

## Projet Web des Données et Web Sémantique

Professeurs référents : M. Guillet

Polytech Nantes  
Spécialité INFO

Rapport de projet pour le 01/03/2024

Sean Bogosavac

# 1 Introduction

Ce projet a pour but de créer un site internet exploitant les possibilités du web sémantique. L'objectif est de récupérer une donnée source, de la transformer en un graph de données au format TURTLE, puis de créer un site exploitant, via des requêtes SPARQL, ces données ainsi que des données publiques (e.g. Wikidata).

Ce rapport va passer en revue un projet de site exploitant des données portant sur des conflits historiques. Le code complet du projet est trouvable au dépôt github suivant : Github - Sean Bogosavac Projet Web Sémantique.

Le compte-rendu du mini-projet comprendra :

dans un fichier au format pdf à vos noms et prenom, un rapport (au format pdf) de 5-8 pages présentant vos données, votre application et son interface, qui contiendra : la source (url) de vos données, le squelette openrefine (rdf schema alignment), le schéma rdfs (turtle) et un extrait du fichier rdf (turtle) obtenu, les requetes sparql utilisées, des photo d'écran de l'interface de votre application. l'adresse du git (sources + doc + exemples - voir ci-dessous NE PAS OUBLIER LES INSTRUCTIONS D'INSTALLATION ET DE LANCEMENT)

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Données et génération du fichier turtle</b>	<b>2</b>
<b>3</b>	<b>Design du site et requêtes SPARQL</b>	<b>4</b>
3.1	Page d'accueil . . . . .	4
3.2	Page Guerre . . . . .	4
3.3	Page Bataille . . . . .	5
3.4	Page Commandant . . . . .	6
<b>4</b>	<b>Site final</b>	<b>7</b>

## 2 Données et génération du fichier turtle

Les données utilisées dans ce projet sont tirées du CDB90 Battle Dataset disponible en libre accès sur github à cette adresse. Seuls les fichiers battle\_actors.csv, battle\_durations.csv, battle.csv et commanders.csv seront utilisés, les autres ne représentant que peu d'intérêt.

```
data > battles.csv > data
1 "isqno","war","name","locn","campgn","postype","post1","post2","front","depth","time","aeroa","surpa","cea","leada","trnga","morala","logsa","momnta","intela","tec
2 1,"NETHERLAND'S WAR OF INDEPENDENCE","NIEUPOORT","SPANISH FLANDERS","NIEUPOORT 1600",0,"HD",0,0,0,0,0,0,0,0,1,0,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,1,"NETHERLAND'S W
3 2,"THIRTY YEAR'S WAR","WHITE MOUNTAIN","BOHEMIA","BOHEMIA 1620",1,"HD",1,0,0,0,0,0,0,0,0,0,0,1,0,0,1,1,1,1,1,0,0,0,0,1,0,0,0,0,0,1,"THIRTY YEAR'S WAR","THIR
4 3,"THIRTY YEAR'S WAR","WIMPFEN","PALATINATE","PALATINATE 1622",0,"HD",0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,1,0,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY Y
5 4,"THIRTY YEAR'S WAR","DESSAU BRIDGE","ANHALT","DANISH INVASION OF GERMANY 1625-26",0,"PD",0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,"THIRTY YEAR'S WAR"
6 5,"THIRTY YEAR'S WAR","LUTTER","BRUNSWICK","DANISH INVASION OF GERMANY 1625-26",0,"HD",0,0,0,0,0,0,0,1,0,1,0,1,0,0,0,1,1,1,0,0,0,0,0,0,1,0,0,0,0,0,1,"THIRTY YEA
7 6,"THIRTY YEAR'S WAR","BREITENFELD I","SAXONY","LEIPZIG 1631",0,"HD",0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY YEA
8 7,"THIRTY YEAR'S WAR","THE LECH","BAVARIA","BAVARIA 1632",0,"PD",0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,1,1,0,0,0,0,0,0,1,0,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY YEA
9 8,"THIRTY YEAR'S WAR","ALTE VESTE","FRANCONIA","NUREMBERG 1632",0,"PD",0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY YEA
10 9,"THIRTY YEAR'S WAR","LUETZEN","SAXONY","SAXONY 1632",0,"HD",0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,1,1,0,0,0,0,0,0,0,1,0,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY YEA
11 10,"THIRTY YEAR'S WAR","NORDLINGEN I","BAVARIA","BAVARIA 1634",0,"PD",0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY YEA
12 11,"THIRTY YEAR'S WAR","MITTSTOCK","BRANDENBURG","E GERMANY 1636",0,"PD",0,0,0,0,1,1,0,0,0,0,0,0,1,1,1,1,0,0,0,0,0,0,1,1,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY
13 12,"THIRTY YEAR'S WAR","BREITENFELD II","SAXONY","SAXONY 1642",0,"HD",0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,1,1,0,0,0,0,0,0,1,0,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY Y
14 13,"THIRTY YEAR'S WAR","ROCROI","FRANCE","NE FRANCE 1643",0,"HD",0,0,0,0,0,0,0,2,0,0,0,1,1,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY YEA
15 14,"THIRTY YEAR'S WAR","TUTTLINGEN","SWABIA","SWABIA 1643",0,"HD",0,0,0,0,0,2,0,2,0,0,0,0,1,0,1,1,1,0,0,0,0,0,0,0,2,2,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY
16 15,"THIRTY YEAR'S WAR","FREIBURG","SWABIA","SWABIA 1644",0,"PD",0,0,0,0,0,0,0,0,0,0,0,1,0,1,1,0,1,0,0,0,0,0,0,1,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY YEA
17 16,"THIRTY YEAR'S WAR","JANKAU","BOHEMIA","BOHEMIA 1645",0,"HD",0,0,0,0,3,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,1,0,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY YEA
18 17,"THIRTY YEAR'S WAR","MERGETHEIM","BAVARIA","BAVARIA 1645",0,"HD",0,0,0,0,3,0,1,0,0,0,1,0,1,1,1,0,0,0,0,0,0,0,1,2,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY Y
19 18,"THIRTY YEAR'S WAR","KALLERHEIM (NORDLINGEN II)","BAVARIA","BAVARIA 1645",0,"PD",0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,"THIRTY YEAR'S W
20 19,"THIRTY YEAR'S WAR","LENS","FRANCE","NE FRANCE 1648",0,"HD",0,0,0,0,2,0,1,0,0,0,1,0,1,1,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,"THIRTY YEAR'S WAR","THIRTY YEA
21 20,"ENGLISH CIVIL WAR","EDGEHILL","ENGLAND","EDGEHILL",0,"HD",0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,1,"ENGLISH CIVIL WAR","ENGLISH CIVIL W
22 21,"ENGLISH CIVIL WAR","MARSTON MOOR","ENGLAND","YORK",0,"HD",0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,1,0,0,0,0,1,"ENGLISH CIVIL WAR","ENGLISH CIVIL W
23 22,"ENGLISH CIVIL WAR","TIPPERMUIR","SCOTLAND","ABERDEEN",0,"HD",0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,"ENGLISH CIVIL WAR","ENGLISH CIVIL WAR"
24 23,"ENGLISH CIVIL WAR","KILSYTH","SCOTLAND","KILSYTH",0,"HD",0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,1,0,0,0,0,1,"ENGLISH CIVIL WAR","ENGLISH CIVIL
25 24,"ENGLISH CIVIL WAR","NEWBURY II","ENGLAND","NEWBURY II",1,"HD",1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,"ENGLISH CIVIL WAR","ENGLISH CI
26 25,"ENGLISH CIVIL WAR","NASEBY","ENGLAND","NASEBY",0,"HD",0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,1,"ENGLISH CIVIL WAR","ENGLISH CIVIL WAR","Firs
27 26,"SECOND ENGLISH CIVIL WAR","PRESTON","ENGLAND","PRESTON",0,"HD",0,0,0,0,2,2,1,0,0,0,1,0,1,1,1,1,0,0,0,0,0,0,1,2,0,0,0,0,1,"SECOND ENGLISH CIVIL WAR","SECON
28 27,"SECOND ENGLISH CIVIL WAR","DUNBAR","SCOTLAND","DUNBAR",0,"HD",0,0,0,0,2,0,1,1,0,0,0,0,0,1,1,0,0,1,1,0,0,0,0,1,1,2,0,0,0,1,"SECOND ENGLISH CIVIL WAR","SECO
29 28,"SECOND ENGLISH CIVIL WAR","WORCESTER","ENGLAND","WORCESTER",1,"HD",1,0,0,0,0,0,1,1,0,0,0,0,0,0,0,1,1,1,0,0,0,0,0,1,0,0,0,0,0,1,"SECOND ENGLISH CIVIL WAR"
30 29,"THE FRONDE","ST. ANTOINE","FRANCE","THE FRONDE",1,"HD",1,1,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,0,2,0,0,1,0,0,0,0,0,0,0,0,1,"THE FRONDE","THE FRONDE","The Fron
31 30,"FRANCO-SPANISH WAR","THE DUNES","SPANISH NETHERLANDS","DUNKIRK 1658",0,"HD",0,0,0,0,1,2,0,0,0,0,0,0,0,1,1,1,1,1,0,0,0,0,2,1,0,0,0,0,1,"FRANCO-SPANISH W
32 31,"AUSTRO-TURKISH WAR","THE RAAB","HUNGARY","HUNGARY 1664",0,"HD",0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,"AUSTRO-TURKISH WAR, 1662-1683","AUSTA
33 32,"AUSTRO-TURKISH WAR","VIENNA","AUSTRIA","AUSTRIA 1683",1,"HD",1,1,0,0,2,1,1,0,0,0,0,0,0,1,1,1,0,0,0,0,0,0,1,1,0,0,0,0,1,"AUSTRO-TURKISH WAR, 1662-1683"
34 33,"POLISH-TURKISH WAR","CHOCIM II","POLAND","CHOCIM",0,"PD",0,0,0,0,3,1,0,0,0,0,0,0,1,1,1,0,0,0,0,0,1,0,1,0,0,0,1,"POLISH-TURKISH WAR","POLISH-TURKISH WA
35 34,"DUTCH WAR","SINSHEIM","GERMANY","RHINELAND 1674",0,"HD",0,0,0,0,0,0,2,0,0,0,0,0,0,1,1,0,0,0,0,0,0,2,0,0,0,0,1,"DUTCH WAR","DUTCH WAR","Franco-Dutch War
36 35,"DUTCH WAR","SENEF","SPANISH NETHERLANDS","SPANISH NETHERLANDS 1674",0,"HD",0,0,0,0,2,1,0,0,0,0,0,0,1,1,0,2,0,0,0,0,0,0,1,2,0,0,0,0,1,"DUTCH WAR","DUTCH
37 36,"DUTCH WAR","ENZHEIM","GERMANY","RHINELAND 1674",1,"HD",1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,2,2,0,0,0,0,0,0,1,0,0,0,0,1,"DUTCH WAR","DUTCH WAR","Franco-Dut
```

FIGURE 1 – Extrait du fichier battles.csv

Toutes les colonnes ne seront pas récupérées, mais on peut déjà voir une partie des données pertinentes (nom de la bataille, lieu, guerre à laquelle elle est associée, etc.).

Les données des différents fichier csv sont ensuite rassemblées et transformées en un fichier TURTLE par le script python csvToRdf.py, via le schéma RDFS suivant :

```
# Define types and relations
types = [
    ":Battle a rdfs:Class .\\n\\n",
    ":Belligerent a rdfs:Class .\\n\\n",
    ":Commander a rdfs:Class .\\n\\n",
    ":Location a rdfs:Class .\\n\\n",
    ":War a rdfs:Class .\\n\\n",
    ":Side a rdfs:Class .\\n\\n\\n"
]
file.writelines(types)

relations = [
    ":hasName a rdf:Property ;\\n",
    "  rdfs:domain :Battle .\\n",
    "  rdfs:range xsd:string .\\n\\n",
    ":hasEndpoint a rdf:Property ;\\n",
    "  rdfs:domain :Battle .\\n",
    "  rdfs:range :Commander .\\n",
    ":isPartOf a rdf:Property ;\\n",
    "  rdfs:domain :Battle .\\n",
    "  rdfs:range :War .\\n\\n",
    ":takesPlaceIn a rdf:Property ;\\n",
    "  rdfs:domain :Battle .\\n",
    "  rdfs:range :Location .\\n\\n",
    ":resultedIn a rdf:Property ;\\n",
    "  rdfs:domain :Battle .\\n",
    "  rdfs:range xsd:integer .\\n\\n",
    ":begins a rdf:Property ;\\n",
    "  rdfs:domain :Battle .\\n",
    "  rdfs:range xsd:date .\\n\\n",
    ":ends a rdf:Property ;\\n",
    "  rdfs:domain :Battle .\\n",
    "  rdfs:range xsd:date .\\n\\n",
    ":isAttacker a rdf:Property ;\\n",
    "  rdfs:domain :Side .\\n",
    "  rdfs:range :Belligerent .\\n\\n",
    ":foughtIn a rdf:Property ;\\n",
    "  rdfs:domain :Side .\\n",
    "  rdfs:range :Battle .\\n\\n",
    ":foughtFor a rdf:Property ;\\n",
    "  rdfs:domain :Commander .\\n",
    "  rdfs:range :Side .\\n\\n",
    ":wasInSide a rdf:Property ;\\n",
    "  rdfs:domain :Belligerent .\\n",
    "  rdfs:range :Side .\\n\\n\\n"
]
file.writelines(relations)
```

FIGURE 2 – Définition des types et relations dans le script python

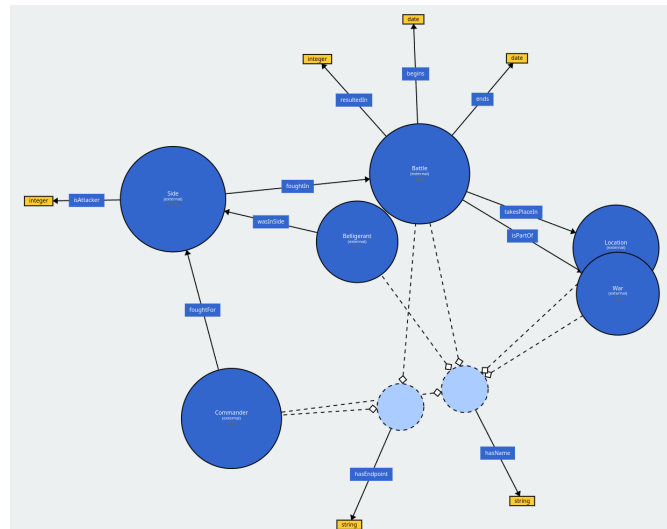


FIGURE 3 – Schéma RDFS (affiché via WebVOWL)

## 3 Design du site et requêtes SPARQL

Le site créé pour afficher ces données sera composé de 4 pages :

### 3.1 Page d'accueil

La page d'accueil listera les guerres dont font partie les batailles de la base de données. L'objectif étant de fournir à l'utilisateur cette liste, sous la forme de boutons, qui l'ammeneront à une page dédiée à la guerre sélectionnée.

Les données utilisées dans cette page pourront être récupérées via la requête SPARQL suivante :

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX : <http://127.0.0.1:3030/>

SELECT DISTINCT ?war ?name
WHERE {
    ?war a :War .
    ?war :hasName ?name .
}
```

### 3.2 Page Guerre

L'objectif de la page Guerre sera de résumer, pour la guerre sélectionnée, les batailles qui s'y sont déroulés. L'affichage de ces informations se fera sous plusieurs formes : une carte, affichant pour chaque bataille un point pour que l'utilisateur puisse se faire une idée de la localisation géographique du conflit, et une liste des batailles, ordonnées chronologiquement, avec des boutons pour que l'utilisateur puisse avoir accès à une page descriptive de la bataille de son choix.

Les coordonnées des batailles ne sont pas présentes dans le dataset, mais elles peuvent être récupérés depuis DBPedia, grâce à la présence des URI DBPedia de chaque bataille au sein du dataset.

Les données utilisées dans cette page pourront être récupérées via la requête SPARQL suivante :

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX georss: <http://www.georss.org/georss/>
PREFIX : <http://127.0.0.1:3030/>
```

```

SELECT DISTINCT ?warName ?battle ?name ?beginDate ?coord ?abstract
WHERE {
    ?battle a :Battle .
    ?battle :hasName ?name .
    ?battle :isPartOf :${warTag} .
    :${warTag} :hasName ?warName .
    ?battle :begins ?beginDate .
    ?battle :hasEndpoint ?endpoint .

    SERVICE <https://dbpedia.org/sparql> {
        ?endpoint georss:point ?coord .
    }
}
ORDER BY ?beginDate

```

### 3.3 Page Bataille

La page Bataille donnera diverse informations sur la bataille sélectionnée : une brève description (récupérée depuis DBPedia), sa localisation sur une carte (de la même manière qu'à la page guerre), ainsi que les différents belligérants impliqués, avec chacun leur commandants (sous la forme de boutons cliquables amenant à la page du commandant voulu), ainsi que le résultat de la bataille.

Les données utilisées dans cette page pourront être récupérées via la requête SPARQL suivante :

```

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX georss: <http://www.georss.org/georss/>
PREFIX : <http://127.0.0.1:3030/>

SELECT DISTINCT ?name ?coord ?abstract ?result ?belName ?commander ?c
WHERE {
    :${battleId} :hasName ?name .
    :${battleId} :hasEndpoint ?endpoint .
    :${battleId} :resultedIn ?result .

    ?side :foughtIn :${battleId} .
    ?side :isAttacker ?isAttacker .

```

```

?belligerant :wasIn ?side .
?belligerant :hasName ?belName .

?commander :foughtFor ?side .
?commander :hasName ?comName .

SERVICE <https://dbpedia.org/sparql> {
  ?endpoint georss:point ?coord .
  ?endpoint dbo:abstract ?abstract .
  FILTER(LANG(?abstract) = "en")
}

```

### 3.4 Page Commandant

Enfin la page Commandant donnera la liste des batailles auquel le commandant sélectionné a participé, ainsi que le camp pour lequel il s'est battu. Les données utilisées dans cette page pourront être récupérées via la requête SPARQL suivante :

```

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX : <http://127.0.0.1:3030/>

SELECT DISTINCT ?name ?abstract ?battleName ?belName ?side ?battle
WHERE {
  :${commanderTag} a :Commander .
  :${commanderTag} :hasEndpoint ?name .

  :${commanderTag} :foughtFor ?side .
  ?belligerant :wasIn ?side .
  ?belligerant :hasName ?belName .
  ?side :foughtIn ?battle .
  ?battle :hasName ?battleName .
}

```

## 4 Site final

Voici une capture d'écran de chaque page mentionnée ci-dessus en guise d'exemple.  
(Le style des pages n'a pas pu être travaillé par manque de temps)



FIGURE 4 – Page d'accueil



FIGURE 5 – Exemple de page Guerre (Guerre de la Deuxième Coalition)





FIGURE 6 – Exemple de page Bataille (Bataille de Novi)

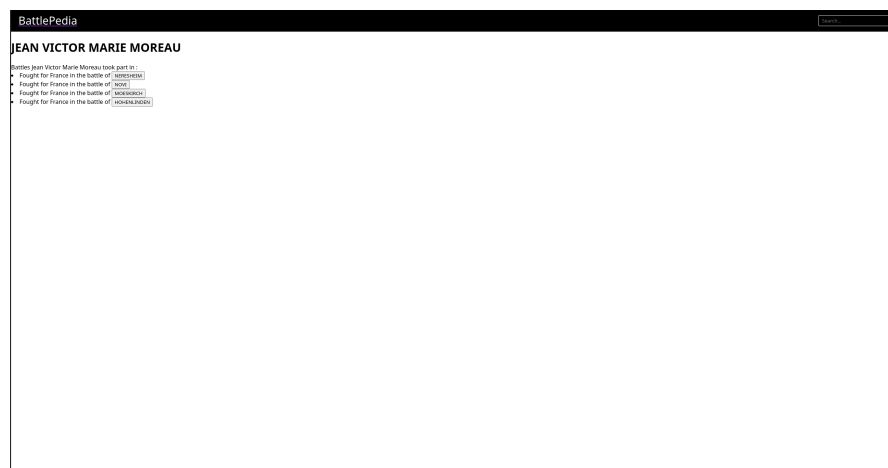


FIGURE 7 – Exemple de page Commandant (Jean Moreau)