - Ænglisc
 - العربية
 - Deutsch
 - Ελληνικά
 - English
 - Esperanto
 - Español
 - Eesti
 - Euskara
 - فارسی
 - Suomi

Article [Discussion](#) Lire [Modifier](#) [Afficher l'historique](#) Rechercher

Musée du Louvre

48° 51′ 40″ N 2° 20′ 09″ E (carte)

Le **musée du Louvre** est l'un des plus grands **musées** du monde et le plus grand **musée de Paris** par sa surface de 210 000 m² dont 60 600 consacrés aux expositions². Situé au cœur de la ville, entre la rive droite de la **Seine** et la **rue de Rivoli**, dans le 1^{er} arrondissement, le bâtiment est un ancien palais royal, le **palais du Louvre**. La **statue équestre de Louis XIV** est le point de départ de l'**axe historique**, mais le palais n'est pas aligné sur cet axe. Le musée possède une longue histoire de conservation artistique en France, depuis les rois **capétiens** jusqu'à nos jours.

Sommaire [\[afficher\]](#)

Variété des œuvres exposées

Musée universaliste, le Louvre couvre une chronologie et une aire géographique larges, depuis l'Antiquité jusqu'à 1848, de l'Europe occidentale jusqu'à l'Iran, via la Grèce, l'Égypte et le Proche-Orient. Il est constitué de huit départements : Antiquités orientales, Antiquités égyptiennes, Antiquités grecques, étrusques et romaines, Arts de l'Islam, Sculptures, Objets d'art, Peintures, Arts Graphiques et présente 35 000 œuvres dans 60 600 m² de salles.

Les œuvres du musée comprennent des sculptures, dessins, céramiques, objets archéologiques et objets d'art entre autres. Parmi les pièces les plus célèbres du musée se trouvent le **Code d'Hammurabi**, la **Vénus de Milo**, **La Joconde** de **Léonard de Vinci**, et **La Liberté guidant le peuple** d'**Eugène Delacroix**. Le Louvre est le musée le plus visité au monde, avec 8,5 millions de visiteurs en 2010.

À Paris, plusieurs musées nationaux sont complémentaires des collections du Louvre*.

Musée du Louvre



Informations géographiques

Pays France

Ville Paris

Adresse 99, rue de Rivoli
75001 Paris

Coordonnées 48° 51′ 40″ Nord
2° 20′ 09″ Est

Informations générales

Date d'inauguration 10 août 1793

Collections Antiquités orientales
Antiquités égyptiennes
Antiquités grecques, étrusques et romaines du Louvre
Arts de l'Islam
Sculptures
Objets d'art
Peintures
Arts graphiques

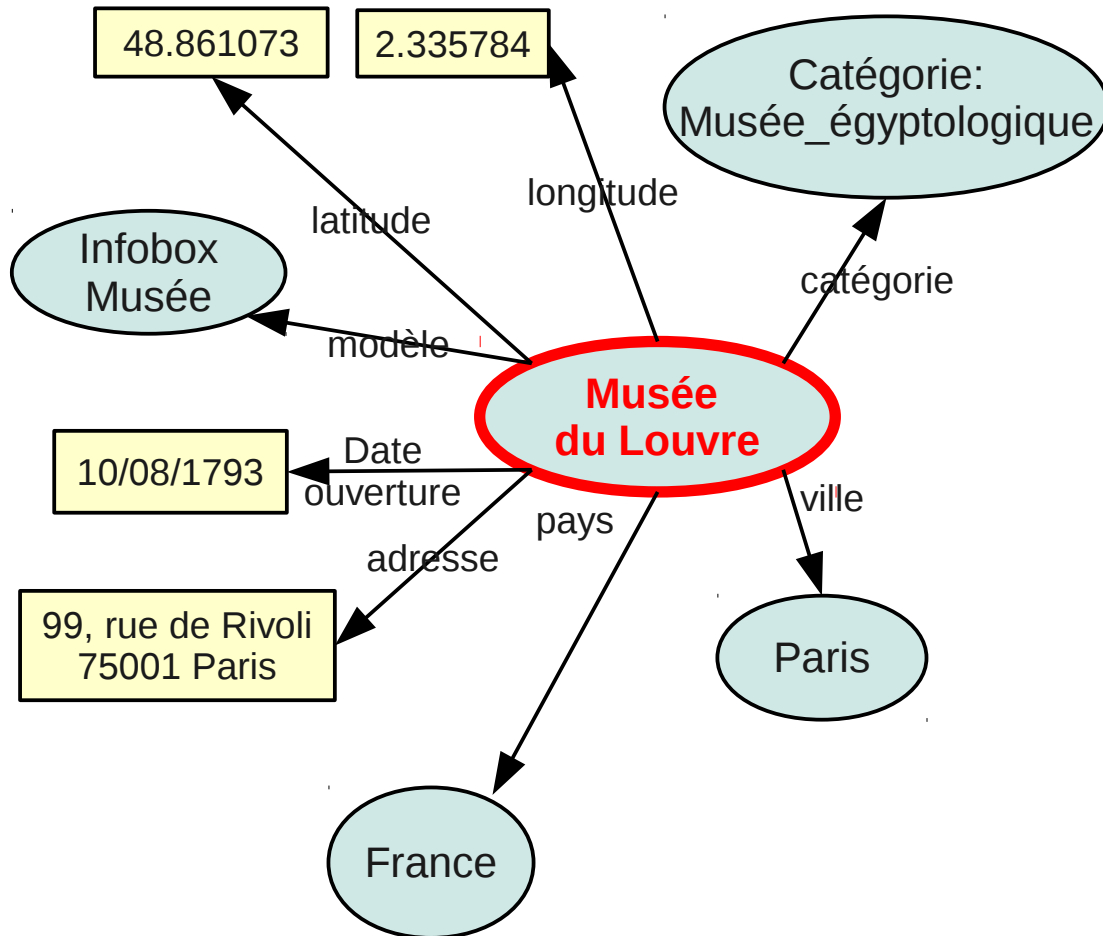
Nombre d'œuvres 35 000 en exposition
445 000 au total (dont 140 000 œuvres graphiques)

Catégories : [Musée national \(France\)](#) | [Musée archéologique en France](#) | [Musée gréco-romain](#) | [Musée égyptologique](#) | [Musée d'art islamique](#) | [Musée fondé en 1793](#) | [Sculpture en France](#) | [Monument parisien](#) | [Patrimoine mondial à Paris](#) | [1er arrondissement de Paris](#) | [Louvre](#) | [\[+\]](#)


categories



Données extraites de Wikipédia



Musée du Louvre



Informations géographiques

Pays	France
Ville	Paris
Adresse	99, rue de Rivoli 75001 Paris
Coordonnées	48° 51′ 40″ Nord 2° 20′ 09″ Est

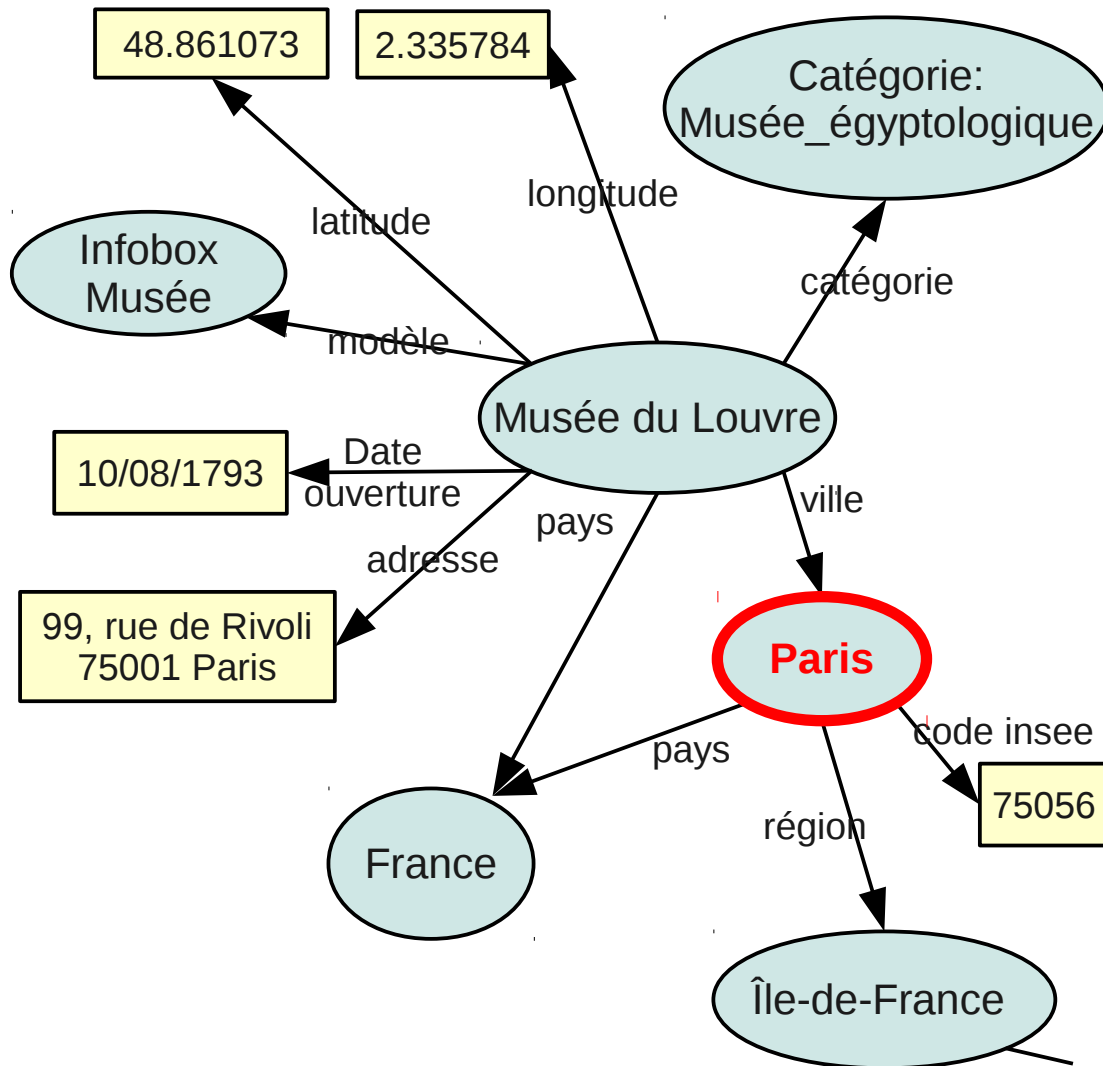
Informations générales

Date d'inauguration	10 août 1793
Collections	<ul style="list-style-type: none"> Antiquités orientales Antiquités égyptiennes Antiquités grecques, étrusques et romaines du Louvre Arts de l'Islam Sculptures Objets d'art Peintures Arts graphiques
Nombre d'œuvres	35 000 en exposition 445 000 au total (dont 140 000 œuvres graphiques)


[Le gréco-romain](#) |
 Musée égyptologique |
 [Patrimoine mondial à Paris](#)



Données extraites de Wikipédia



Paris

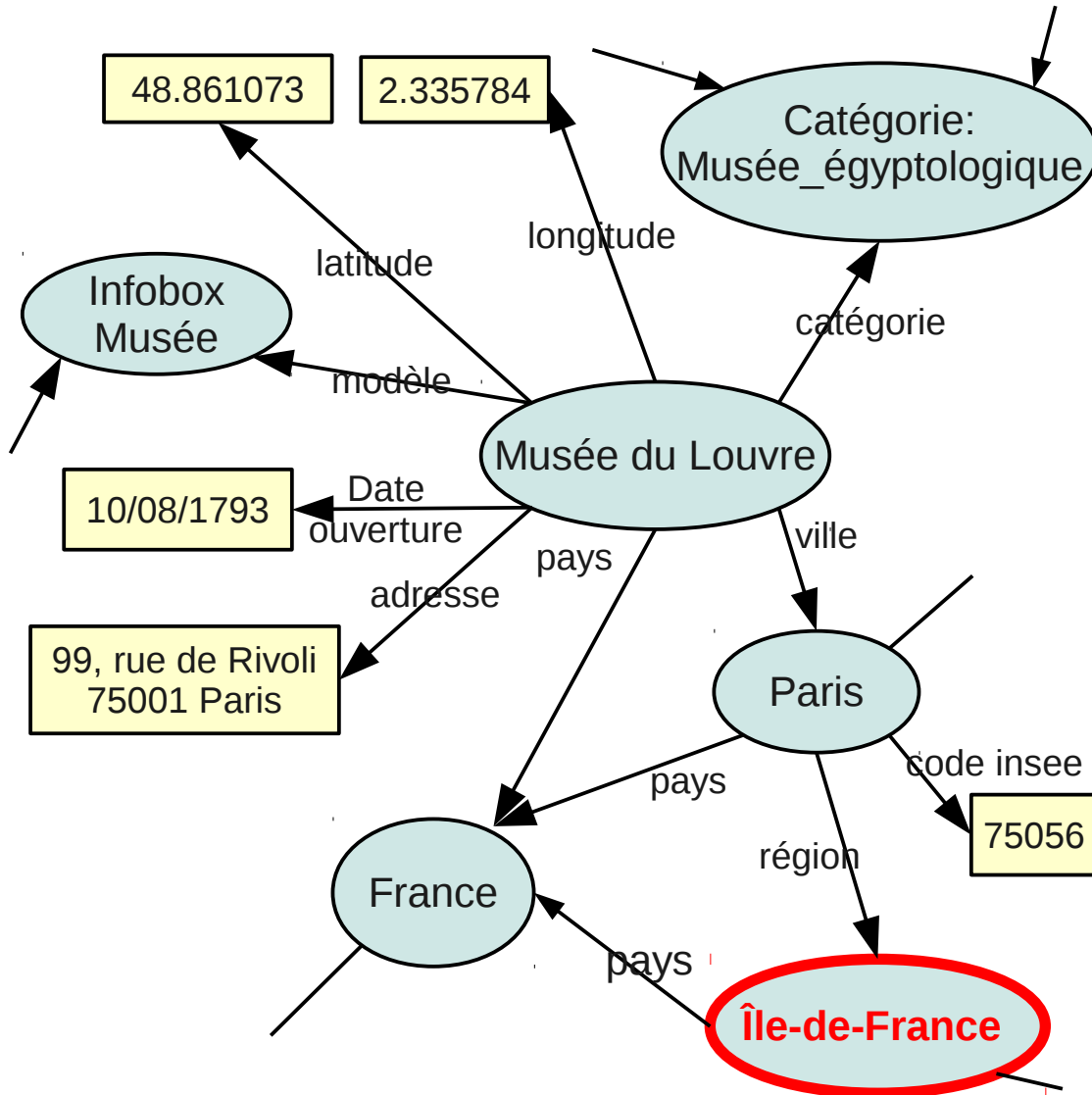




La tour Eiffel (au premier plan) et les gratte-ciel de la Défense (à l'arrière-plan) dominant Paris et sa banlieue ouest.

Administration

Pays	France
Région	Île-de-France (préfecture)
Département	Paris (préfecture)
Arrondissement	Paris (Chef-lieu)
Canton	Chef-lieu de 20 cantons (les arrondissements)
Maire Mandat	Bertrand Delanoë (PS) 2008 - 2014
Code postal	75001 à 75020 et 75116
Code commune	75056

Données extraites de Wikipédia



Île-de-France	
	
<small>Détail</small>	
Administration	
Pays	 France
Préfecture	Paris
Départements	Paris (75) Seine-et-Marne (77) Yvelines (78) Essonne (91) Hauts-de-Seine (92) Seine-Saint-Denis (93) Val-de-Marne (94) Val-d'Oise (95)
Chefs-lieux	Paris Melun Versailles Évry Nanterre Bobigny Créteil Pontoise
Arrondissements	20
Cantons	317
Communes	1 301
Conseil régional	Conseil régional d'Île-de-France
Président	Jean-Paul Huchon (PS) 2010-2016
Préfet	Daniel Canepa

Nommage des ressources

Musée du Louvre

Sujet des articles de DBpedia :

[http://fr.dbpedia.org/resource/\[Nom_Page\]](http://fr.dbpedia.org/resource/[Nom_Page])

dbFr:

Propriétés extraites textuellement :

[http://fr.dbpedia.org/property/\[propriété\]](http://fr.dbpedia.org/property/[propriété])

dbFr-prop:

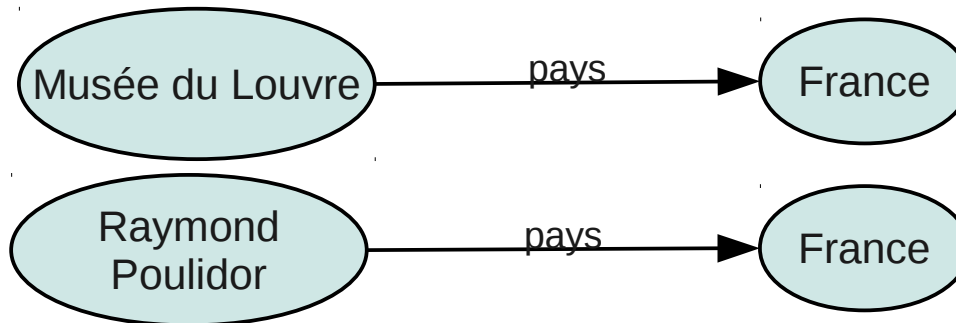
pays

Mais aussi :

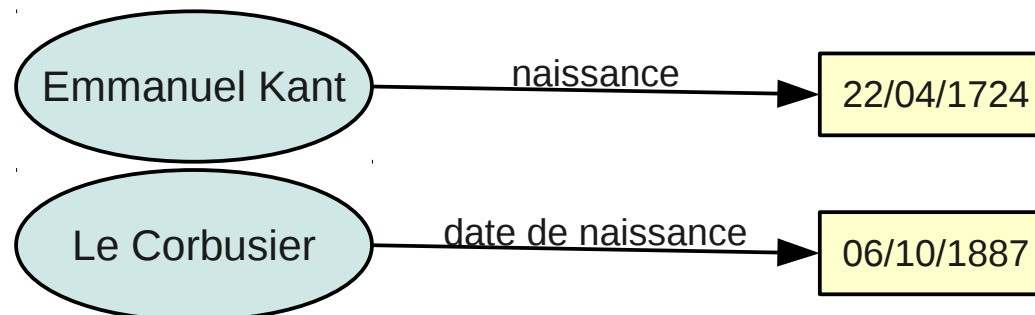
- Vocabulaire normalisé
(Ontologie DBpedia)
- URI stables

Problèmes soulevés par l'extraction textuelle

- Mêmes termes, plusieurs sens



- Différents termes avec le même sens



Vocabulaire normalisé

Vocabulaire défini collaborativement sur
<http://mappings.dbpedia.org>

Ontologie :

- Hiérarchie de classes

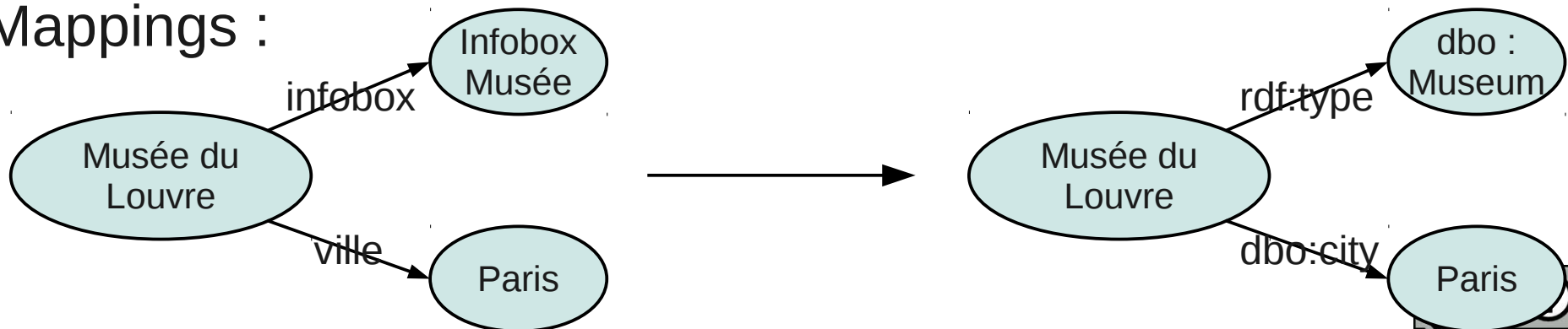
<http://dbpedia.org/ontology/Museum>

- Propriétés

<http://dbpedia.org/ontology/city>

dbo:

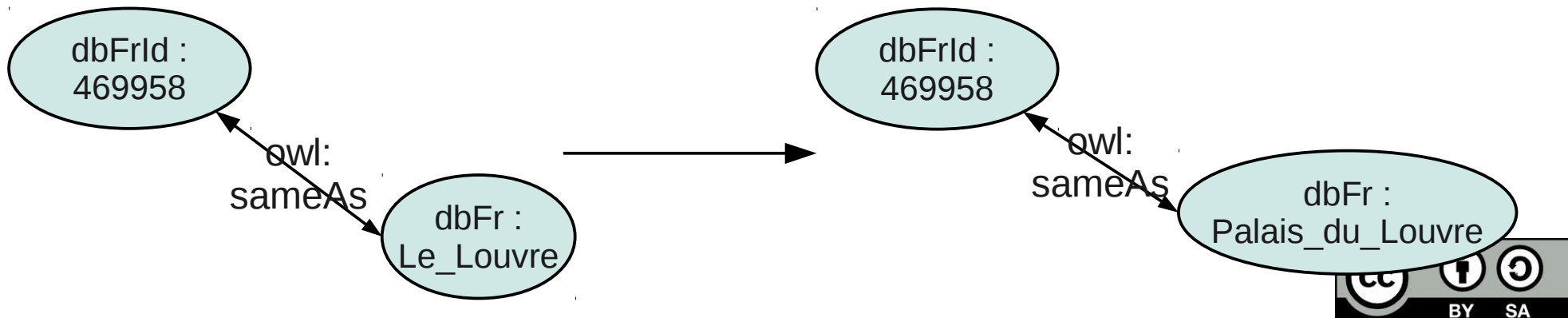
Mappings :



- Lorsqu'une page est renommée
« Le Louvre » → « Palais du Louvre »
~~dbFr :Le_Louvre~~ → dbFr :Palais_du_Louvre

- Introduction d'URI ne indépendante du titre
<http://fr.dbpedia.org/resourceById/469958>

dbFrId:



Volume des données

fichier	nbr de triplets
correspondance avec les pages wikipédia	7574361
id page wikipédia	2863425
titre des pages	2524787
extraction d'infobox textuelle	16579168
extraction d'infobox normalisée	2092711
types des instances	2118305

fichier	nbr de triplets
liens interlangues	1876807
liens entre pages	60849851
redirections	1214339
homonymie	298634
liens vers catégories	3779203
categories	201230
liens externes	2140229
images	1649470

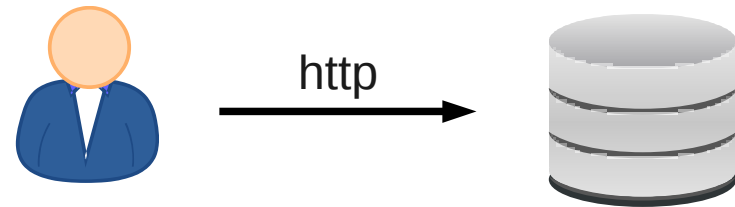
En total 130 millions de triplets.
(soit le second chapitre en taille).

Interroger les données

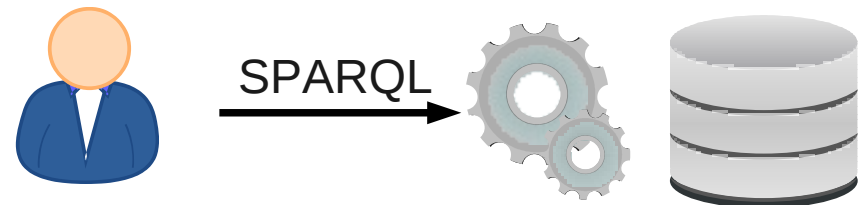
Accès aux données

- Téléchargement de l'ensemble des données
<http://fr.dbpedia.org/download/>
(licence: CC – BY – SA)

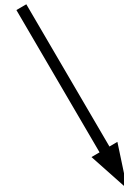
- Par déréférencement



- SPARQL
<http://fr.dbpedia.org/sparql>



http://fr.dbpedia.org/resource/Musée_du_Louvre



About: [Musée du Louvre](#)

An Entity of Type : [architectural structure](#), from Named Graph : <http://fr.dbpedia.org>, within Data Space : fr.dbpedia.org

Property	Value
dbpedia-owl:abstract	<ul style="list-style-type: none"> Le musée du Louvre est l'un des plus grands musées consacrés aux expositions. Situé dans un ancien bâtiment, le palais royal mais le palais n'est pas aligné sur la rue depuis les rois capétiens jusqu'à nos jours.
dbpedia-owl:address	<ul style="list-style-type: none"> 99, rue de Rivoli 75001 Paris
dbpedia-owl:city	<ul style="list-style-type: none"> dbpedia-fr:Paris
dbpedia-owl:collection	<ul style="list-style-type: none"> Arts de l'Islam Peintures Sculptures Arts graphiques Antiquités grecques, étrusques et romaines Objets d'art Antiquités orientales Antiquités égyptiennes
dbpedia-owl:country	<ul style="list-style-type: none"> dbpedia-fr:France

```

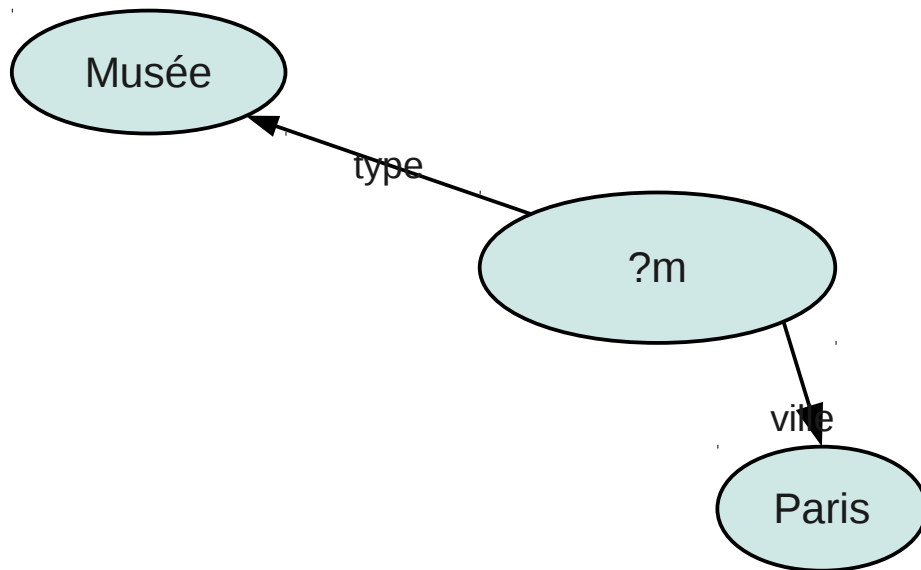
<rdf:RDF>
- <rdf:Description rdf:about="http://fr.dbpedia.org/resource/Musée_du_Louvre"
  <rdf:type rdf:resource="http://schema.org/Place"/>
  <rdf:type rdf:resource="http://dbpedia.org/ontology/Place"/>
  <rdf:type rdf:resource="http://www.opengis.net/gml/Feature"/>
  <rdf:type rdf:resource="http://dbpedia.org/ontology/Building"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Thing"/>
  <rdf:type rdf:resource="http://dbpedia.org/ontology/ArchitecturalStructure"/>
  <rdf:type rdf:resource="http://dbpedia.org/ontology/Museum"/>
  <rdf:type rdf:resource="http://schema.org/Museum"/>
  <rdfs:label xml:lang="fr">Musée du Louvre</rdfs:label>
  <geo:lat rdf:datatype="http://www.w3.org/2001/XMLSchema#float">48.8629</geo:lat>
  <geo:long rdf:datatype="http://www.w3.org/2001/XMLSchema#float">2.3328</geo:long>
  <dbpedia-owl:country rdf:resource="http://fr.dbpedia.org/resource/France"/>
  <dbpedia-owl:city rdf:resource="http://fr.dbpedia.org/resource/Paris"/>
  <dbpedia-owl:address xml:lang="fr">99, rue de Rivoli</dbpedia-owl:address>
  <dbpedia-owl:address xml:lang="fr">75001 Paris</dbpedia-owl:address>
  <dbpedia-owl:openingDate rdf:datatype="http://www.w3.org/2001/XMLSchema#date">1793-10-18</dbpedia-owl:openingDate>
  </rdf:Description>
  
```

SPARQL

- Langage de requête sur des données RDF
- Standard du W3C
- Accessible à <http://fr.dbpedia.org/sparql>

Exemple SPARQL 1

Musées à Paris



```

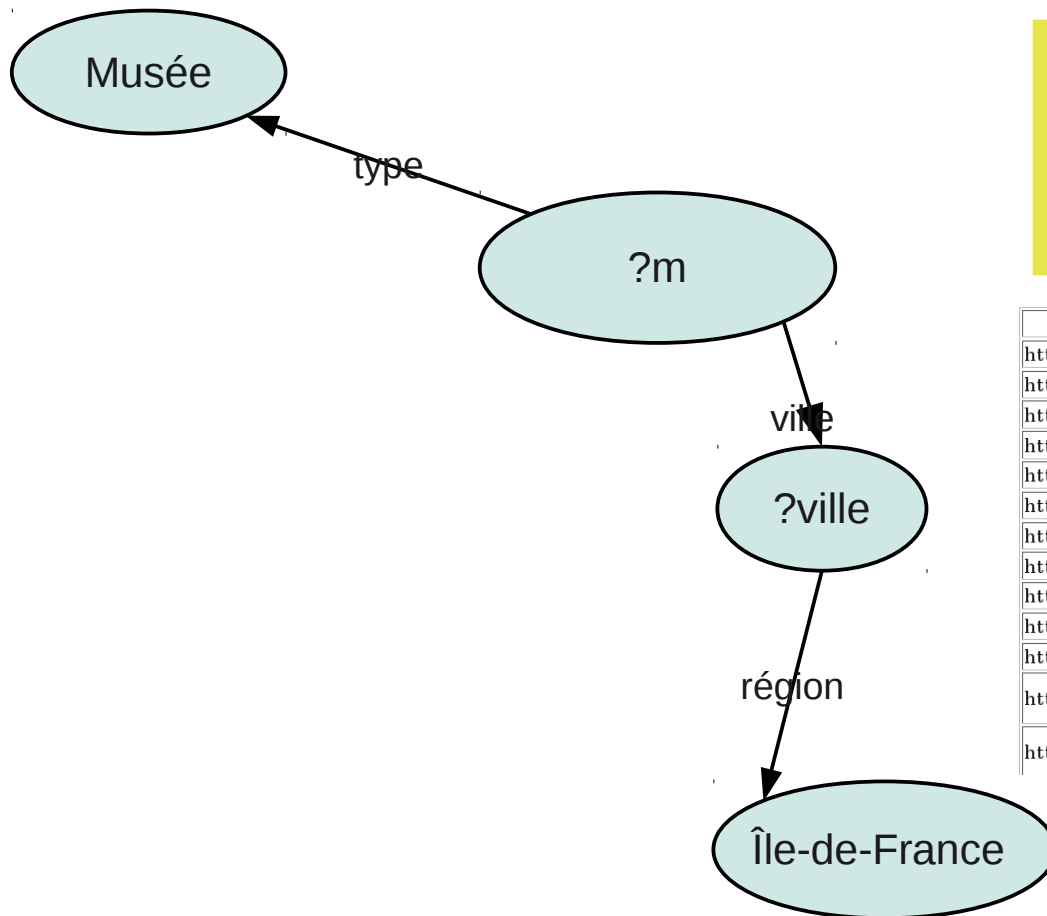
SELECT DISTINCT ?m WHERE {
  ?m rdf:type dbpedia-owl:Museum .
  ?m dbpedia-owl:city dbpedia-fr:Paris .
}
  
```

m
http://fr.dbpedia.org/resource/Catacombes_de_Paris
http://fr.dbpedia.org/resource/Bibliothèque_nationale_de_France
http://fr.dbpedia.org/resource/Musée_de_Cluny
http://fr.dbpedia.org/resource/Musée_Picasso_(Paris)
http://fr.dbpedia.org/resource/Musée_du_Louvre
http://fr.dbpedia.org/resource/Musée_Grévin
http://fr.dbpedia.org/resource/Musée_d'Orsay
http://fr.dbpedia.org/resource/Centre_national_d'art_et_de_culture_Georges-Pompidou
http://fr.dbpedia.org/resource/Musée_Nissim-de-Camondo
http://fr.dbpedia.org/resource/Muséum_national_d'histoire_naturelle
http://fr.dbpedia.org/resource/Palais_de_la_découverte
http://fr.dbpedia.org/resource/Musée_du_quai_Branly
http://fr.dbpedia.org/resource/Bibliothèque_historique_de_la_Ville_de_Paris
http://fr.dbpedia.org/resource/Cité_des_sciences_et_de_l'industrie
http://fr.dbpedia.org/resource/Musée_des_arts_et_métiers
http://fr.dbpedia.org/resource/Musée_national_d'art_moderne
http://fr.dbpedia.org/resource/Musée_Rodin
http://fr.dbpedia.org/resource/Musée_national_Eugène-Delacroix
http://fr.dbpedia.org/resource/Grand_Palais_(Paris)
http://fr.dbpedia.org/resource/Petit_Palais
http://fr.dbpedia.org/resource/Musée_Bourdelle
http://fr.dbpedia.org/resource/Jeu_de_Paume_(musée)



Exemple SPARQL 2

Musées dans une ville d'Île de France



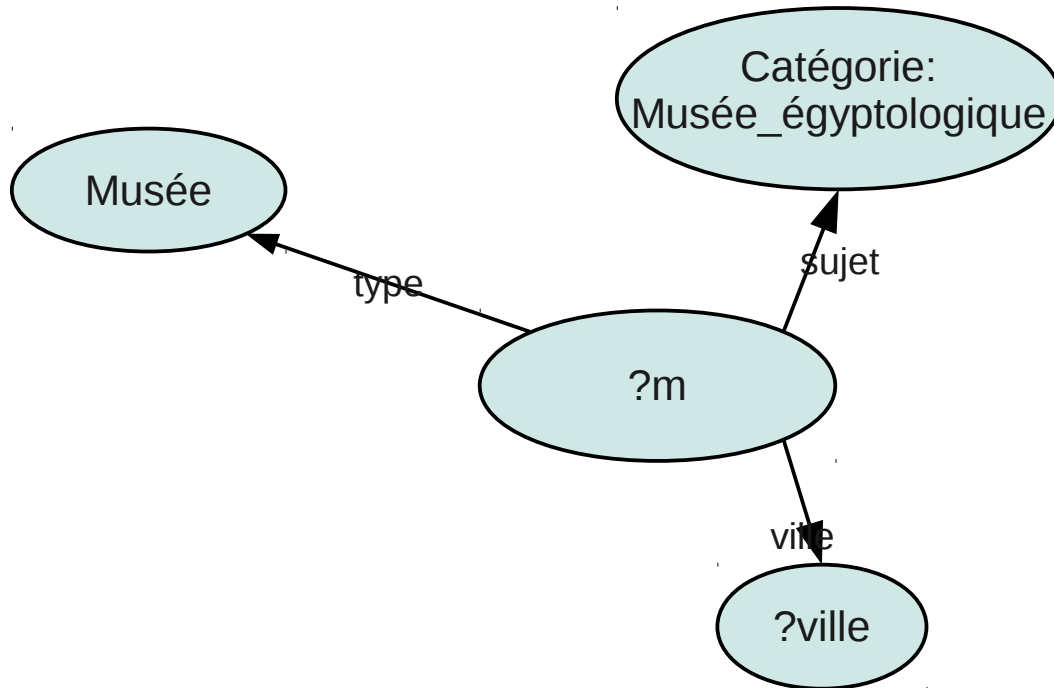
```

SELECT DISTINCT ?m?ville WHERE {
  ?m rdf:type dbpedia-owl:Museum .
  ?m dbpedia-owl:city ?ville .
  ?ville dbpedia-owl:region dbpedia-fr:Île-de-France
}
  
```

m	ville
http://fr.dbpedia.org/resource/Catacombes_de_Paris	http://fr.dbpedia.org/resource/Paris
http://fr.dbpedia.org/resource/Bibliothèque_nationale_de_France	http://fr.dbpedia.org/resource/Paris
http://fr.dbpedia.org/resource/Musée_de_Cluny	http://fr.dbpedia.org/resource/Paris
http://fr.dbpedia.org/resource/Musée_Picasso_(Paris)	http://fr.dbpedia.org/resource/Paris
http://fr.dbpedia.org/resource/Musée_du_Louvre	http://fr.dbpedia.org/resource/Paris
http://fr.dbpedia.org/resource/Musée_Grévin	http://fr.dbpedia.org/resource/Paris
http://fr.dbpedia.org/resource/Musée_d'Orsay	http://fr.dbpedia.org/resource/Paris
http://fr.dbpedia.org/resource/Musée_Nissim-de-Camondo	http://fr.dbpedia.org/resource/Paris
http://fr.dbpedia.org/resource/Muséum_national_d'histoire_naturelle	http://fr.dbpedia.org/resource/Paris
http://fr.dbpedia.org/resource/Palais_de_la_découverte	http://fr.dbpedia.org/resource/Paris
http://fr.dbpedia.org/resource/Musée_du_quai_Branly	http://fr.dbpedia.org/resource/Paris
http://fr.dbpedia.org/resource/Musée_d'archéologie_nationale	http://fr.dbpedia.org/resource/Saint-Germain-en-Laye
http://fr.dbpedia.org/resource/Rambolitrain	http://fr.dbpedia.org/resource/Rambouillet

Exemple SPARQL 3

Dans quelles villes y a-t-il des musées égyptologiques ?
Combien ?



```

select ?ville count(distinct ?m) as ?nb where {
  ?m rdf:type dbpedia-owl:Museum .
  ?m dbpedia-owl:city ?ville .
  ?m dcterms:subject <http://fr.dbpedia.org/resource/Catégorie:Musée_égyptologique>
} group by ?ville order by desc(?nb)
  
```

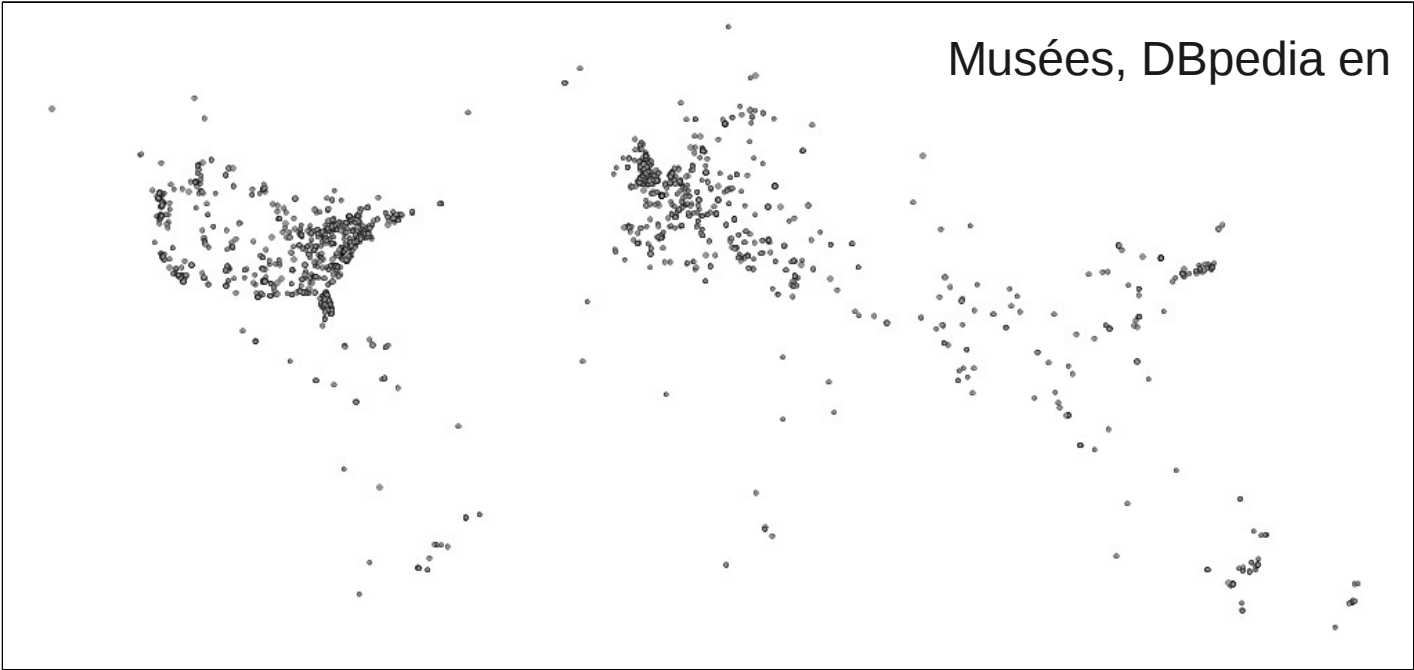
ville	nb
http://fr.dbpedia.org/resource/Le_Caire	2
http://fr.dbpedia.org/resource/Paris	1
http://fr.dbpedia.org/resource/Moscou	1
http://fr.dbpedia.org/resource/New_York	1
http://fr.dbpedia.org/resource/Brooklyn	1
http://fr.dbpedia.org/resource/Alexandrie	1
http://fr.dbpedia.org/resource/Nantes	1
http://fr.dbpedia.org/resource/Louxor	1
http://fr.dbpedia.org/resource/Naples	1
http://fr.dbpedia.org/resource/Turin	1

À venir

- Outils
 - Lookup en français
Recherche de termes par mot clés
 - Spotlight en français
Reconnaissance d'entités nommées dans un texte
- Historisation des données

Merci

Musées, DBpedia en



Musées, DBpedia fr

