



# Gestion de données complexes

## GMIN 332

ABBOUTE Amayas  
ENTRINGER Gilles

- Web sémantique
- Différentes technologies de stockages
- Approche NoSQL
- Langage de requêtes Sparql



## Plan :

- Technologies utilisées
- Ontologies
- Jeux de données
- Réalisation
- Conclusion
- Démonstration

- API Jena
  - TDB
  - SDB (BDD utilisée : MySQL)
- D2RQ (BDD utilisée : Oracle)
- NoSQL
  - Hbase
  - Neo4j

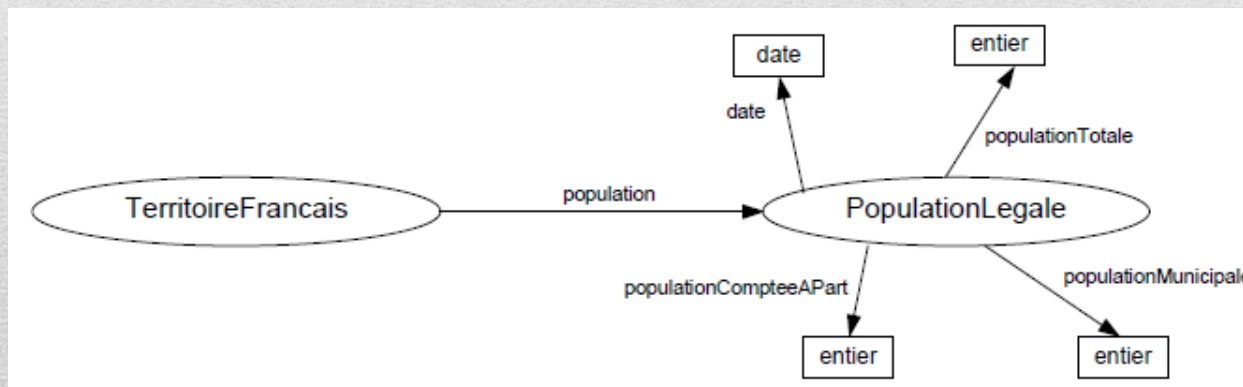


## INSEE COG

- Code officiel géographique
- Ontologie OWL

## INSEE Population

- Ontologie OWL



Cf. <http://rdf.insee.fr/def/index.html#onto-demo>

## GeoNames

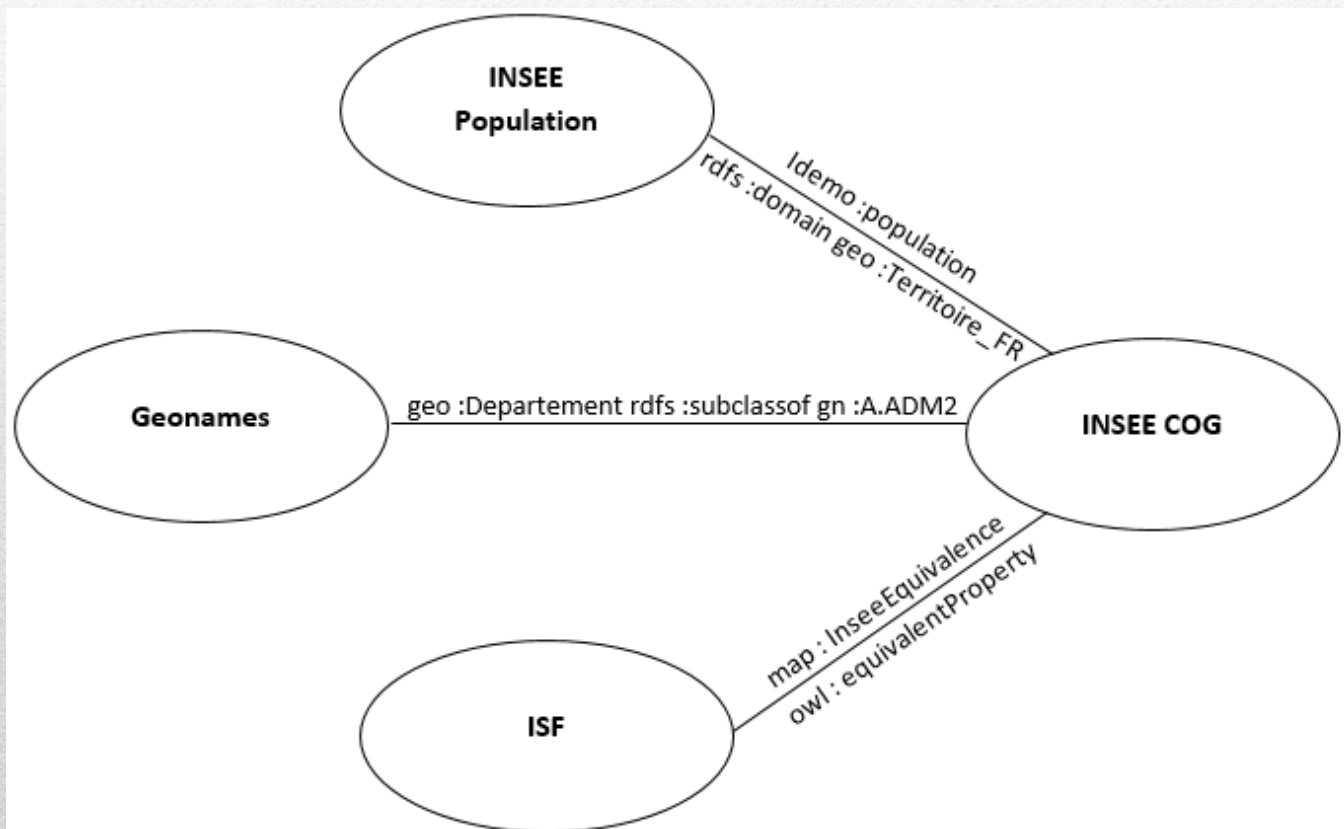
- Ontologie OWL

## Mapping D2RQ

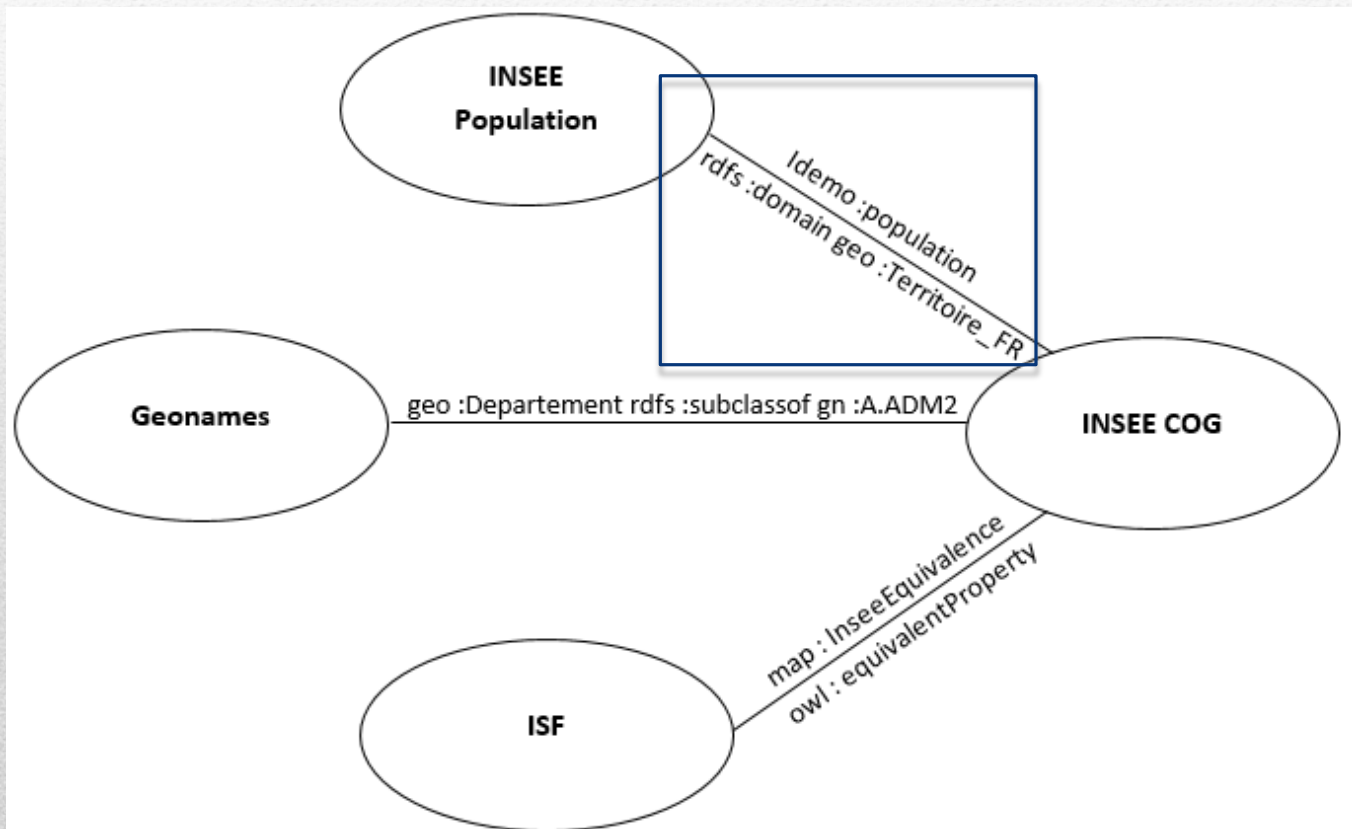
- Fichier de mapping : Format Turtle



## Liaison des ontologies :



## Liaison des ontologies :





## INSEE Population => INSEE COG

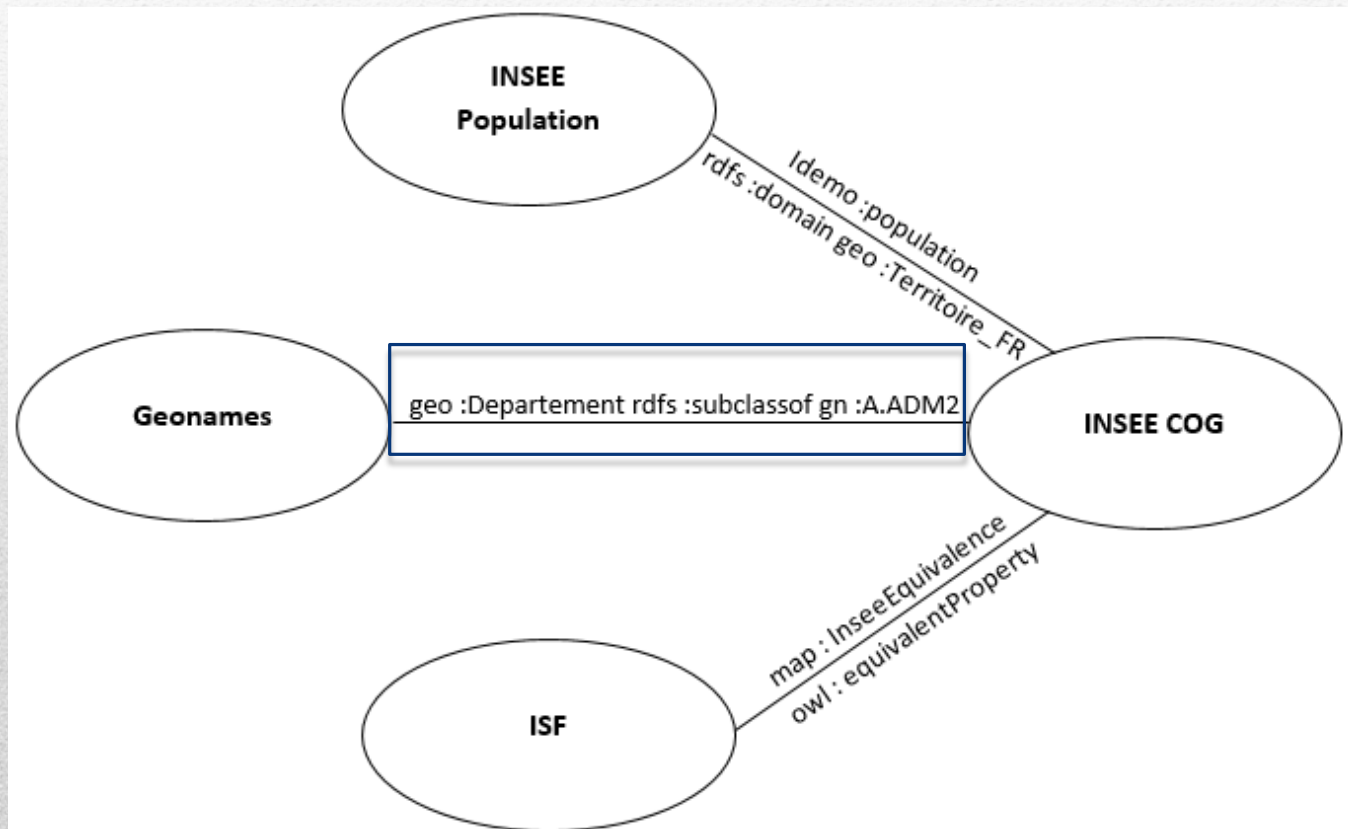
Avant :

```
idemo:population
  a owl:ObjectProperty ;
  rdfs:label "population"@fr ;
  rdfs:isDefinedBy <http://rdf.insee.fr/def/demo> ;
  rdfs:domain igeo:TerritoireFrancais ;
  rdfs:range idemo:PopulationLegale .
```

Après :

```
idemo:population
  a owl:ObjectProperty ;
  rdfs:label "population"@fr ;
  rdfs:isDefinedBy <http://rdf.insee.fr/def/demo> ;
  rdfs:domain geo:Territoire_FR ;
  rdfs:range idemo:PopulationLegale .
```

## Liaison des ontologies :

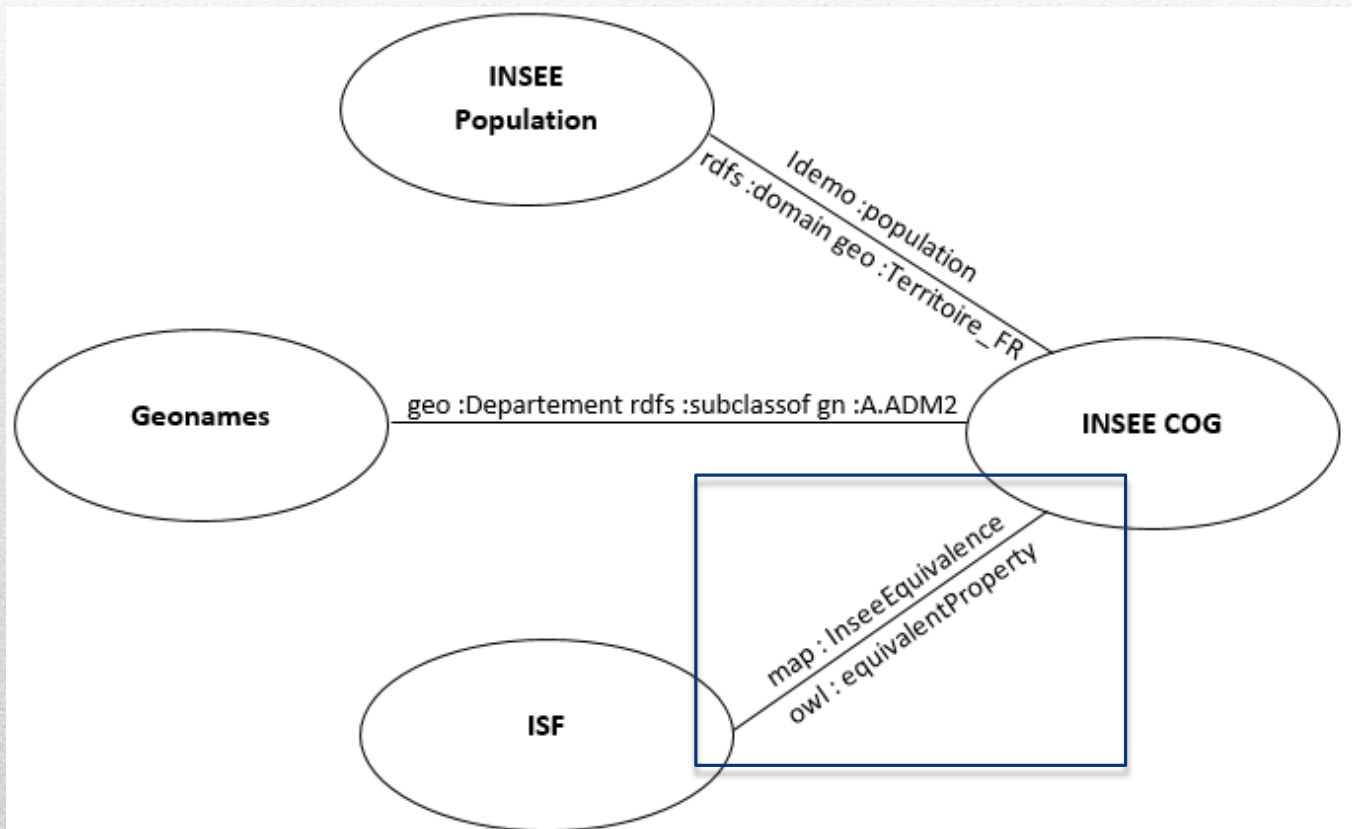




## GeoNames => INSEE COG

```
<owl:Class rdf:about="http://rdf.insee.fr/geo/Departement">
  <rdfs:subClassOf rdf:resource="http://www.geonames.org/ontology#A.ADM2"/>
  <rdfs:subClassOf>
    <owl:Class rdf:about="http://rdf.insee.fr/geo/Territoire_FR"/>
  </rdfs:subClassOf>
  <rdfs:subClassOf rdf:resource="http://rdf.insee.fr/geo/NUTS_3"/>
  <rdfs:label>Département</rdfs:label>
</owl:Class>
```

## Liaison des ontologies :





## D2RQ Mapping

```
map:IMPOT_CODEINSEE a d2rq:PropertyBridge;  
    d2rq:belongsToClassMap map:IMPOT;  
    d2rq:property impots:IMPOT_CODEINSEE;  
    d2rq:additionalClassDefinitionProperty map:InseeEquivalence;  
    d2rq:column "IMPOT.CODEINSEE";  
.  
map:InseeEquivalence a d2rq:AdditionalProperty;  
    d2rq:propertyName owl:equivalentProperty;  
    d2rq:propertyValue geo:code_INSEE;
```

## INSEE COG

- Arrondissements, Cantons, Départements, Régions, Communes
- Format tabulaire (.csv) + RDF
- Stockage : Hbase, Neo4J, TDB

## INSEE Population

- Format Turtle
- Stockage : SDB (Mysql)



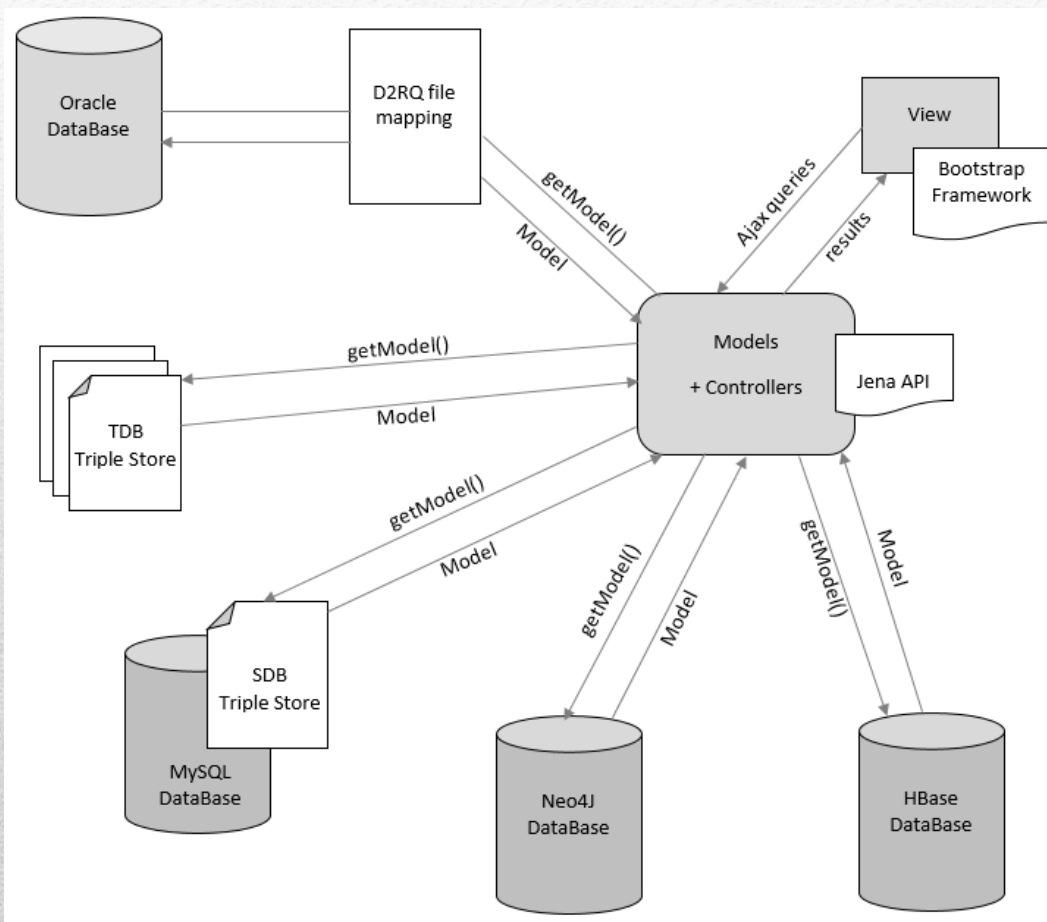
## GeoNames

- Format RDF
- Stockage : TDB

## ISF (Impôts de solidarité sur la fortune)

- Format tabulaire (.csv)
- Stockage : D2RQ (Oracle)

## Architecture générale de l'application :



- Apache Tomcat
- Ajax



## Génération dynamique du modèle

- 1 modèle généré par requête (modèle global)
- Chargement des sources de données sélectionnées par l'utilisateur

### Exemple : TDB (Régions)

1. Chargement du modèle à partir du Triplestore
2. Ajout au modèle global de l'application

## Interface utilisateur

### Gestion de données Complexes

Accueil

Sources

Résultats

Ontologies

Send

Prefixes



Introduction	Technologies utilisées	Ontologies	Jeux de données	Réalisation	Conclusion
--------------	---------------------------	------------	--------------------	-------------	------------

Difficultés rencontrées :

- Compréhension des ontologies GeoNames et INSEE
- Différentes versions d'ontologies de l'INSEE
- Lenteur de Neo4J => Utilisation de Cypher

Expériences apportées

- Découverte de nombreuses technologies (Jena API, NoSQL, Triplestores etc.)
- Evaluation et comparaison des technologies SDB et TDB
- Approche web sémantique

Introduction	Technologies utilisées	Ontologies	Jeux de données	Réalisation	Conclusion
--------------	---------------------------	------------	--------------------	-------------	------------

# Démonstration