

TP 1 – Web de données

1. DBpedia

DBpedia est un projet collaboratif et une ressource de données ouverte qui extrait des informations semi-structurées à partir de Wikipédia et les rend disponibles au public sous une forme interrogeable. Il transforme les articles de Wikipédia en données RDF (Resource Description Framework), créant ainsi un ensemble de données liées. DBpedia permet de consulter des informations sur des millions d'entités, telles que des personnes, des lieux, des œuvres d'art, des concepts, etc., qui sont décrites dans Wikipédia.

1. Ouvrez le lien suivant : <https://www.dbpedia.org/>
2. Ouvrez le lien suivant : <https://dbpedia.org/resource/Oran> , que remarquez-vous ?
3. Que se passe-t-il en ouvrant le lien suivant : <https://dbpedia.org/data/Oran.xml>
4. Ouvrez le lien suivant : <https://dbpedia.org/page/Oran> , comment les données sont-elles organisées ? Extrayez 2 triplets, dont l'un avec une valeur littérale, et un autre dont la valeur est une ressource.
5. Dans la même page, observez les différents formats existants.
6. En se basant sur Wikipedia, citez trois URIs différentes.
7. Ouvrez le site <https://issemantic.net/rdf-visualizer> .
 - Copiez le code RDF/XML suivant, soumettez-le dans l'onglet « VISULIZE CODE SNIPPET » et observez le résultat.

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dbo="http://dbpedia.org/ontology/"
  xmlns:dbr="http://dbpedia.org/resource/">

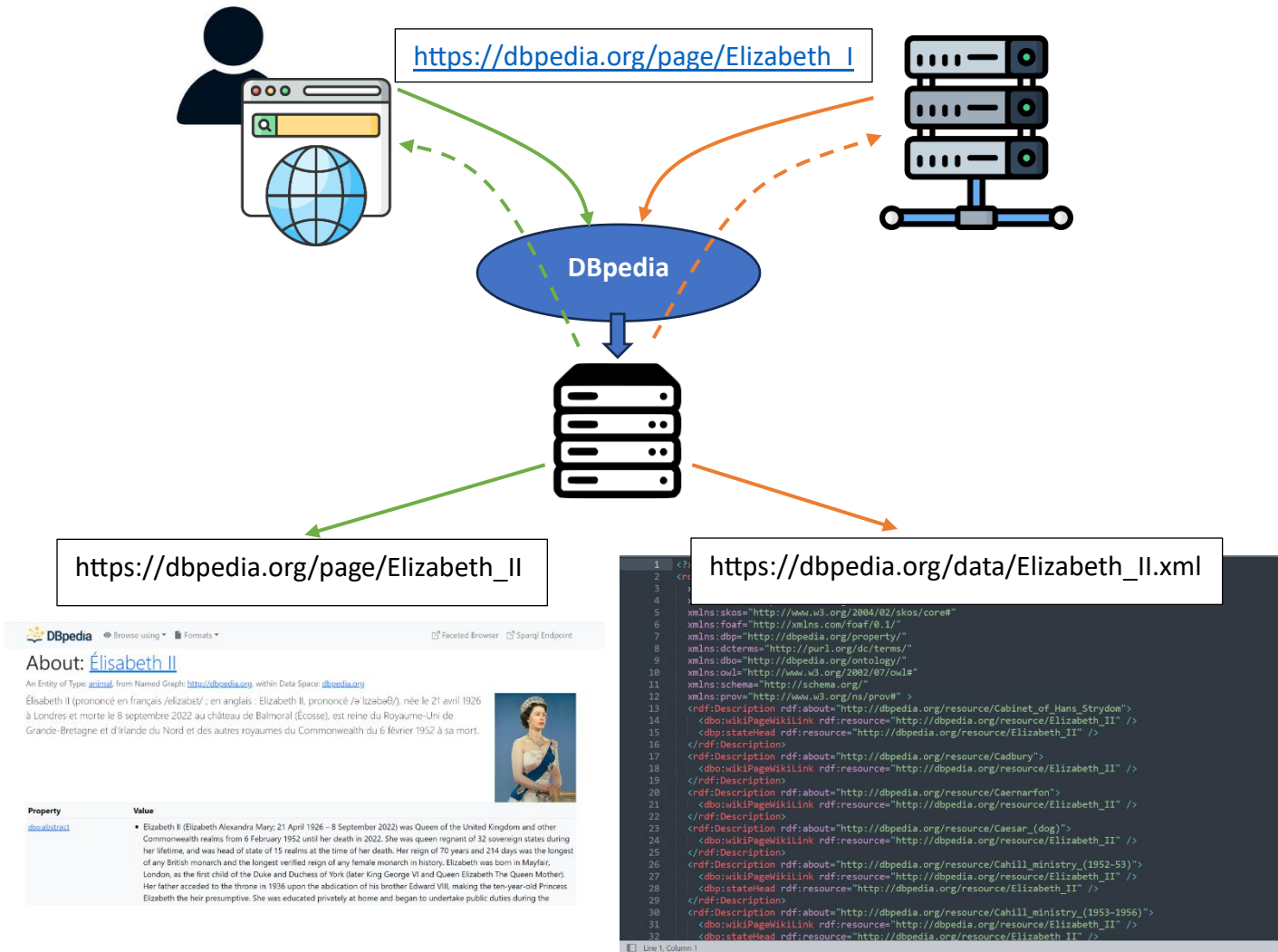
  <rdf:Description rdf:about="http://dbpedia.org/resource/Algeria">
    <rdf:type rdf:resource="http://dbpedia.org/ontology/Country"/>
    <dbo:populationTotal>40400000</dbo:populationTotal>
    <dbo:capital rdf:resource="http://dbpedia.org/resource/Algiers"/>
  </rdf:Description>

</rdf:RDF>
```

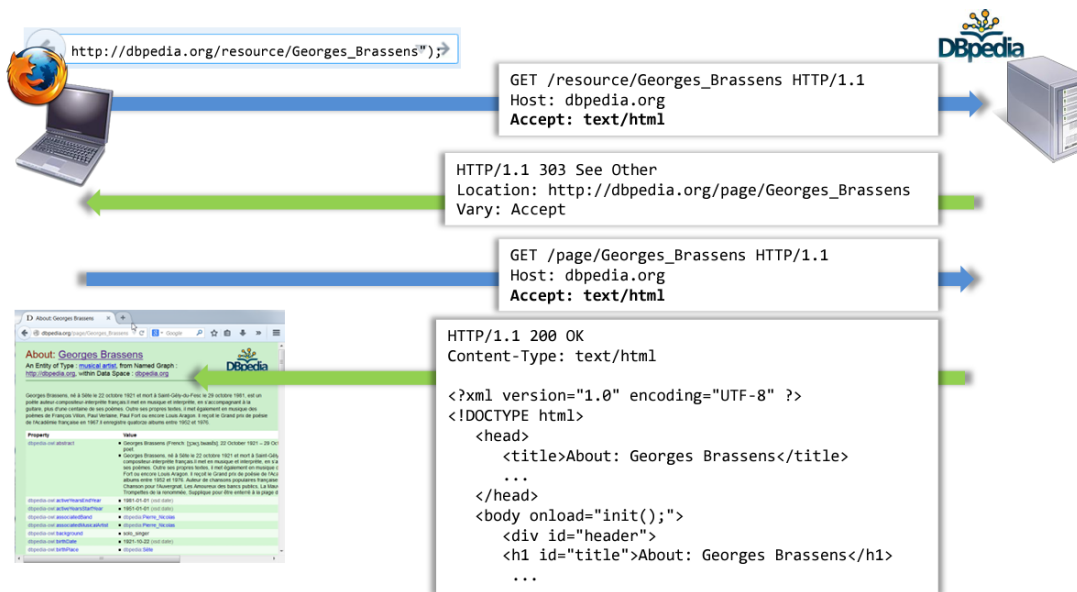
- Copiez le lien https://dbpedia.org/page/Malek_Bennabi, soumettez le dans l'onglet « VISUALIZE WEB PAGE ». Observez le résultat.

2. Négociation de contenu

La même URI peut être utilisé pour récupérer différentes représentations.



L'image suivante décrit une négociation de contenu. Il y a deux requêtes et deux réponses.



1. Dans la première requête quel est le format de données demandé ?
2. Dans la première réponse que signifie le code de réponse HTTP 303 ?
3. Dans la deuxième requête quel est le format de données demandé ?
4. Dans la deuxième réponse que signifie le code de réponse HTTP 200
5. Cherchez trois autres codes de réponse http.

- Simuler le fonctionnement d'une architecture client/serveur

Vous pouvez simuler le fonctionnement d'un navigateur interagissant avec un serveur en utilisant « curl » qui est un outil en ligne de commande largement utilisé pour effectuer des requêtes réseau dans un environnement en ligne de commande ou terminal. Son nom est une contraction de "Client for URLs" (Client pour les URL). Curl permet d'envoyer et de recevoir des données via différents protocoles réseau, notamment HTTP et HTTPS.

Pour effectuer des requêtes HTTP vers un serveur et récupérer des réponses, nous allons utiliser curl. Pour simuler une requête de navigateur, vous pouvez utiliser la commande curl pour envoyer une requête GET à un site web et afficher la réponse. Par exemple, pour simuler l'accès à la page d'accueil de Google, vous pouvez utiliser la commande suivante :

```
curl https://www.google.com
```

1. Exécuter cette requête et observez la réponse.
2. Que fait la requête suivante. Expliquer les options utilisés.

```
curl -L -H "Accept: text/html" http://dbpedia.org/resource/Algeria
```

3. Observez la requête suivante:

```
curl -L -H "Accept: application/rdf+xml" http://dbpedia.org/resource/Algeria
```

Quel est la différence avec la précédente requête. Executer la et observez la réponse.