Ontology & Application Query Interface

STREETS & BINS AROUND DCC

Knowledge and Data Engineering

GROUP – C

Divyanshu Marwah – 19304526
Prateek Tulsyan – 19303677
Piyush Saxena – 19300369
Ankit Taparia – 19302722
Akash Verma – 19308805
Chen Wang - 19302124

APPROACH TO ONTOLOGY MODELLING

Description of competency question that ontology answers

Litter bins are an essential public infrastructure for modern life. Reasonably setting the location of litter bin is helpful for building a good living environment maintaining a clean and tidy cityscape. Urban planners can improve the utilization rate of resources by designing the location and distribution of litter bins reasonably with the limited hardware resources, thus achieving the goal of saving resources. Therefore, it is of great significance to establish an Ontology Modelling about the location of litter bins. The Ontology Modelling can help urban planners to understand the distribution of litter bins more intuitively and provide an accurate factual basis for planning and designing the location and distribution litter bins.

Description of datasets selected for application

Litter bins are an essential public infrastructure for modern life. Reasonably setting the location of litter bin is helpful for building a good living environment maintaining a clean and tidy cityscape. Urban planners can improve the utilization rate of resources by designing the location and distribution of litter bins reasonably with the limited hardware resources, thus achieving the goal of saving resources. Therefore, it is of great significance to establish an Ontology Modelling about the location of litter bins. The Ontology Modelling can help urban planners to understand the distribution of litter bins more intuitively and provide an accurate factual basis for planning and designing the location and distribution litter bins.

Assumptions

In "Roads and Streets in Dublin City" dataset:

- Replaced missing values in "year built" column by 0
- Replaced blank values in "route no" column by "missing"
- Replaced missing values in "line length" column by 0
- Replaced missing values in "street class code", "new area code", "surface type code" columns by 0
- Replaced '?' and blank values in "road start" column by "missing"
- Replaced blanks in "road finish" column by "missing"
- Replaced blanks in "Irish" column by "missing"

References to sources used/reused: (e.g. SIOC, FOAF for people)

PREFIX foaf:<http://xmlns.com/foaf/0.1/>
PREFIX rdf-schema:<http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf-syntax-ns:http:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>

Discussion of data mapping process

Considering the original dataset for streets and bins, there are a lot of references used such as street class coding, surface type coding, bin type, etc. We have used these references as separate classes in our ontology.
Another thing which we have done in mapping process is breaking down the columns into more closely coupled scenario. For instance, in the bins dataset, there are certain columns specifying the accessibility of the litter bins with columns such as school, footpath, residential, etc., so we have created a separate class for accessibility area. Same logic goes for zonal area.
The streets dataset has multiple columns which give us the features for the street, so we have created a separate class as street features containing the properties such as road start, road finish, line length, etc.
In the same way, bins dataset has certain column which give us the features of the bins, so we have created a separate class for bin features containing properties such as cigarette butt container, is damaged, etc.

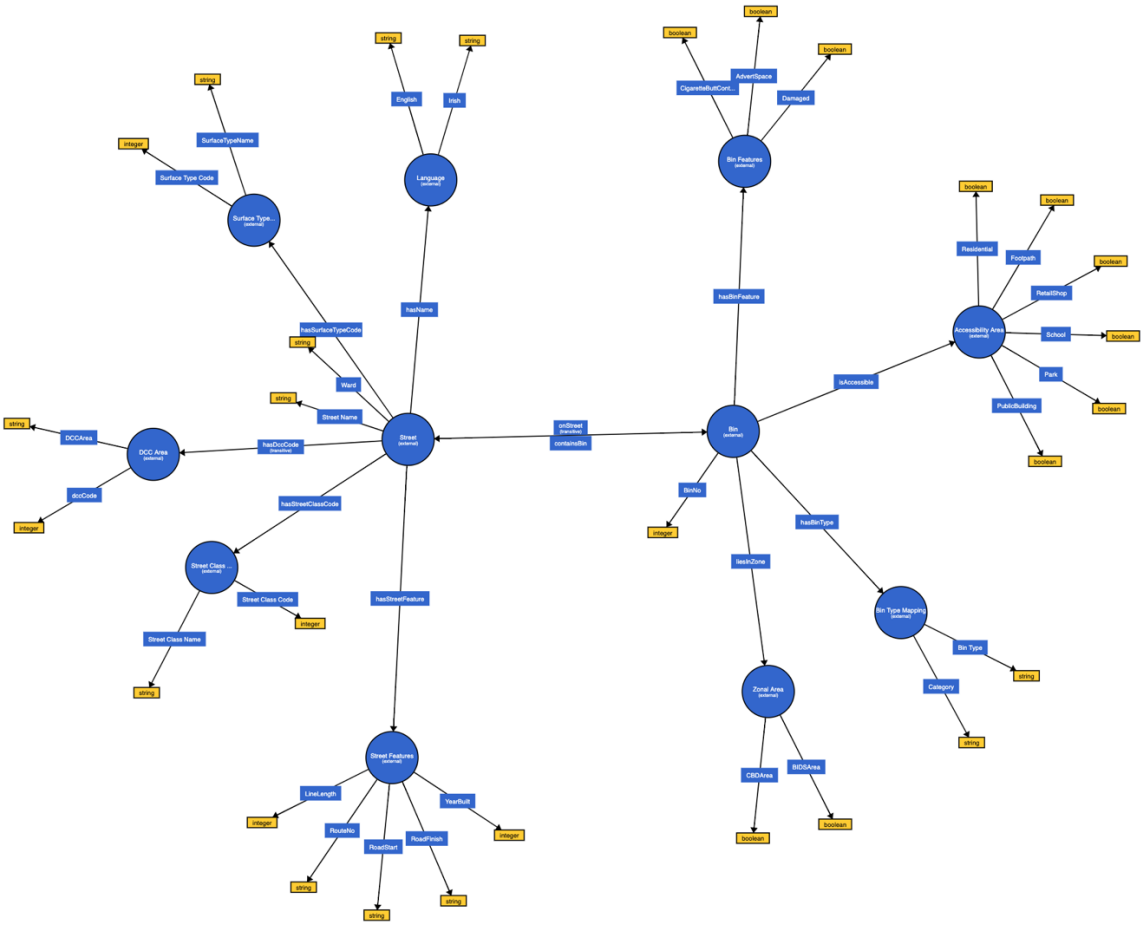
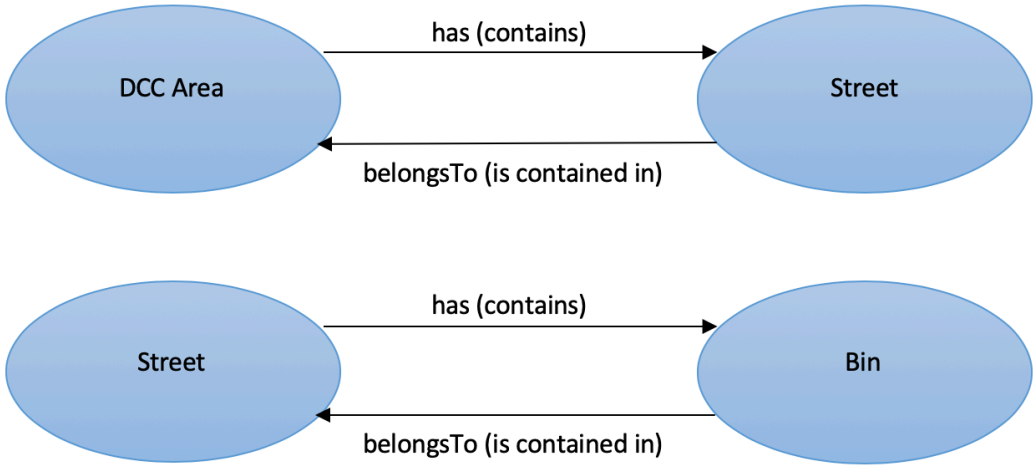


Fig 1: OWL Ontology Schema

Explanation of use of inverse, symmetric and transitive properties

- **Inverse property**
  Street and DCC Area holds inverse relationship, as a particular DCC area say X has many streets of which one of them is say Y. Therefore, we can say that X contains Y and Y is contained in X.
  Similarly, the above will hold for a Bin and a Street.



- **Transitive Property**
  Entities - Bin, Street and DCC Area have transitive property. A bin is located on a street. A street belongs to a DCC Area. Hence, we can say that a bin belongs to or is present in a DCC area.

OVERVIEW OF DESIGN

Description of Application Query Interface

Tools Used

- Virtuoso TripleStore
- Python with Flask
- Bootstrap3
- Query Editor using YASGUI library
- Query Formator using CodeMirror library

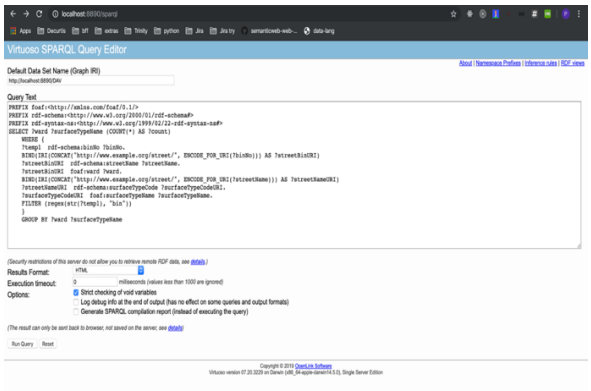


Fig 2: Virtual Triple Store

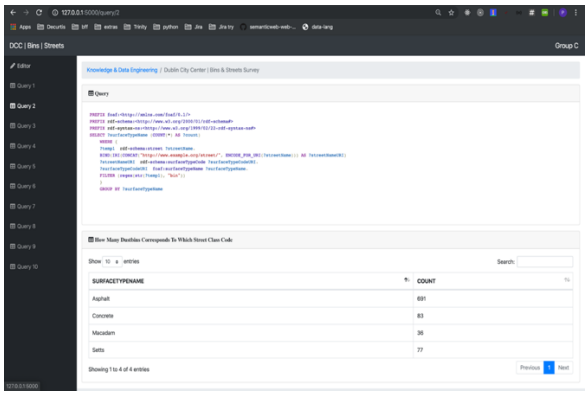


Fig 3: 10 Default Queries

The application query interface contains 10 Default Queries and a Query Editor where a query can be entered and run on the go.

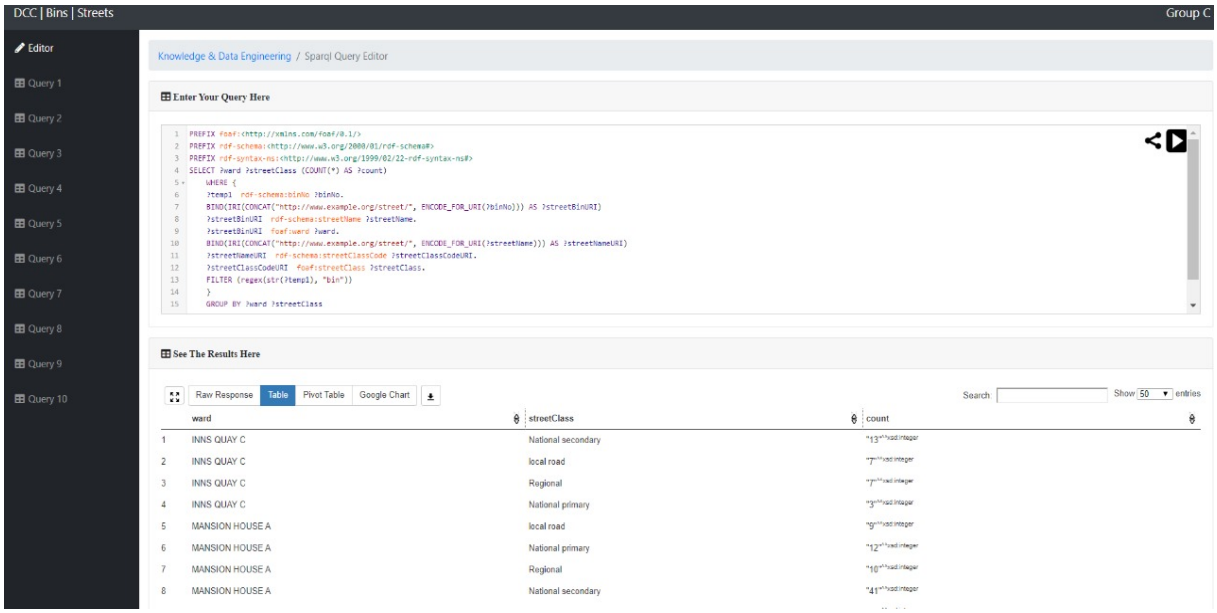


Fig 4: Query Editor to write and execute query on the fly

Description of Queries

- How many bins are present on each street class type (local road, national primary, etc.)?
- Find the count of each street surface type (concrete, macadam, etc.) for every District Electoral Division (Ward-DED)
- Find the count of each street type for every District Electoral Division (Ward-DED)
- How many roads were built in each Dublin City Council (DCC) area over past 5 years?
- Where the damaged bins are located along with the street type of the corresponding street?
- How many bins are available for advertisement on a particular street surface type?
- Calculate the bin density on each street. Formula used = No of bins/ Line Length of the street
- How many bins are present in each DCC area?
- Which city center streets were built in and before 2000 and compute the total number of bins present on them?
- Where are the damaged bins located along with the surface type of the corresponding street?

Challenges faced while ontology modelling or creating queries and mappings

- It took time to find datasets which were relatable with each other, could create at least 10 classes and could provide 40 different instances.
- Finding the inverse, transitive and symmetric characteristics in ontology.
- Creating mapping between classes and properties. As sometimes, properties might seem like a subclass.

CONCLUSION

Strengths of ontology model, queries & interface

- Diverse dataset - Datasets selected are very informative and combined datasets give us huge details about the data. For instance, streets dataset gives us the information about the class coding of the roads/streets in Dublin and further many such features of the streets. When combined with bins dataset, it gives us a diverse data, like the types of bins on which types of streets, how many bins in which zonal area etc.
- Responsive interface

Weaknesses of ontology model, queries & interface

- Our application uses local SPARQL endpoint created by Virtuoso (Open source). Therefore, to run the application, Virtuoso should be installed on the system.

Code Reference: https://github.com/tulsyanp/linked-data-streets-bins-survey