

SPARNATURAL

Un explorateur de graphes de connaissances intuitif et configurable

thomas.francart@sparna.fr / florence.clavaud@culture.gouv.fr

SemWeb.Pro 2022

Combien de participants à semweb.pro savent
écrire une requête SPARQL ?



Combien de participants à semweb.pro savent
requêter data.bnf.fr en SPARQL
pour trouver les éditions des
œuvres de Victor Hugo en anglais ?



Pour tous les autres « vrais gens »

<http://data.bnf.fr/sparnatural>

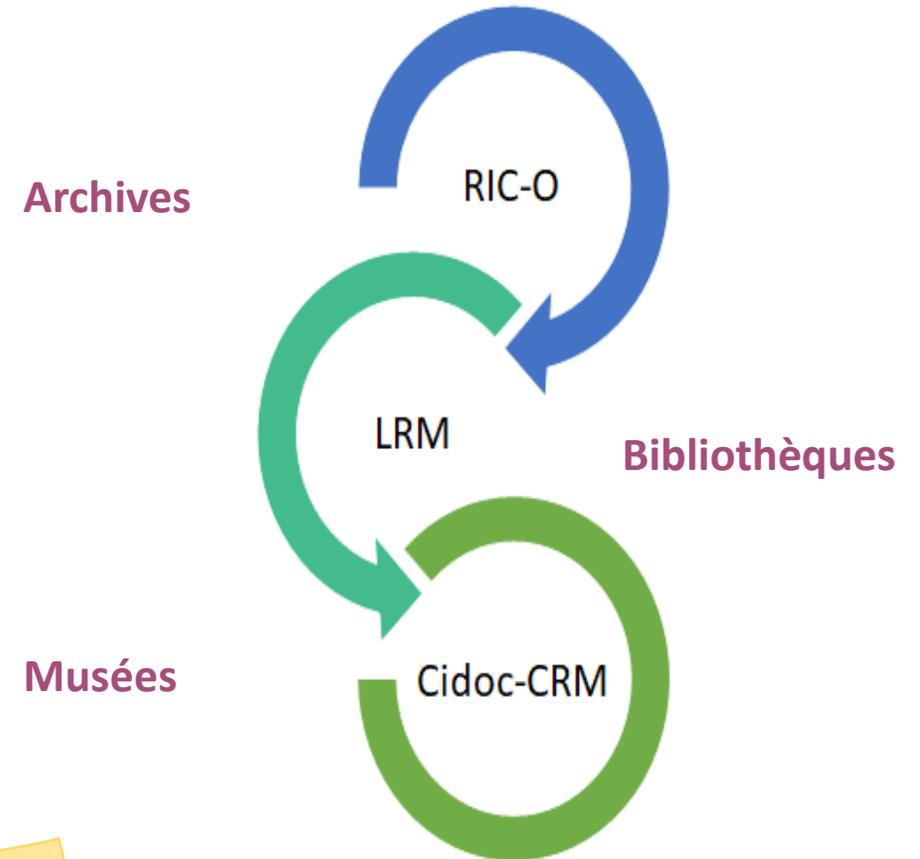
Interrogez data.bnf.fr sans connaître ni SPARQL ni le modèle de données

Demo : <https://youtu.be/-yEQdvtBvCg>

Le projet : comment donner à
voir les graphes de
connaissances de deux
institutions patrimoniales
nationales : AN & BNF ?

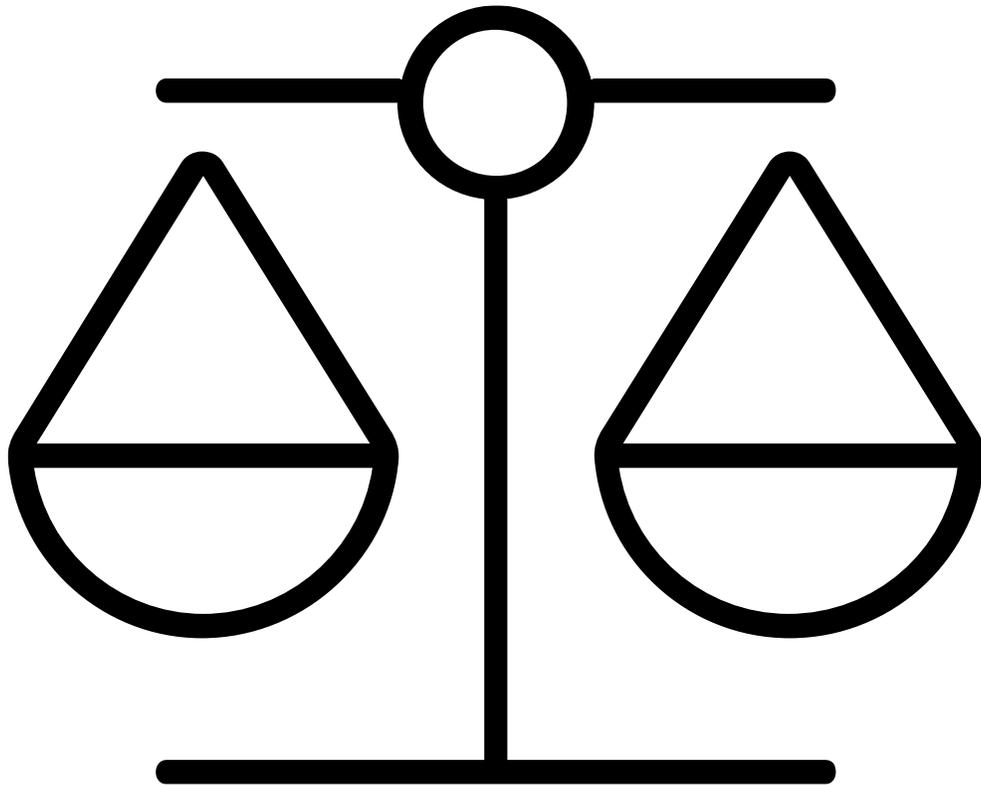
Contexte du projet

- Stratégie nationale pour la mise en œuvre des principes du Linked Open Data pour les **métadonnées culturelles...**
- ... qui sont décrites sur la base de **modèles métiers** qui se structurent en graphes de connaissances...
- ... qui nécessitent de nouvelles interfaces pour les **utilisateurs finaux**

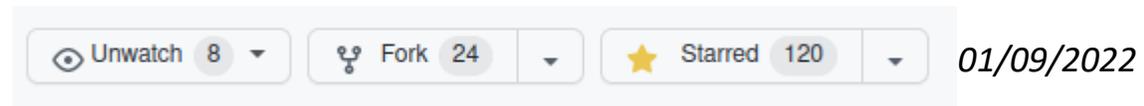


des vrais gens !

Innovation ouverte



- Sparnatural : explorateur de graphes de connaissances reconfigurable
- License **LGPL 3.0** : Utilisation, y compris commerciale
 - <https://github.com/sparna-git/Sparnatural>



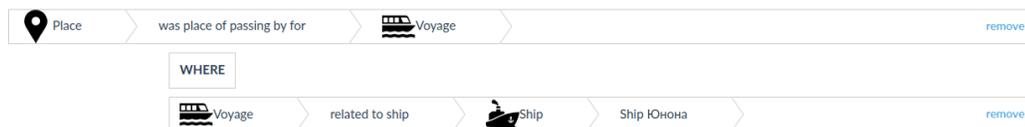
- But : créer un commun et une communauté
 - Donc : doc ! <https://docs.sparnatural.eu>

Historique de Sparnatural

- Inspiration : ResearchSpace / British Museum
 - <http://researchspace.org>
 - On monte sur les épaules des géants
- Exemple:
<http://rs.sealitproject.eu/resource/sealit:Search>

- Développement de Sparnatural dans le cadre du projet **OpenArcheo** du consortium MASA en 2018
 - <http://openarchaeo.huma-num.fr>
 - Présenté à semweb.pro 2019

Find: Places WAS PLACE OF PASSING BY FOR Voyage where voyages related to ship Ship Юнона



Le projet AN/BNF

1. Evolutions fonctionnelles de Sparnatural (discutées en amont)

1. *Sélection des colonnes de résultat*
2. *Support du OPTIONAL / NOT EXISTS*
3. *Sélection dans des arbres*
4. *Tooltips*
5. *Chargement de requêtes sauvegardées*
6. ...

2. Réaliser et déployer 2 **démonstrateurs** AN / BNF

1. Démonstrateur AN : <https://sparna-git.github.io/sparnatural-demonstrateur-an>

3. Faire des ateliers avec des **utilisateurs finaux**

4. **Ecrire de la doc**, doc, doc : <https://docs.sparnatural.eu>
pour faciliter les déploiements dans d'autres institutions



des vrais gens !

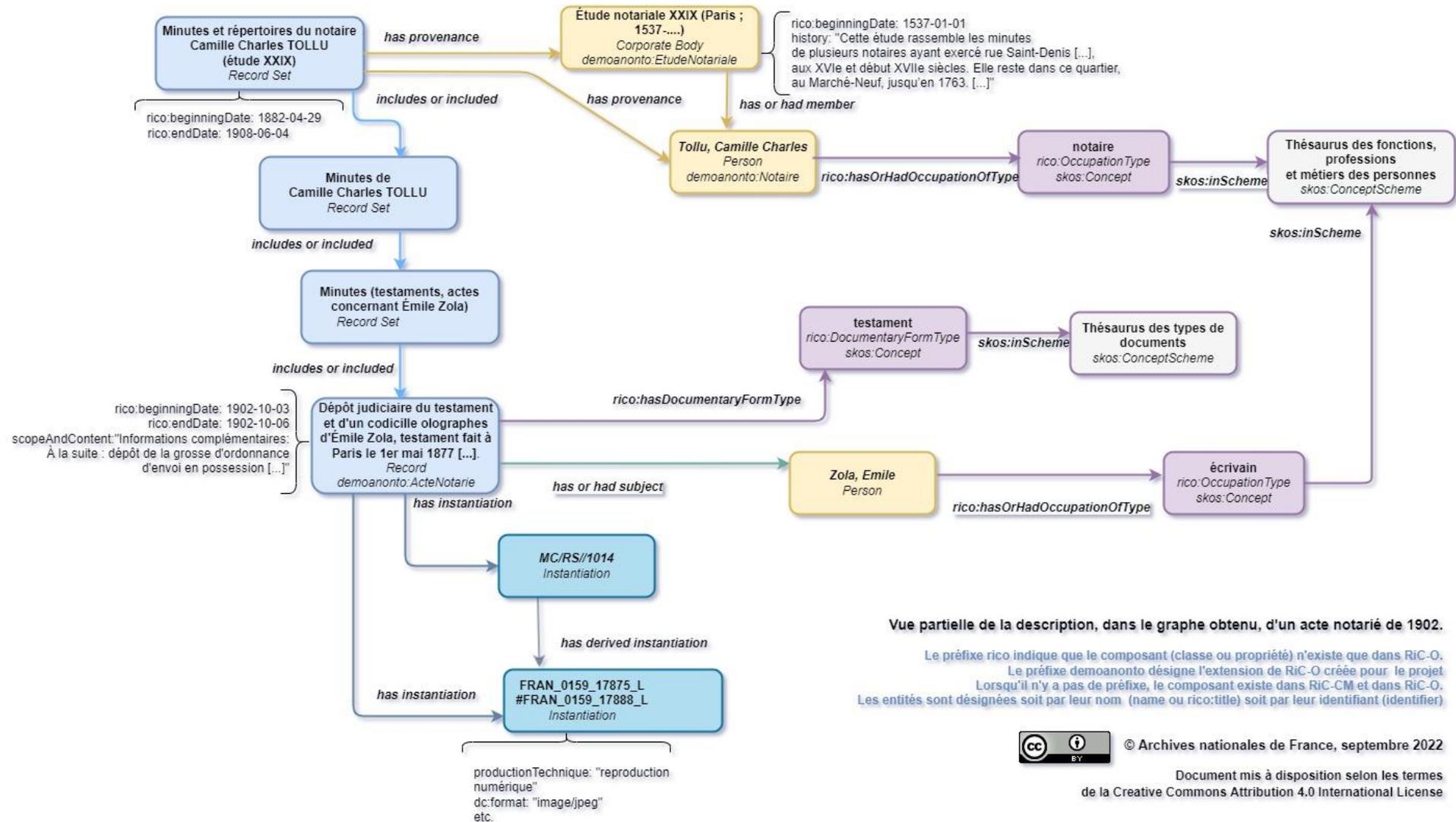
Les données des AN

- Le jeu de données des AN a été constitué pour l'occasion
- Partie des métadonnées décrivant les archives des **notaires de Paris du XVe siècle à nos jours (soit environ 9 km linéaires d'archives papier)**
 - *Cohérentes / riches / très consultées / des archives de même nature existent dans les services d'archives départementales*
 - 40 études notariales (sur 122), 1079 notaires, 1577 instruments de recherche
 - + les vocabulaires contrôlés

Les données des AN

- Conversion de fichiers EAD et EAC-CPF
 - Utilisation pour l'essentiel du logiciel open source [RiC-O Converter](https://github.com/ArchivesNationalesFR/rico-convert), développé pour les Archives nationales en 2019, disponible sur GitHub depuis avril 2020 :
<https://github.com/ArchivesNationalesFR/rico-convert>
- Intégration dans le graphe de certains référentiels des Archives nationales (eux-mêmes convertis en SKOS et RiC-O ; voir <https://github.com/ArchivesNationalesFR/Referentiels/tree/main/concepts/rdf>)
- Données par la suite enrichies via des requêtes SPARQL de type INSERT
 - Environ 58 millions de triplets en tout actuellement
 - Première utilisation de RiC-O à grande échelle en France
- Tous les détails sur la documentation du démonstrateur :
 - <https://sparna-git.github.io/sparnatural-demonstrateur-an/presentation-fr.html>
- Les données sont disponibles sur demande (florence.clavaud@culture.gouv.fr)

Structure RiC-O



<https://sparnatural-demonstrateur-an/presentation-fr.html>

Vue partielle de la description, dans le graphe obtenu, d'un acte notarié de 1902.

Le préfixe rico indique que le composant (classe ou propriété) n'existe que dans RiC-O.
 Le préfixe demoanonto désigne l'extension de RiC-O créée pour le projet.
 Lorsqu'il n'y a pas de préfixe, le composant existe dans RiC-CM et dans RiC-O.
 Les entités sont désignées soit par leur nom (name ou rico:title) soit par leur identifiant (identifier)

Bilan : dispositif d'exploration et de recherche

- « exploration du graphe »
- « **nouvelles possibilités** de recherche »
- « une fois passé l'effet de surprise, on peut mentionner le caractère à la fois **intuitif, souple et interactif** de ce type d'interface »
- « **plus grand engagement intellectuel** pour l'utilisateur que de remplir un classique formulaire de recherche »
- « l'utilisateur découvre le contenu du graphe en même temps qu'il construit cette recherche ; il prend véritablement la main »
- « il peut **partager la requête exécutée** ou encore exporter les résultats de sa recherche »

Il y aurait beaucoup à dire sur

1. les fonctionnalités d'exploration offertes par Sparnatural
2. l'UX et la « ludification » de l'expérience de recherche

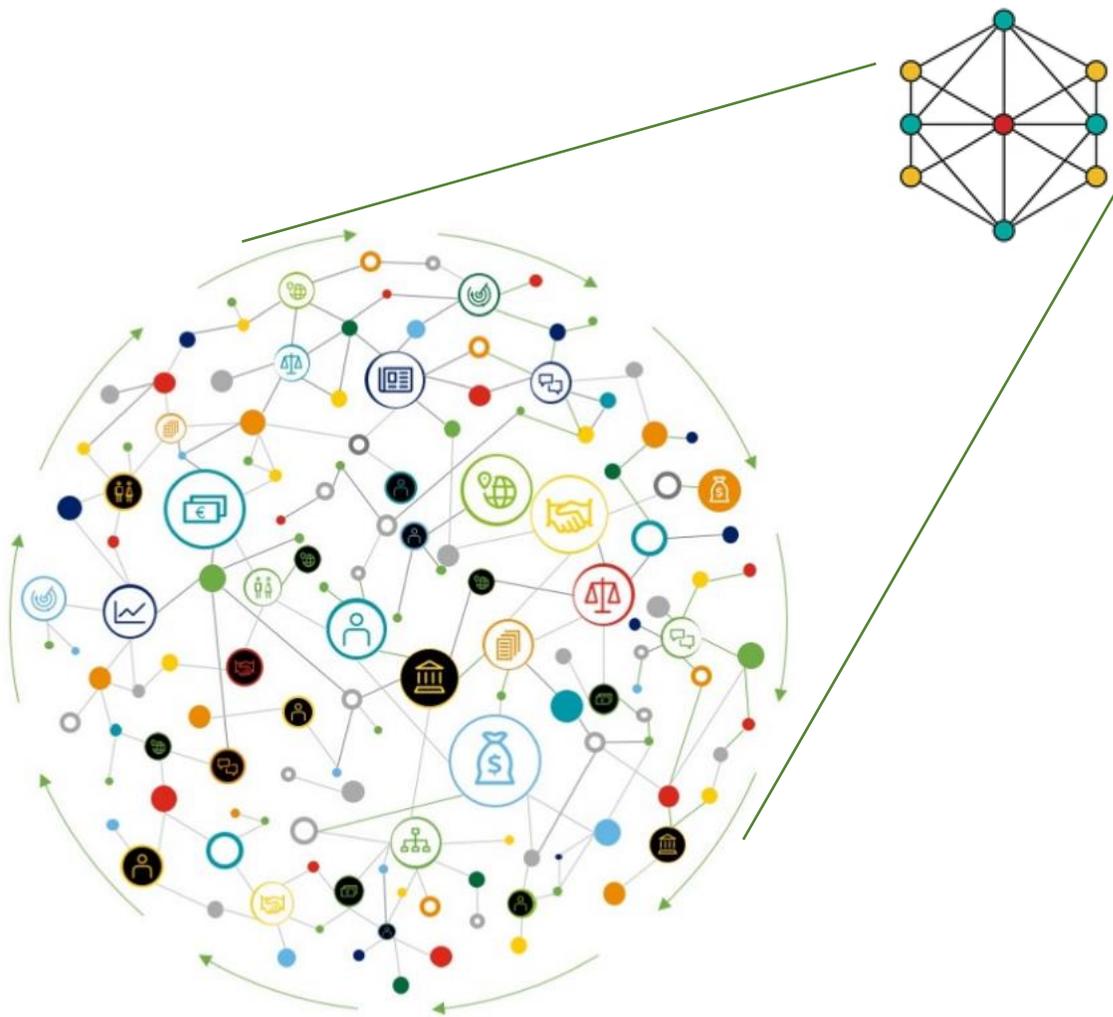
Interactions « essai/erreur », « retour arrière », requêtage par l'exemple, etc.

100 % côté client



- Composant javascript **côté client**, dans le navigateur
- Aucun déploiement de serveur supplémentaire
- Seul prérequis: un service SPARQL
 - (*« CORS-enabled »*)
- Un déploiement facilité
 - *« Votre démo en 15 minutes ! »*

Configuration : Comment
réconcilier des modèles
conceptuels abstraits avec
les points de vue des
utilisateurs finaux ?



Montrer un graphe de connaissances nécessite une étape d'**éditorialisation**

- Choisir des icônes, des libellés, des « tooltips »
- Ne pas tout montrer
- Prendre des raccourcis dans le graphe
- Montrer les liens « à l'envers »
- Généraliser ou au contraire préciser
- ...

les vrais sens !

Modèle conceptuel d'interopérabilité : FRBR, CIDOC-CRM, RiC-O, etc.

- Niveau d'abstraction élevé
- Vocabulaire conceptuel
- Logique d'interopérabilité métier



Utilisateurs finaux

- Niveau d'abstraction "usuel dans la communauté"
- Vocabulaire précis
- Logique de recherche

Ontologie de configuration en OWL

On définit donc *un modèle ontologique pour la recherche et ses correspondances, simples ou plus élaborées, avec les classes et propriétés de l'ontologie métier.*

Voir la documentation à : <https://docs.sparnatural.eu/OWL-based-configuration>

Edition de la configuration en OWL

Exemple de configuration très simple pour une classe.

Fichier OWL, donc : 1/ éditable dans Protégé et 2/ partageable pour d'autres projets utilisant la même ontologie

On envisage de proposer une ontologie générique de navigation basée sur RiC-O.

The screenshot displays the Protégé interface for editing an ontology. The left pane shows a class hierarchy with 'Archives' selected under 'SpamaturalClass'. The right pane shows the configuration for the 'Archives' class, including annotations and a description.

Annotations: Archives

- `rdfs:label` [language: fr] Archives
- `rdfs:label` [language: en] Records
- `defaultLabelProperty` `aPourIntitule`
- `falcon` `fad fa-archive`
- `order` [type: xsd:integer] 1
- `sparqlString` `<https://spamatural-demo-anf.huma-num.fr/ontology#Archive>`
- `tooltip` [language: fr] Tout type de ressource archivistique, qu'il s'agisse d'un item (comme une minute ou un répertoire) ou d'un groupe de documents (série des minutes ou des répertoires d'un notaire, liasse de minutes, dossier de client...).
- `tooltip` [language: en] Any type of archival resource, whether it is an item (such as a minute or a registry) or a group of documents (series of minutes or registry of a notary, bundle of minutes, client file...).

Description: Archives

- Equivalent To** +
- SubClass Of** +
 - `SpamaturalClass`

Le principe du mapping

La génération de la requête finale se fait en 2 étapes :

1. D'abord une requête SPARQL est générée en utilisant seulement les identifiants de classes et de propriétés connus dans la configuration Sparnatural

2. Les identifiants de classes et de propriétés dans la requête sont **remplacées** par des identifiants de classes/propriétés du modèle de graphe **ou des expressions d'identifiant plus complexes**

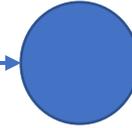
Créer un raccourci

Musée



situé dans

Pays



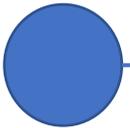
Config :

« Situé dans » = chemin « localisé dans » puis « fait partie de »
sparqlString = *<http://.../localiseDans>/<http://.../faitPartieDe>*

Ce qu'on veut montrer

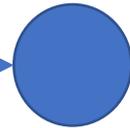
Le graphe

Musée



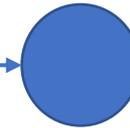
localisé dans

Ville



fait partie de

Pays

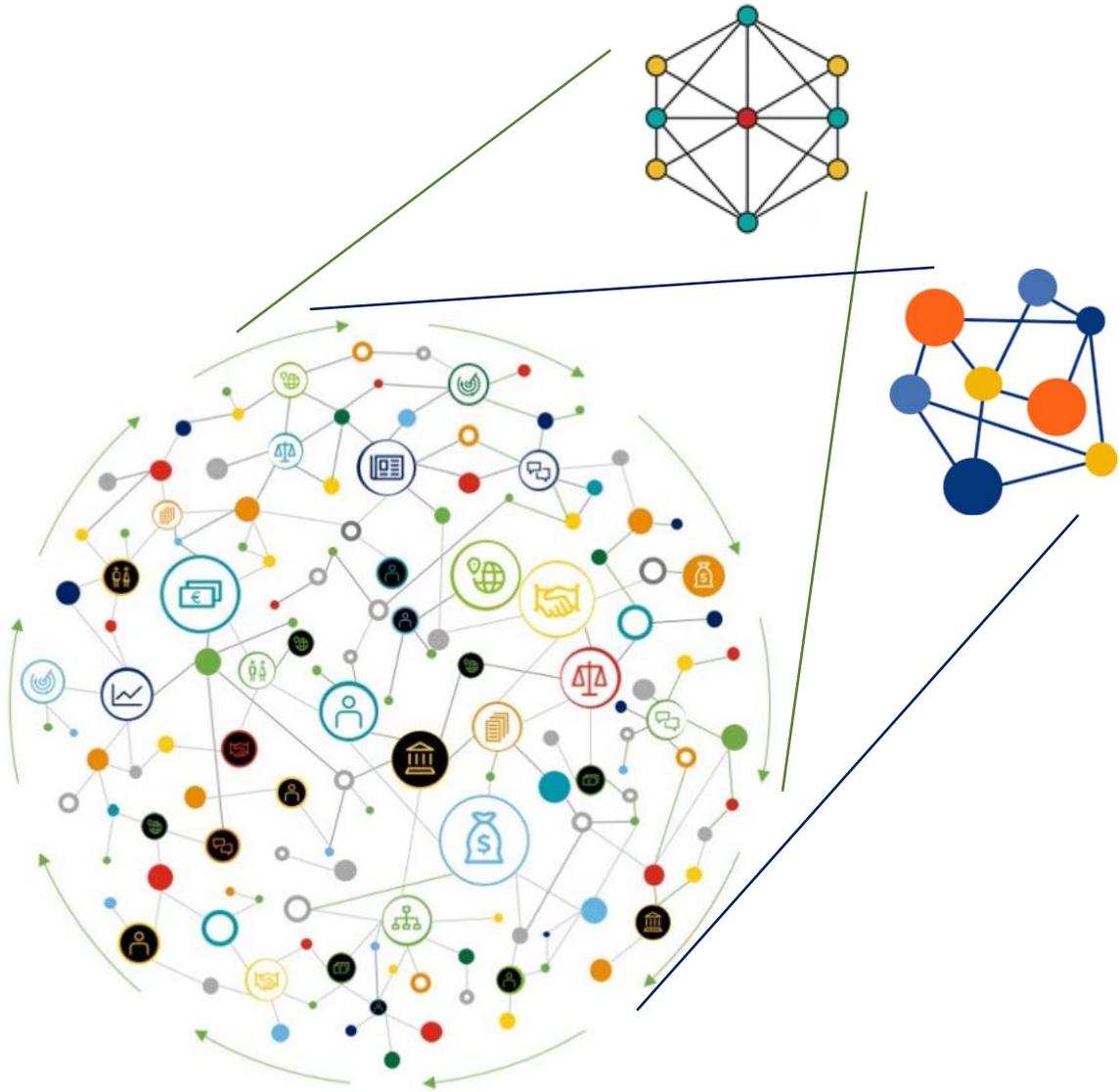


```
1 SELECT ?this
2 WHERE {
3     ?this a <http://exemple.fr/sparnatural-config/Musee> .
4     ?this <http://exemple.fr/sparnatural-config/situeDans> ?Pays .
5 }
```

La requête initiale

La requête transformée (« backward-chaining »)

```
1 SELECT ?this
2 WHERE {
3     ?this a <http://data.moi.fr/ontologie/Musee> .
4     ?this <http://data.moi.fr/ontologie/localiseDans>/<http://data.moi.fr/ontologie/faitPartieDe> ?Pays .
5 }
```



Plusieurs points de
vues peuvent
coexister sur les
mêmes données

Le démonstrateur

<https://youtu.be/mdhIEKSkgs>

<https://sparna-git.github.io/sparnatural-demonstrateur-an/>

Acte notarié *concernent* Édifice parisien Louvre (palais du) +

Une copie numérique fidèle et intégrale d'une ressource archivistique, produite à des fins de communication aux publics comme substitut de l'original.

et

Acte notarié Optionnel Négatif *ont pour copie numérique* Reproduction numérique Tous-tes

et

Acte notarié *a ou ont pour cote* Cote Tous-tes

et

Acte notarié *a ou ont pour provenance* Notaire

où

Notaire *a pour date de début d'exercice* Date A partir de 1750



this Cote_3 Notaire_4

Tri: A Z Z A

Montrer le nom des colonnes



D'autres démos

<https://sparnatural.eu> :

- *Partitions musicales (Cité de la Musique)*
- *Référentiel d'interopérabilité du médicament (maquette) (ANS)*
- *DBPedia*
- *La vôtre ?*

Un outil de valorisation des
données des institutions
s'inscrivant dans leurs missions
de service public.

Pour les vrais gens !

Merci de votre attention