

# Web Sémantique

## Rapport final du mini-projet

Geneviève CIRERA - Tony BULTE

Le 10 février 2015

# Introduction Rappel du contexte

L'objectif de ce projet est de développer une application web en mettant en avant l'utilisation du web sémantique pour la structuration des connaissances et l'utilisation du web de données pour la récupération de connaissances déjà existantes sur l'internet.

## *Sujet*

Pour le développement de notre application Web sémantique, nous avons choisi le thème des Pays. Ce thème est très riche et peut faire appel à beaucoup d'autres connaissances rattachées.

Pour cette application web, nous allons développer un site d'apprentissage sur le thème des pays.

Notre site sera divisé en deux parties :

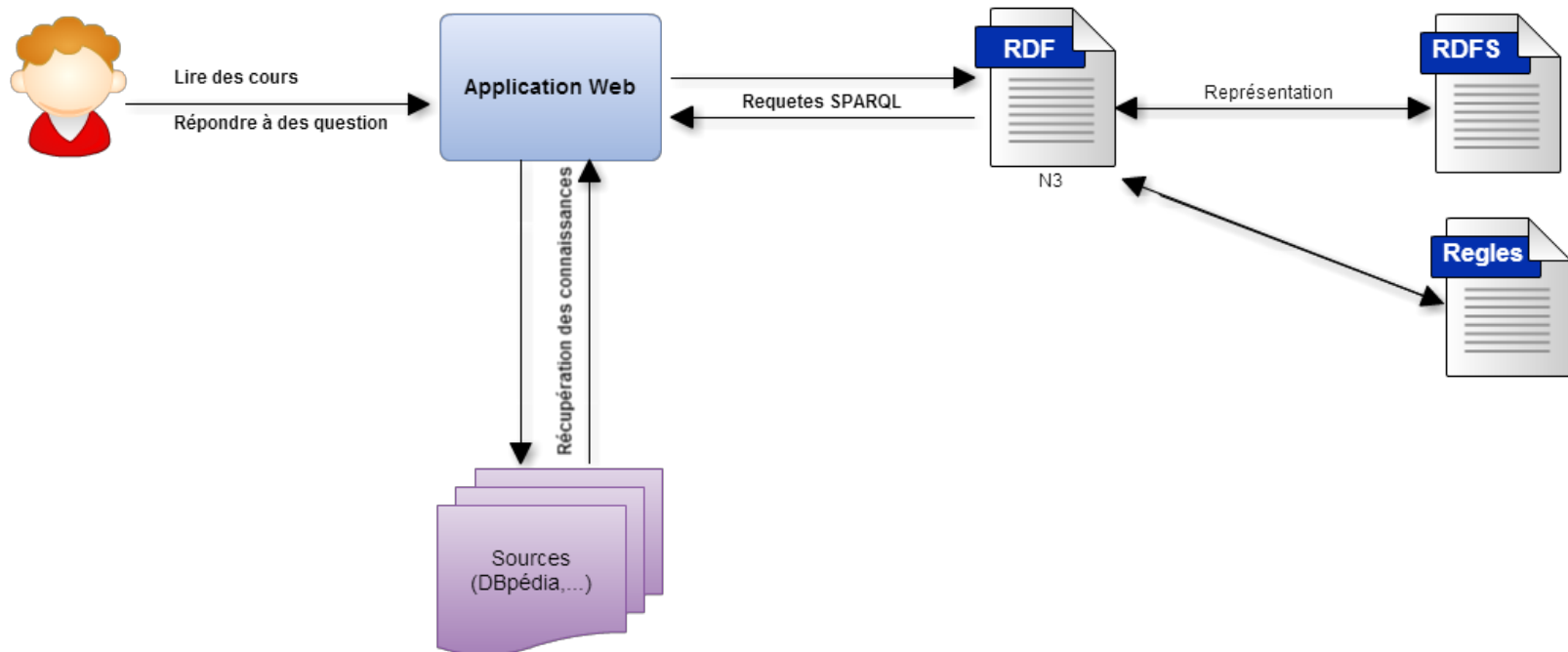
Une partie cours : Pour mettre en avant le côté "apprentissage"

Une partie exercices : Pour tester les connaissances de l'utilisateur

Pour rappel, voici le schema fonctionnel qui a été implémenté pour notre application.

# Application Web Sémantique

Schéma fonctionnel



## Les données

Les données dont nous disposions étaient DBpédia et notre fichier RDF où nous avons ajouté des données sur trois pays, la France, l'Espagne et l'Allemagne telles que l'origine du nom du pays, des villes situées dans le pays, leurs classements au niveau économique, les groupes auxquels ils appartiennent (OTAN, G20,...),...

Lors de l'implémentation de notre application, nous nous sommes rendus compte que la commande `java -jar corese-server.jar -load notreFichier.rdf` ne nous permettait pas de faire des requêtes sur notre fichier RDF. En revanche charger notre fichier en N3 fonctionnait parfaitement. Nous avons donc décidé de traduire le RDF en N3.

Quelques modifications ont été apportées aux données. Tout d'abord la modification des questions qui sont définies par un "string" qui est la question et un "paramètre" afin que dans la partie exercice, les questions sont récupérées puis une aléatoire est choisie

et est affichée à l'utilisateur sur un pays. Le paramètre sert d'indication pour la requête associée à la question qui choisi un pays et la réponse à la question sélectionnée. Ensuite, des relations d'héritage ont été ajoutées par exemple "Pays" est sous-classe de la classe "Country" de DBpédia. Des "comments" et des "labels" ont été ajoutés pour une meilleure compréhension du RDFS.

Les règles définies sont appliquées avant le démarrage de l'application et le résultat est conservé dans un fichier car le serveur corese n'exécute pas les règles.

## Partie Cours

La partie "Cours" a pour but de donner la possibilité à l'utilisateur de consulter toutes les données sur lesquelles les questions de la partie "Exercice" peuvent porter.

C'est pourquoi, les données sont classées en catégorie, "Allemagne", "Espagne" et "France" (qui porte sur les données N3 que nous avons définies) et "Capitale", "Devise", "Président", "Surface" et "Langue" qui donne respectivement la capitale, la devise, le président, la surface en km2 et la langue de chaque pays trouvé sur DBpédia.

Les catégories "Allemagne", "Espagne" et "France" ont des sous-catégories suivant l'information que l'on souhaite qui s'affiche lors du clique sur "Allemagne", "Espagne" ou "France". Ces sous-catégories sont : "Origine du nom", "Continent", "Groupe", "Classement éco" et "Ville".

"Origine du nom" donne l'origine du nom du pays, "Continent", dans quel continent se situe le pays, "Groupe", les groupes auxquels appartient le pays (G20, Espace Schengen,...), "Classement éco" donne le rang du classement économique au niveau européen et/ou mondial. Enfin, "Ville" liste toutes les villes du pays définies dans le fichier nt.

Toutes les requêtes sont exécutées par l'intermédiaire de javascript sur des données de DBpédia où les données que nous avons définies et chargé en locale sur le serveur corese.

Le résultat de chaque requête est ensuite parsé afin l'afficher sur la page html de manière correct, suppression des guillemets, espaces, underscore...en trop. De plus, afin d'afficher les noms sans l'url, les requêtes vont jusqu'à chercher le nom de la ressource ce qui renvoie une chaîne de caractère et non une url.

# Partie Exercice

La partie Exercice va permettre à l'utilisateur de tester ses connaissances. Deux type d'exercices ont été implémentés :

- **Quizz**

Principe : Une question aléatoire est posée à propos d'un pays aléatoire. L'utilisateur dispose d'une zone de texte pour répondre à la question et peut vérifier si sa réponse est la bonne.

- **QCM**

Principe : Une question aléatoire est posée à propos d'un pays aléatoire. Trois réponses sont proposées, dont une qui est la bonne réponse à la question posée.

Pour ces exercices, quatre questions ont été implémentées :

“Quelle est la surface de %pays% ? (en km2)”

“En %pays%, quelle est la capitale ?”

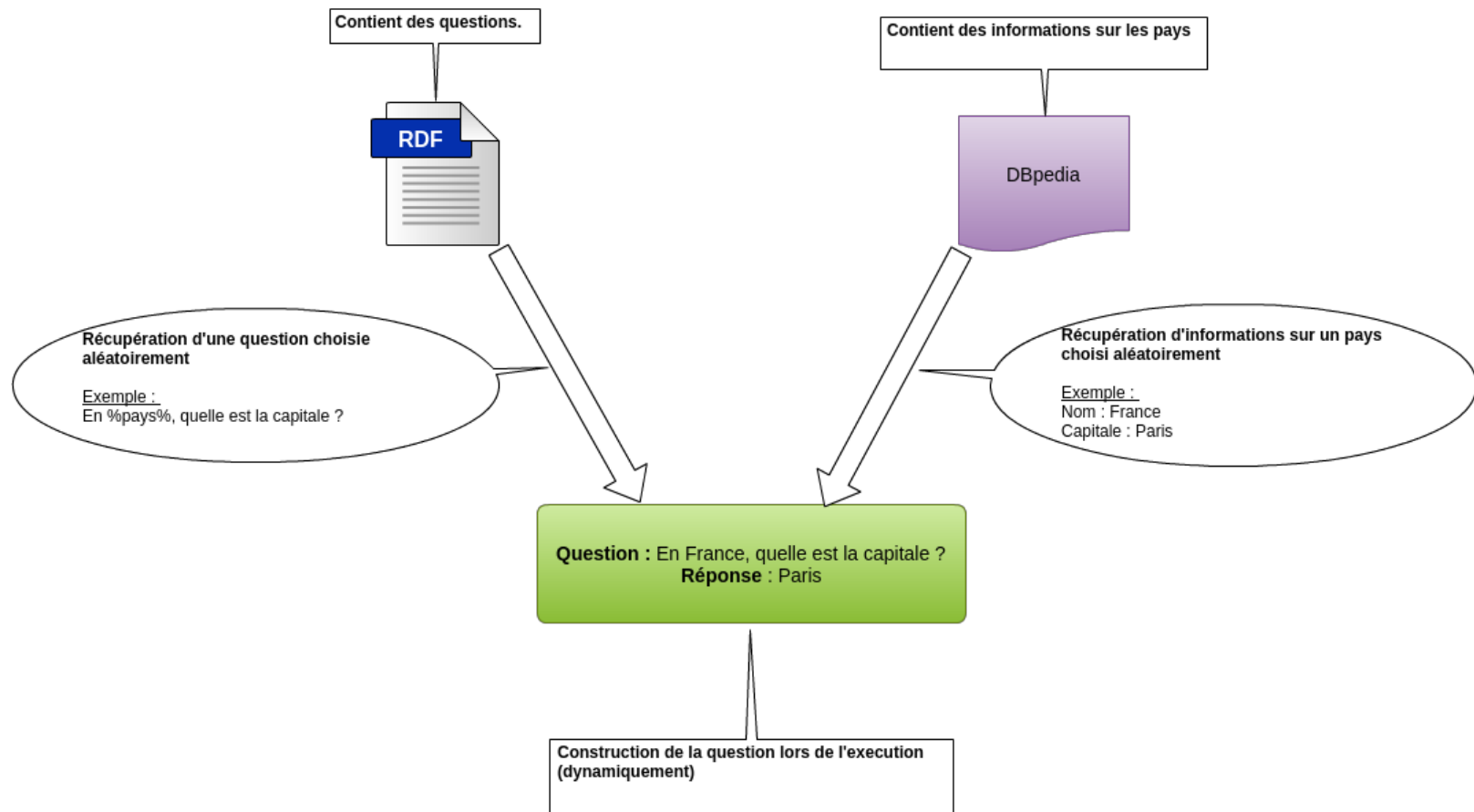
“En %pays%, qui est le président ?”

“En %pays%, quelle est la langue ?”

Le schéma ci-dessous présente la façon de générer les question :

# Application Web Sémantique

## Partie exercice



La construction des questions se passe comme suit :

1. Des squelettes de questions sont stockés dans un fichier RDF.
2. Une question est choisie aléatoirement
3. Une requête SPARQL est exécutée vers DBpedia pour récupérer des informations sur un seul pays choisi aléatoirement
4. La question est construite à l'exécution
5. La question est affichée dynamiquement dans l'application web

Une fois la question affichée, l'utilisateur peut tenter de répondre à la question qui lui est posée. Il peut corriger sa réponse. Cette correction va lui permettre de savoir s'il a répondu juste ou non, et s'il n'a pas répondu juste, la bonne réponse est affichée.

## Les difficultés rencontrées

Les principales difficultés nous sont venues de la récupération des résultats des requêtes. En effet, lorsque l'on envoie une requête via Corese, nous récupérons en résultat une page HTML contenant les réponses. Cela est très peu pratique car nous devons parser la page HTML reçue pour récupérer la liste des résultats.

## Conclusion

Ces cinq semaines de projet nous ont permis de pouvoir proposer une application web d'apprentissage des pays avec des valeurs ajoutées comme des informations toujours à jour car elles sont récupérées de la base de connaissance DBpedia, des questions stockées dans un fichier RDF et un noyau de connaissances qui peut être enrichi facilement par des potentiels successeurs à notre projet.