Table of Contents

Aim

The aim of this course work is to define a plausible ontology by implementing Semantic web principles learnt and to populate the ontology with data from two sets of external data source and to create SPARQL query demonstrate the correctness of the ontology and the data loaded into the ontology.

Ontology

The project is all about defining an Ontology for Companies. Once the ontology is created, data from external source like DBpedia will be loaded. To demonstrate the correctness of the ontology, the ontology can be queried using SPARQL queries.

Creating the Ontology

First the data source needs to be identified. The source should essentially provide SPARQL endpoint to fetch the data and load it into the ontology. The data and its properties should be studied properly so a meaningful ontology can be created.

Once the properties are analyzed, the ontology needs to be structured so it can be a source of cohesive and wholistic data about the subject. Significant research is required to model the ontology to establish the purpose. To create the ontology, I have used Protégé tool which is one of the most popular tools available to create ontologies.

Otology Structure

The three main structural units of the ontology are as follows

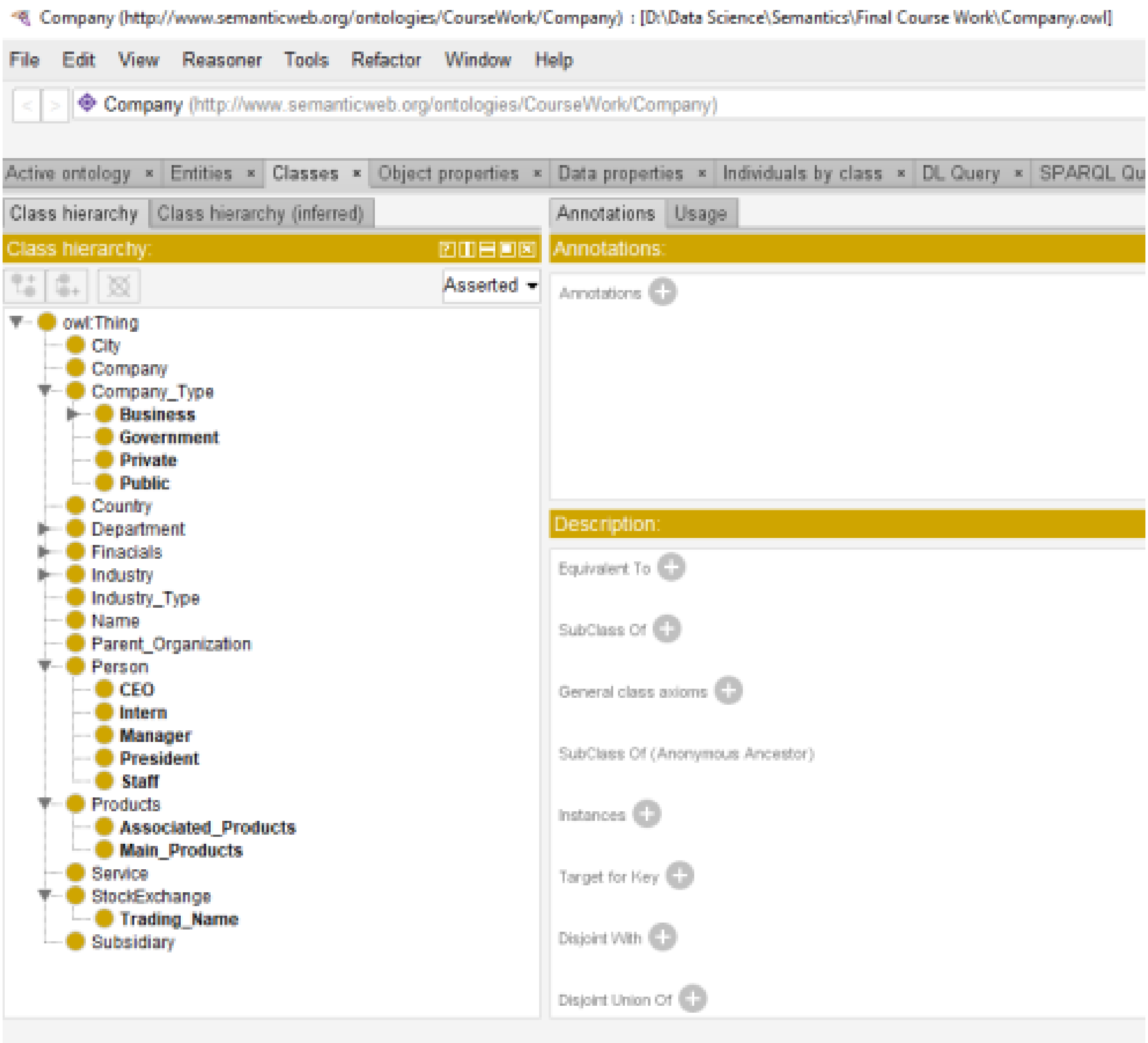
1. Class Hierarchy
2. Object Properties
3. Data Properties

Let us analyze them in detail

- **Class hierarchy**

The success of the ontology lies in identifying the classes and to create object and data properties to connect the classes. Classes needs to be structed so the subject is represented properly and are well connected through object and data properties.

It was quite challenging for me to decide and structure the classes. To overcome this, I studied extensively the data and its properties available in the data source using SPARQL. I have created 15 top level classes and 22 sub-classes. The below picture (Picture1) depicts the class hierarchy of the ontology.



Company (http://www.semanticweb.org/ontologies/CourseWork/Company) : [D:\Data Science\Semantics\Final Course Work\Company.owl]

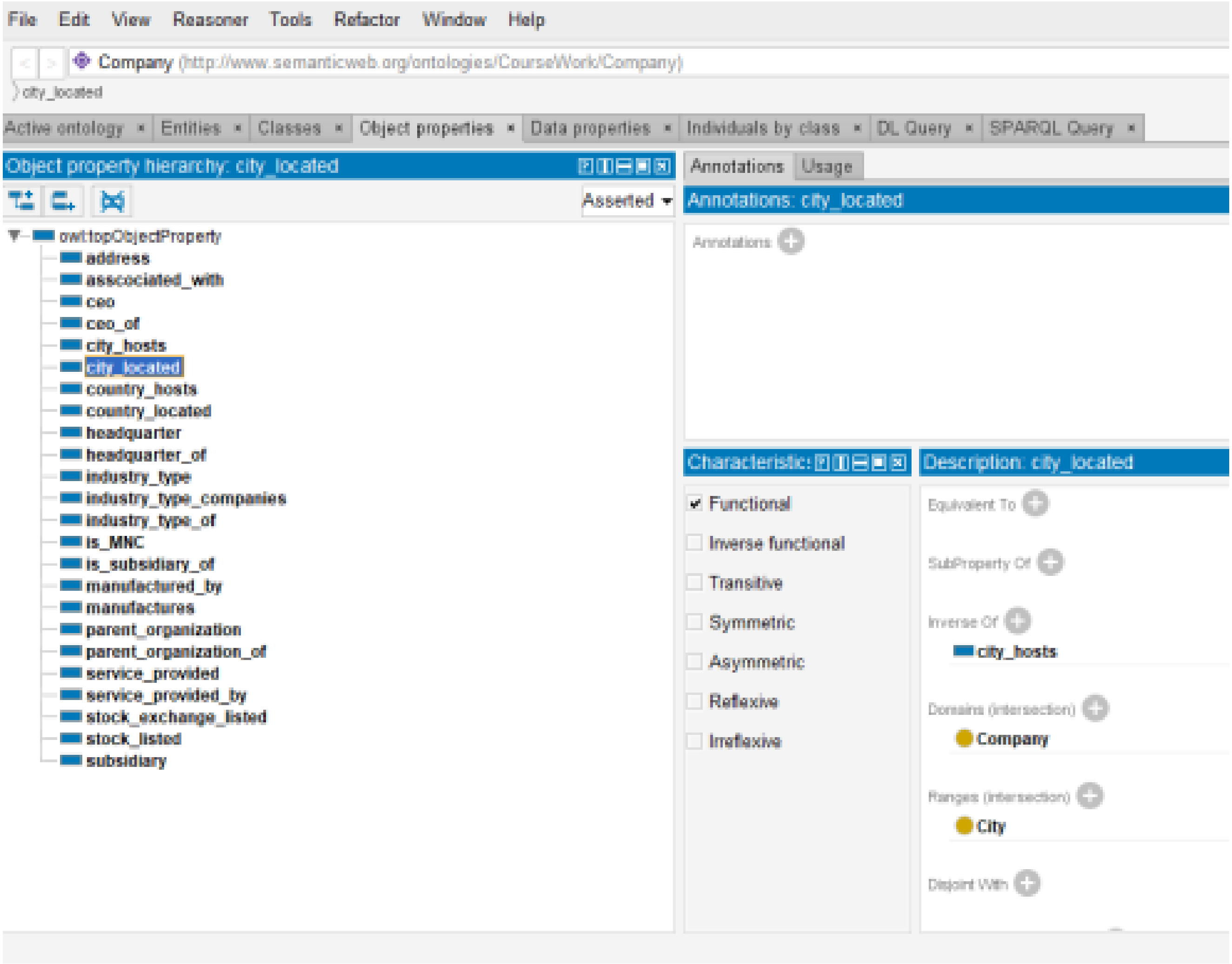
Company (http://www.semanticweb.org/ontologies/CourseWork/Company)

- owl:Thing
  - City
  - Company
  - Company_Type
    - **Business**
    - **Government**
    - **Private**
    - **Public**
  - Country
  - Department
  - Finacials
  - Industry
  - Industry_Type
  - Name
  - Parent_Organization
  - Person
    - **CEO**
    - **Intern**
    - **Manager**
    - **President**
    - **Staff**
  - Products
    - **Associated_Products**
    - **Main_Products**
  - Service
  - StockExchange
    - **Trading_Name**
  - Subsidiary

Picture 1 – Class hierarchy

- **Object properties**

Object properties are essential to relate different classes by linking them. For instance I have created object properies like ceo, city_located, county_located, headquarter, industry_type, manufactures, subsidiary, and reverse functional properties like city_hosts, county_hosts, manufactured_by, is_subsidiary_of etc. Picture depicts the object properties of the onltology.



owl:topObjectProperty
- address
- asscociated_with
- ceo
- ceo_of
- city_hosts
- city_located
- country_hosts
- country_located
- headquarter
- headquarter_of
- industry_type
- industry_type_companies
- industry_type_of
- is_MNC
- is_subsidiary_of
- manufactured_by
- manufactures
- parent_organization
- parent_organization_of
- service_provided
- service_provided_by
- stock_exchange_listed
- stock_listed
- subsidiary

☑ Functional
☐ Inverse functional
☐ Transitive
☐ Symmetric
☐ Asymmetric
☐ Reflexive
☐ Irreflexive

Equivalent To

SubProperty Of

Inverse Of
city_hosts

Domains (intersection)
Company

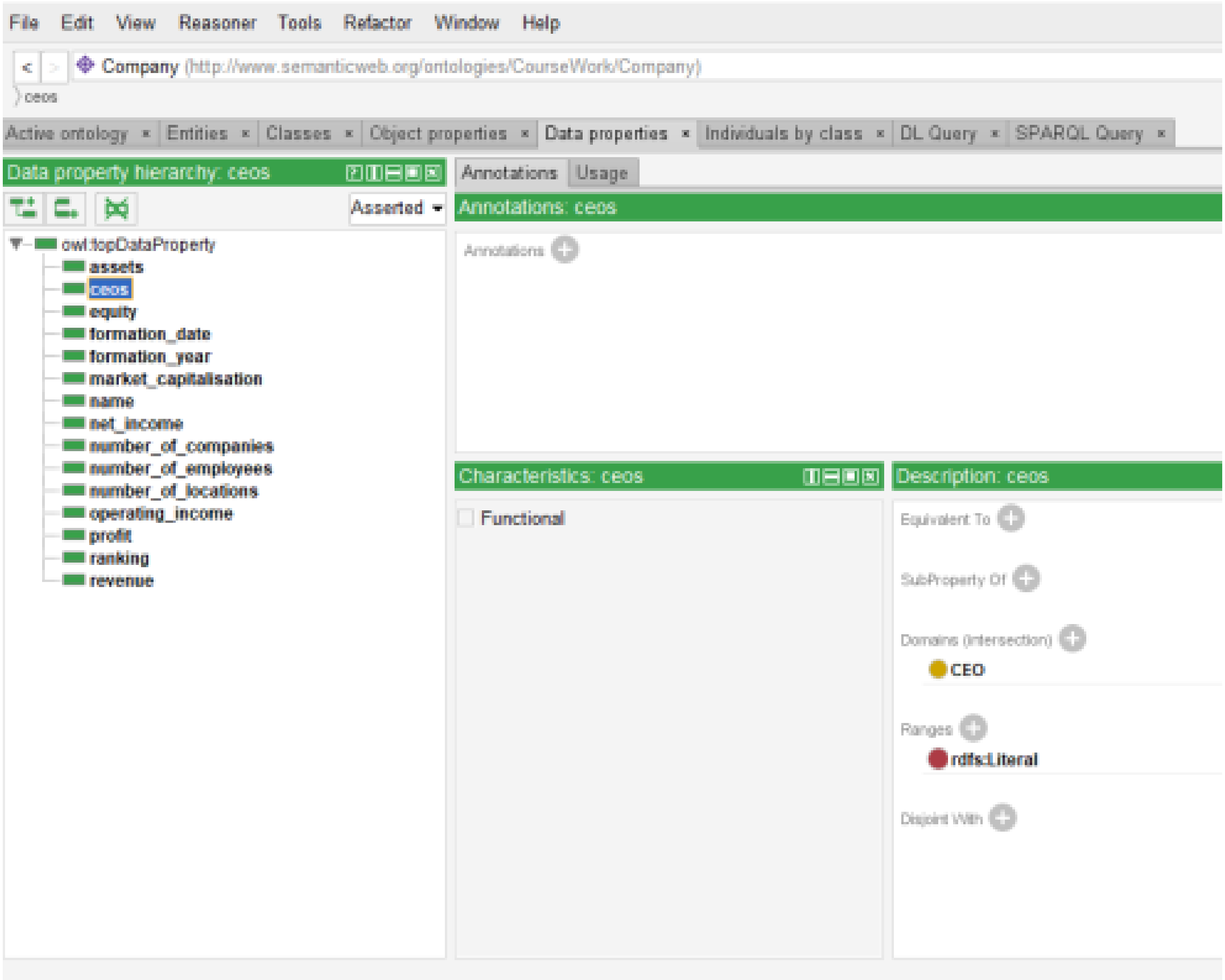
Ranges (intersection)
City

Disjoint With

Picture 2 – Object properties

- **Data properties**

Data properties are defined to hold the values for properties like profit, revenue, number of employees, net income, assets etc. Picture 3 illustrates the Data properties of the ontology.



Picture 3 – Data properties

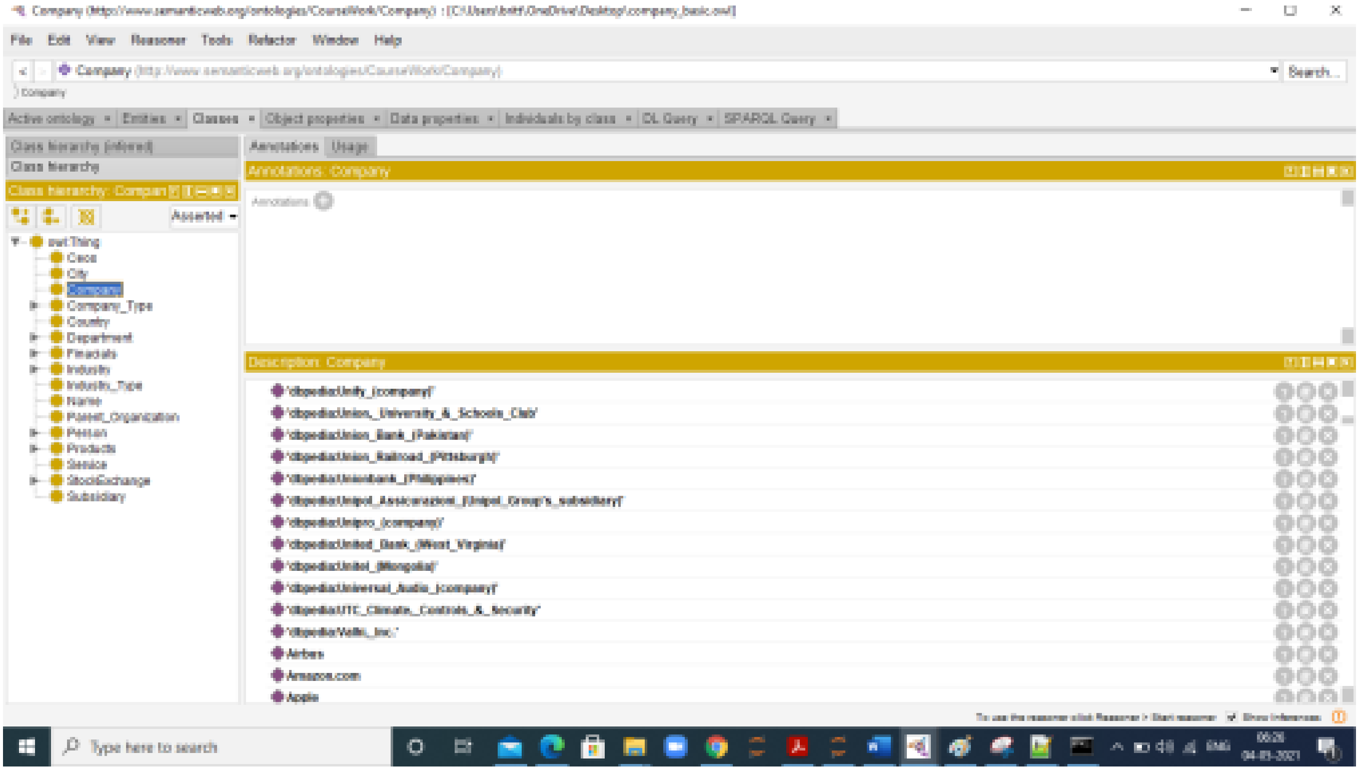
Data population into Ontology

Now that we have created the ontology, we need to populate the ontology with data. First to check the correctness of the ontology I manually input data for 25 Fortune 500 companies. I collected the data for the companies from various sources including Wikipedia.

Once the manual population is completed, I tested the correctness of the ontology by executing a few queries in SPARQL tab and I was convinced that the ontology is properly constructed and data are represented in a wholistic manner.

Once the correctness is verified, I loaded the data from DBpedia. For this I constructed SPARQL query and embedded it in a python script and executed. I faced a few of challenges like a data for a few classes did not load. I overcame by analyzing the data and object structure and making suitable correction in object hierarchy, adding few more classes to hold the data and checking the correctness of the characteristics of the object.

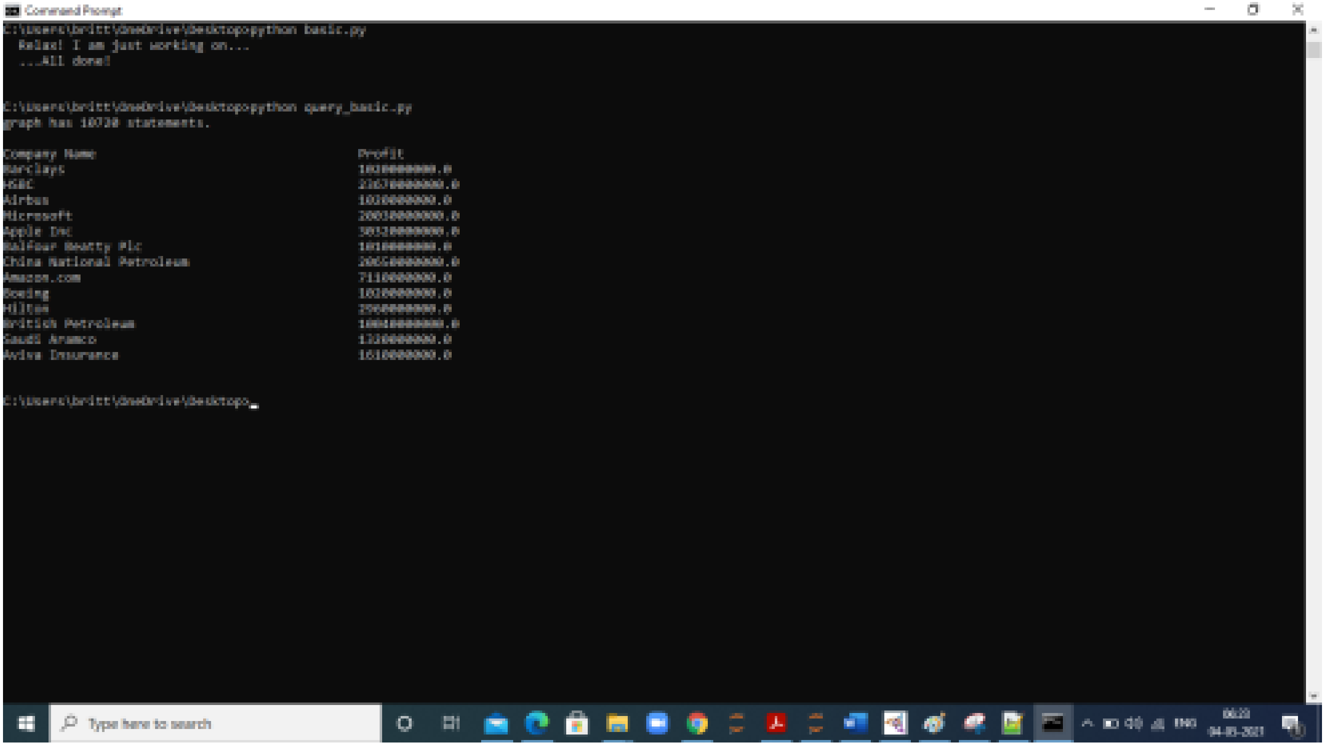
The below are some of the pictures depicting the ontology loaded with external data.



'dbpedia:Unify_(company)'
'dbpedia:Union_University_&_Schools_Club'
'dbpedia:Union_Bank_(Pakistan)'
'dbpedia:Union_Railroad_(Pittsburgh)'
'dbpedia:Unionbank_(Philippines)'
'dbpedia:Unipol_Assicurazioni_(Unipol_Group's_subsidiary)'
'dbpedia:Unipro_(company)'
'dbpedia:United_Bank_(West_Virginia)'
'dbpedia:Unitel_(Mongolia)'
'dbpedia:Universal_Audio_(company)'
'dbpedia:UTC_Climate_Controls_&_Security'
'dbpedia:Vallic_Inc.'
Airbus
Amazon.com
Apple

Querying the Ontology populated from DBpedia

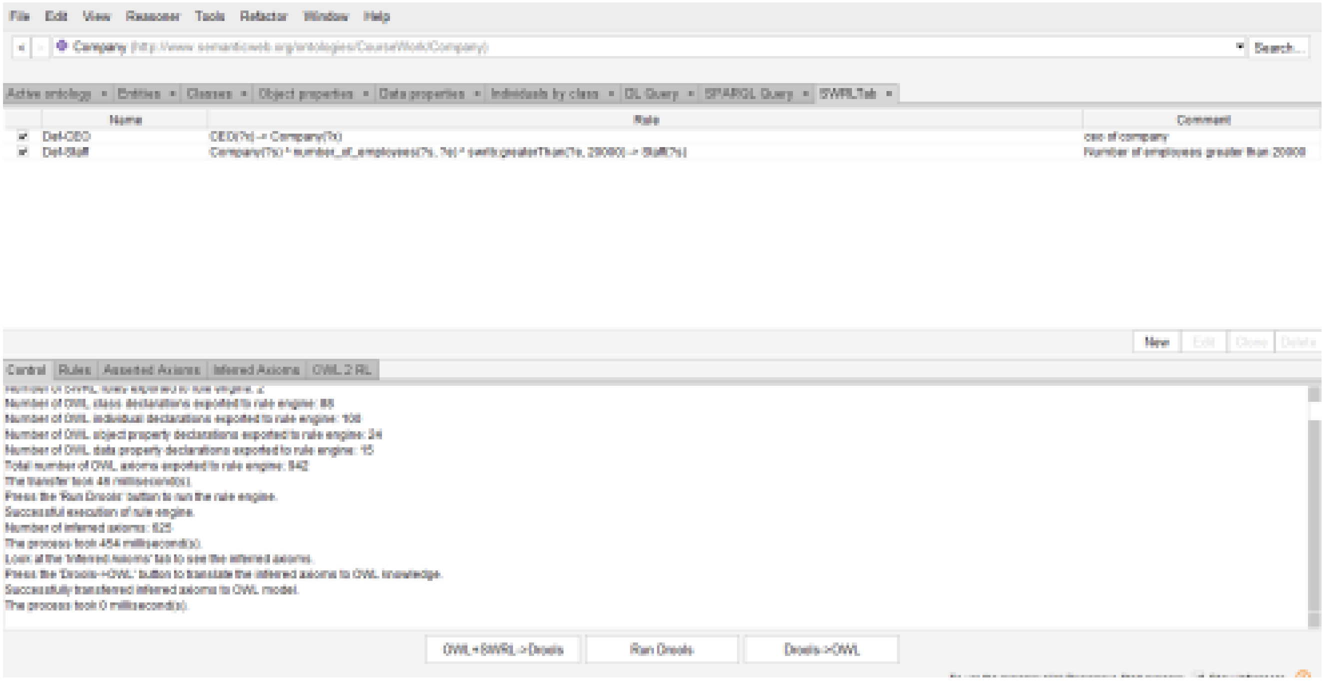
Now that data from DBpedia is loaded into the ontology. I used a simple query to fetch the name and profit of all the companies in the ontology. The query was executed using Python script. The below is the snapshot of the query result.



```
C:\Users\britt\OneDrive\Desktop>python basic.py
  Relax! I am just working on...
  ...All done!


C:\Users\britt\OneDrive\Desktop>python query_basic.py
graph has 10720 statements.

Company Name                                        Profit
Barclays                                            1020000000.0
HSBC                                                23670000000.0
Airbus                                              1020000000.0
Microsoft                                           20030000000.0
Apple Inc                                           30320000000.0
Balfour Beatty Plc                                  1010000000.0
China National Petroleum                            20650000000.0
Amazon.com                                          7110000000.0
Boeing                                              1020000000.0
Hilton                                              2960000000.0
British Petroleum                                   10040000000.0
Saudi Aramco                                        1320000000.0
Aviva Insurance                                     1610000000.0


C:\Users\britt\OneDrive\Desktop>_
```

SWRL Rule

The below SWRL rule has been set up and tested in Company ontology. The result output is as shown below.



	Name	Rule	Comment
☑	Def-CEO	CEO(?x) -> Company(?x)	ceo of company
☑	Def-Staff	Company(?x) ^ number_of_employees(?x, ?y) ^ swrlb:greaterThan(?y, 20000) -> Staff(?x)	Number of employees greater than 20000

```
Number of SWRL rules exported to rule engine: 2
Number of OWL class declarations exported to rule engine: 88
Number of OWL individual declarations exported to rule engine: 108
Number of OWL object property declarations exported to rule engine: 24
Number of OWL data property declarations exported to rule engine: 15
Total number of OWL axioms exported to rule engine: 842
The transfer took 48 millisecond(s).
Press the 'Run Drools' button to run the rule engine.
Successful execution of rule engine.
Number of inferred axioms: 625
The process took 454 millisecond(s).
Look at the 'Inferred Axioms' tab to see the inferred axioms.
Press the 'Drools->OWL' button to translate the inferred axioms to OWL knowledge.
Successfully transferred inferred axioms to OWL model.
The process took 0 millisecond(s).
```

List of Files

1. CompanyOntology.pdf – Project report

2. Company.owl – Ontology file
3. basic.py – Python script to load data from DBpedia
4. query_basic.py – Python script to execute basic query

Instructions to run the code

To load the ontology in Protégé, please select Company.owl file through Open file option.

Execute basic.py in command prompt to populate the ontology DBpedia data

Execute query_basic.py to test the ontolog

9

10