EireBnB - An ontology-driven application for tourism in Ireland

Claire Farrell
School of Computer Science & Statistics,
Trinity College Dublin,
Dublin, Ireland
farrec26@tcd.ie or 16319148

Joshua Cassidy
School of Computer Science & Statistics,
Trinity College Dublin,
Dublin, Ireland
cassidj3@tcd.ie or 20300057

Yifan Pei
School of Computer Science & Statistics,
Trinity College Dublin,
Dublin, Ireland
peiy@tcd.ie or 20303063

Siddhesh Kankekar
School of Computer Science & Statistics,
Trinity College Dublin,
Dublin, Ireland
kankekas@tcd.ie or 20300810

Declan O'Sullivan
School of Computer Science & Statistics,
Trinity College Dublin,
Dublin, Ireland
declan.osullivan@scss.tcd.ie

Albert Navarro
School of Computer Science & Statistics,
Trinity College Dublin,
Dublin, Ireland
navarral@tcd.ie

I. INTRODUCTION

The purpose of this project is to create the ultimate ontological application for Tourists seeking alternative accommodation and activities in Ireland. We identified this problem statement as we had personally encountered issues choosing accommodation in Ireland during our staycations. As hotels in Ireland had become overpriced due to popular demand, younger, more price-sensitive generations searched for alternative accommodations. The problem then continued as we struggled to identify Airbnb's in areas that would have suitable activities. The premise behind this project is to combine datasets that answer complex questions that are unanswerable by querying a single dataset alone, i.e. will find accommodation and activities in a given area. The first dataset utilised contained data points about listings on Airbnb in Ireland; various attributes include location, price, and room type. The second dataset obtained was from Fáilte Ireland, called Activities, and contains information about the various activities in Ireland. Following sections will delve deeper into the dataset chosen. [1] [2]

II. ONTOLOGY MODELLING APPROACH

A. Ontology Competency Questions

Each of these questions were designed to highlight the benefits and capabilities of our ontology model given we have combined various pre-processed datasets.

1. What room types are available where there are 'Fishing' activities?

This question will allow the user to understand what type of rooms are available in an area of fishing activities, e.g. there are entire apartments and single rooms in the area of Alpine Fishery. Cleaned datasets used: Activities, Accomodations, Counties, RoomType

2. What 'Counties' can you go 'Fishing' in and stay in an apartment?

This question allows the user to narrow their search as they have two preferences: fishing activities and apartment room type. The question should return the counties that satisfy those requirements.

Cleaned datasets used: Activities, Accomodations, Counties, RoomType

3. What 'accommodations' can you stay in to go 'Fishing' that have more than '3 Beds'?

This competency question demonstrates where a user now has a separate set of requirements where they can go fishing and need more than three beds. The question returns what kind of accommodations are available, e.g. house, apartment, hotel.

Cleaned datasets used: Activities, Accomodations, Counties, PropertyType

4. Where can someone stay to go 'Fishing' next summer on '2021-01-18'?

This question defines when a user wants to find somewhere to stay in order to go fishing on a specific date. This returns the name of the accommodation.

Cleaned datasets used: Activities, Accomodations, Counties, Calendar

5. What are the amenities offered in accommodations for someone that wishes to go golfing?

This question allows a user who wants to go golfing and wants to gain understanding of the types of amenities e.g. wifi, soap, extra pillows, will be available to them. This will return the name of the activity in the golf category and the types of amenities available in the vicinity.

Cleaned datasets used: Activities, Accomodations, Counties, Amenities

6. What 'activities' can you do in an area where the host has a 'communication rating' of 9?

The question addresses when some users may find good communication with the host important. This question shows the various activities available in an area where there is accomodation with hosts with great communication ratings.

Cleaned datasets used: Activities, Accomodations, Counties, Host

7. What properties does hostid 234243 have in the same county as the Shopping Activity?

This competency question demonstrates that given a specific host id and the interest of Shopping, this question returns the various properties owned by the host in the same

county that has Shopping Activities e.g. shopping centres, outlets.

Cleaned datasets used: Activities, Accomodations, Counties, Host, PropertyType

8. Where can someone stay to go on a fishing holiday for under €250?

This question allows a user to understand that when given a daily rate budget and the interest of going on a fishing trip, the user can see results of the various accommodation options that fit the criteria.

Cleaned datasets used: Activities, Accomodations, Counties, Calendar

9. What accommodations have available dates to visit "Ardgillan Demesne" where the host has an acceptance rate above 70%?

This question defines the situation where a user would like to visit the Ardgillan Demesne but needs to know which dates are available and where the host has a high bookings acceptance rate. This question should return the accommodations and respective dates they are available in the area.

Cleaned datasets used: Activities, Accomodations, Counties, Host

10. Where can someone go on a fishing holiday for at least 3 nights?

This competency question demonstrates usability where the user would like to find accommodation where they are able to go on a fishing trip where they would be able to stay in their accommodation for at least three nights.

Cleaned datasets used: Activities, Accomodations, Counties, RoomType, Calendar

B. Datasets Used

As we knew we wanted to create an application to aid tourism in Ireland, we understood that there were two fundamental topics that comprised a typical tourist's time: accommodation and activities. Therefore, we identified two open-source datasets for our application.

1. Activities Dataset

The original dataset was obtained from https://data.gov.ie/dataset/activities which was last updated in October 2020. It contained Name, URL, Telephone, Longitude, Latitude, AddressRegion, AddressCountry, and, Tags. There is a continuous frequency of updates. [1]

2. Listings Dataset

This dataset was obtained from http://insideairbnb.com/get-the-data.html. InsideAirbnb is an open-source toolkit and data platform for users to understand how Airbnb is being used in their cities. The listings CSV file contains id, name, host_id, host_name, neighbourhood_group, neighbourhood, latitude, longitude, room_type, price, minimum_nights, number_of_reviews, last_review, reviews_per_month, calculated_host_listings_count, availability_365. [2]

These datasets were then cleaned so that any irrelevant columns for our application were dropped. These cleaned CSV files were then separated into various files for convenience. These files include: Accomedation.csv, AccomedationRating.csv, Activities.csv, ActivitiesJoin.csv, ActivitiesTypes.csv, Amenity.csv, AmenityJoin.csv, Calendar.csv, Counties.csv, Host.csv, PropertyType.csv, Reviews.csv, and, RoomType.csv. These datasets were cleaned to help with the R2RML mapping process and because some columns such as amenities and activities were poorly structured and had many values for the one column meaning that it would be significantly more challenging to uplift the raw downloaded data. This can be visualised in our UML diagram of our ontology containing the various classes and their respective properties.

III. ASSUMPTIONS MADE

Due to the size of our dataset, the project team was not able to load the entire dataset into the triple store using our laptops thus only a sample of the datasets were loaded into the store. The assumptions that were made for in the development of this ontology are:
1. The prices for all locations are listed in dollars.
2. The data retrieved is up to date.
3. The activities listed in the activities dataset are still open for business.
4. The airbnb listings are still available for stay in.
5. There are no two counties with the same name in one country.
6. One activity can only happen in one location.

IV. SOURCES USED/REUSED

No external data sources were reused although we attempted to use DBPedia to describe the types of activities, however, it was decided not to use these descriptions for the activities as DBPedia pages did not exist for some activities in our dataset an example of this there is no DBPedia page describing Pitch and Putt. [4]

The datasets that were used for the purposes of this project were the activities dataset which contains information such as the name, URL, telephone number of activities to do in Ireland and the AirBnB listings dataset which contains information such as the host name, price, availability and locations of AirBnB's in Ireland. [1] [2]

V. DATA MAPPING PROCESS

A. Initial Design

The initial design process followed an in-depth survey of the original downloaded CSV files and decided as a group which properties would have value in our graph.
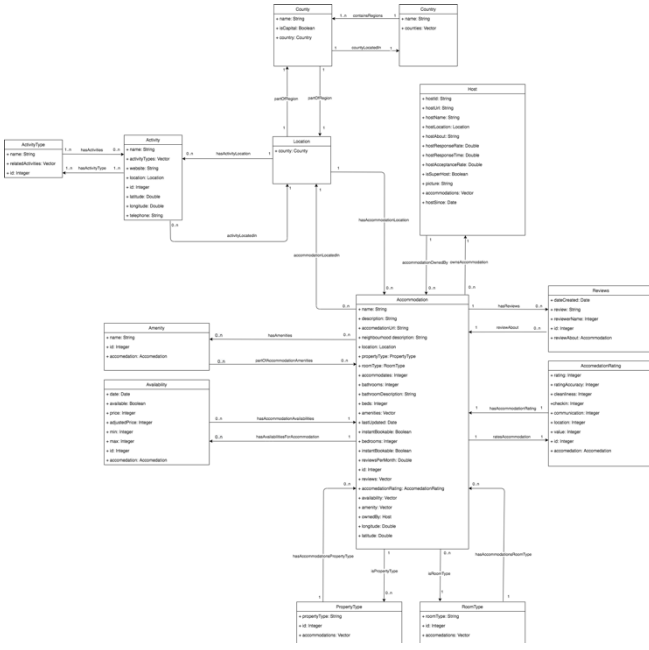
Figure 1: UML class diagram representing the ontology model

Figure 1 shows the UML class diagram that represents the ontology model, from this we can see how each of the Classes relate to each other and their respective cardinalities, with the relation being stored as a property in each class for clarity of the type of the relation. A separate png file containing figure 1 can be found in our submission zip folder for convenience.

B. Mapping Process

Once the UML class diagram representing the ontology was created, Protege was used in building the ontology structure, storing the ontology structure in a turtle file. All of the information that was modelled in the UML class diagram was reflected in the protege modelling process, this included the data properties and their types, the classes, the classes relationship and the classes cardinalities and the type constraints between the classes.
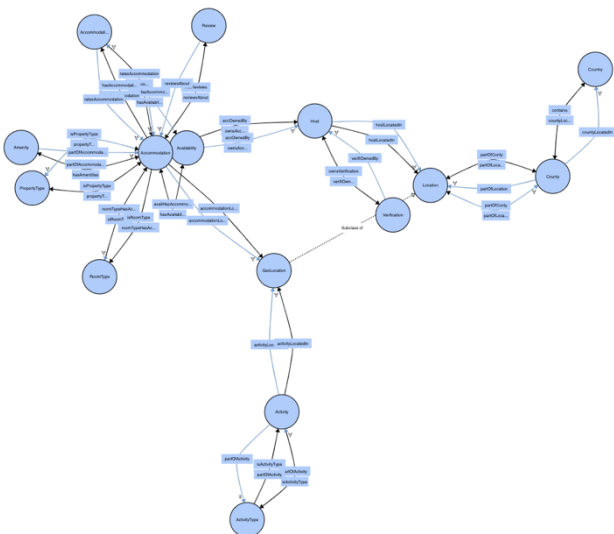

Figure 2: Representation of the created ontology structure

Figure 2 shows a representation of the modeled ontology showing the representation of the ontology structure showing the classes and the relations between them.

A separate png file containing figure 2 can be found in our submission zip folder for convenience.

C. R2RML

The R2RML engine that was provided in this module was used to generate the mappings between the ontology model and the CSV data files to populate the ontology. Due to the amount of data and the project teams limited access to hardware only a sample of the data was inserted into the ontology. The R2RML process stored the ontologies data in a file called "eirebnb_ontology.ttl", the data in this file was inserted into an Apache Jena Fuseki triple store.

VI. USE OF INVERSE, SYMMETRIC AND TRANSITIVE PROPERTIES

A. Inverse Properties

Relationships that are described as inverse properties illustrate that if there is a relationship between a and b, there will also be an inverse relationship between b and a given that the relationship between b is the mathematical opposite to a. The following relationships are examples of the inverse functional properties:

Accommodation and Location
An accommodation is located in a location denoted by "accommodationLocatedIn" in the ontology, so the inverse relationship of this in the ontology is "hasAccommodationLocation" which means that a location has accommodations

Host and Accommodation
An accommodation is owned by a host this relationship is represented by "accommodationOwnedBy" in the knowledge graph, the inverse of this relationship is a host owns accommodations this is represented by "ownsAccommodation" in the ontology

Activity and Location
An activity is located in a location this relationship is represented by "activityLocatedIn" in the knowledge graph, the inverse of this relationship location contains activities this is represented by "hasActivityLocation" in the ontology

Country and County
An country contains many counties this relationship is represented by "containsRegions" in the knowledge graph, the inverse of this relationship county is located in a country this is represented by "countyLocatedIn" in the ontology

Accommodation and Availability
An availability has an accommodation this relationship is represented by "hasAccommodationAvailabilities" in the knowledge graph, the inverse of this relationship an accommodation has availabilities this is represented by "hasActivityAccommodation" in the ontology.

Accommodation and AccommodationRating

An Accommodation has ratings for the accommodation this relationship is represented by "hasAccommodationRating" in the knowledge graph, the inverse of this relationship is a rating has an accommodation that is represented by "ratesAccommodation" in the ontology.

Accommodation and PropertyType
A property type has accommodations this relationship is represented by "hasAccommodationsPropertyType" in the knowledge graph, the inverse of this relationship is an accommodation has a property type that is represented by "isPropertyType" in the ontology.

Accommodation and RoomType
A room type has accommodations this relationship is represented by "hasAccommodationsRoomType" in the knowledge graph, the inverse of this relationship is an accommodation has a room type that is represented by "isRoomType" in the ontology.

Activity and ActivityType
An activity type has activities this relationship is represented by "hasActivities" in the knowledge graph, the inverse of this relationship is an activity has activity types that is represented by "hasActivityType" in the ontology.

Activity and ActivityType
An activity type has activities this relationship is represented by "hasAmenities" in the knowledge graph, the inverse of this relationship is an activity has activity types that is represented by "partOfAccommodationAmenities" in the ontology.

Amenity and Accommodation
Accommodation have amenities this relationship is represented by "hasAmenities" in the knowledge graph, the inverse of this relationship is amenities have accommodations that is represented by "partOfAccommodationAmenities" in the ontology.

Amenity and Accommodation
Accommodations have amenities this relationship is represented by "hasAmenities" in the knowledge graph, the inverse of this relationship is amenities have accommodations that is represented by "partOfAccommodationAmenities" in the ontology.

Reviews and Accommodation
An accommodation has reviews this relationship is represented by "hasReviews" in the knowledge graph, the inverse of this relationship is a review has an accommodation that is represented by "reviewAbout" in the ontology.

B. Symmetric Properties

A symmetric property is identified when the relationship between a and b is the exact same relationship between b and a.

County and Location
The relationship between county and location is modelled to be symmetric using the partOfRegion object property the reason that this relationship is modelled to be symmetric is that a county is a part of a location and a location is part of a county.

C. Transitive Properties

A transitive property occurs when the relationship between a and b is also apparent between b and c. We can say that a relationship is transitive between a and c via the property b.

Host and AccommodationRating
The host and the AccommodationRating are connected with a transitive property with the relationship being denoted by the "hostAccommedationRating" property in the ontology.

Host and Amenity
The Host and the Amenity are connected with a transitive property with the relationship being denoted by the "hostAmenities" property in the ontology.

Host and Availability
The Host and the Availability are connected with a transitive property with the relationship being denoted by the "hostAvailabilties" property in the ontology.

Host and PropertyType
The Host and the PropertyType are connected with a transitive property with the relationship being denoted by the "hostPropertyType" property in the ontology.

Host and Review
The Host and the Review are connected with a transitive property with the relationship being denoted by the "hostReviews" property in the ontology.

Host and RoomType
The Host and the RoomType are connected with a transitive property with the relationship being denoted by the "hostRoomType" property in the ontology.

Accommodation Activity
The Accommodation and the Accommodation are connected with a transitive property with the relationship being denoted by the "hasAccommodationActivities" property in the ontology.

Accommodation and ActivityType
The Accommodation and the ActivityType are connected with a transitive property with the relationship being denoted by the "hasAccommodationActivityTypes" property in the ontology.

ActivityType and Location
The ActivityType and the Location are connected with a transitive property with the relationship being denoted by the "activityTypeLocations" property in the ontology.

Activity and Accommodation
The Activity and the Accommodation are connected with a transitive property with the relationship being denoted by the "hasActivityAccommodation" property in the ontology.

Activity and Accommodation

The Activity and the Accommodation are connected with a transitive property with the relationship being denoted by the "hasActivityAmenities" property in the ontology.

Activity and Availability
The Activity and the Availability are connected with a transitive property with the relationship being denoted by the "hasActivityAvailability" property in the ontology.

Activity and RoomType
The Activity and the RoomType are connected with a transitive property with the relationship being denoted by the "hasActivityRoomType" property in the ontology.

ActivityType and Accommodation
The ActivityType and the Accommodation are connected with a transitive property with the relationship being denoted by the "hasActivityTypeAccommodation" property in the ontology.

ActivityType and Amenity
The ActivityType and the Amenity are connected with a transitive property with the relationship being denoted by the "hasActivityTypeAmenity" property in the ontology.

ActivityType and Availability
The ActivityType and the Availability are connected with a transitive property with the relationship being denoted by the "hasActivityTypeAvailabilies" property in the ontology.

ActivityType and PropertyType
The ActivityType and the PropertyType are connected with a transitive property with the relationship being denoted by the "hasActivityTypePropertyType" property in the ontology.

ActivityType and RoomType
The ActivityType and the RoomType are connected with a transitive property with the relationship being denoted by the "hasActivityTypeRoomType" property in the ontology.

VII. OVERVIEW OF DESIGN

A. Application Query Interface

As the ontology and output file had been completed, the next phase of our project was to create a user interface so that any user could query our model.

Figure 3: UI for competency Question 1: "What room types are available where there are fishing activities?"

Figure 3 shows the UI for the competency question 1 we can see from in the answer to this question that the room type and the location of where the holiday goer can stay is listed.

Figure 4 : UI for competency Question 2: "What counties can you go fishing in and stay in an apartment?"

Figure 4 shows the UI for the competency question 2 we can see from in the answer to this query lists the county in which the user can stay in.

Figure 5: UI for competency Question 3: "What accommodations can you stay in to go fishing that have more than 3 beds?"

Figure 5 shows the UI for the competency question 3 we can see from in the answer to this query lists activity location and the property type that the user can stay in.

Figure 6: UI for competency Question 4: "Where can someone go fishing next summer on 2021-01-18"

Figure 6 shows the UI for the competency question 4 we can see from in the answer to this query lists name of the accommodation, the url of where the accommodation can be accessed, the name of the fishing activity and the location of where the user can stay.

Competency Question 5: Accommodation Amenities Activity

What are the amenities offered in accomodations for someone that wishes to visit Golf activity?

Activity Location	Property Type	Activity Location	Accomodation Id
Tigi shampoo	Ballybofey and Stranorlar Golf Club	Donegal	http://foo.example/Accommodation/23750298

Figure 7: UI for competency Question 5: "What are the amenities offered in accommodations for someone that wishes to go on a golfing holiday?"

Figure 7 shows the UI for the competency question 5 we can see from in the answer to this query lists activity location, the amenity and the property type that the user can stay in.

Competency Question 6: Activities Rating Communication

What activities can you do that have an communication rating of 9?

Activities	Communication Rating
Gardens	9

Figure 8: UI for competency Question 6: "What activities can you do that have a communication rating of 9?"

Figure 8 shows the UI for the competency question 5 we can see from in the answer to this query lists activity location, the amenity and the property type that the user can stay in.

Competency Question 7: Activity Host

What properties does host 234243 have in the same county as the Shopping Activity?

Accomadation Name	Host Name	Location	Activity
City Centre Spacious 1 Bed Apartment - sleeps 4	Johnny	Dublin	Shopping

Figure 9: UI for competency Question 7: "What properties do host 234243 have in the same county as a shopping activity"

Figure 9 shows the UI for the competency question 7 we can see from in the answer to this query lists accommodation name, the host name, the location and the activity.

Competency Question 8: Activity Type Price

Where can someone stay to go on a fishing holiday for under 250 euro?

Accomadation Name	Availability Date	Activity
Self-contained ensuite bedroom with private access	2021-04-06	Angling Bohernabreena Reservoirs

Figure 10: UI for competency Question 8: "Where can someone stay to go on a fishing activity for under 250 euro?"

Figure 10 shows the UI for the competency question 8 we can see from in the answer to this query lists accommodation name, the date the accommodation is available and the activity.

Competency Question 9: Activity Acceptance

What are the accommodations are available dates to visit "Ardgillan Demesne" where the host has an acceptance rate of 7?

Host Name	Acceptance Rate	Room Type	Activity Location
Ronan	89	Shared room	http://foo.example/Location/Dublin

Figure 11: UI for competency Question 9: "What accommodations are available dates to visit "Ardgillan Demesne" where the host has an acceptance rate of 7?"

Figure 11 shows the UI for the competency question 9 we can see from in the answer to this query lists host name, the acceptance rate, the room type and the activity location.

Competency Question 10: Activity Days

Where can someone go on a fishing holiday for atleast 3 nights?

Accomadation Name	Location Name	Activity	Availability Date	Days Available
Self-contained ensuite bedroom with private access	Dublin	Angling Bohernabreena Reservoirs	2021-06-21	3

Figure 12: UI for competency Question 10: "Where can someone go on a fishing holiday for at least 3 nights?"

Figure 12 shows the UI for the competency question 10 we can see from in the answer to this query lists accommodation name, the location name, the activity, the availability date and the days the accommodation is available.

Search the EireBnB Ontology

Query

SUBMIT QUERY ➤

roomType	activityLocation
Shared room	Dublin

Figure 13: UI for creating custom queries

Figure 13 shows the that allows users to create their own queries to query the ontology.

B. Queries

The context behind the following queries is clearly defined in the explanations of the questions in a previous section. The purpose of this section is to highlight the design of the query design and the different properties returned.

#Q1. What room types are available where there are fishing activities?
PREFIX
HostURL:<http://www.semanticweb.org/ontologies/2020/10/Host>
PREFIX Activity:
<http://www.semanticweb.org/ontologies/2020/10/Activity>
PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX rdfs:<http://www.w3.org/2007/01/rdf-schema#>
PREFIX xsd:<http://www.w3.org/2001/XMLSchema#>
SELECT ?roomType ?activityLocation {
?location
<http://www.semanticweb.org/ontologies/2020/10/activityLocation> ?activityLocation .
?accomLocation
<http://www.semanticweb.org/ontologies/2020/10/accommodationLocatedIn> ?accomodationLocation .
?room
<http://www.semanticweb.org/ontologies/2020/10/isRoomType> ?roomTypeId .
?roomTypeId
<http://www.semanticweb.org/ontologies/2020/10/roomType> ?roomType .
?activityId
<http://www.semanticweb.org/ontologies/2020/10/hasActivityType> <http://foo.example/ActivitiesTypes/Fishing> .
?activityId
<http://www.semanticweb.org/ontologies/2020/10/activityName> ?activity .
} GROUP BY ?roomType ?activityLocation

This query relates to the first query. This is beneficial to the user as it returns the type of rooms that are available to a user where they also wish to go fishing.

#Q2. What counties can you go fishing in and stay in an apartment?
PREFIX
MostActivity:<http://www.semanticweb.org/ontologies/2020/10/Activity>
PREFIX County:
<http://www.semanticweb.org/ontologies/2020/10/Location>
PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX rdfs:<http://www.w3.org/2007/01/rdf-schema#>
PREFIX xsd:<http://www.w3.org/2001/XMLSchema#>
SELECT ?activityLocation WHERE{
?activityType
<http://www.semanticweb.org/ontologies/2020/10/activityTypeName> ?Activities .
?location
<http://www.semanticweb.org/ontologies/2020/10/activityLocation> ?activityLocation .

?name
<http://www.semanticweb.org/ontologies/2020/10/propertyType> ?propertyType .
FILTER (?Activities = "Fishing" && ?propertyType = "Entire condominium")
} GROUP BY ?activityLocation

This query uses the group by feature as we want to be able to return unique combinations of the properties. This gives the user information about the counties that are available to stay in given that they want to go fishing and stay in an apartment.

#Q3. What accomodations can you stay in to go fishing that have more than 3 beds?
PREFIX
MostActivity:<http://www.semanticweb.org/ontologies/2020/10/Activity>
PREFIX County:
<http://www.semanticweb.org/ontologies/2020/10/Location>
PREFIX Accommodation:
<http://www.semanticweb.org/ontologies/2020/10/Accomedation>
PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX rdfs:<http://www.w3.org/2007/01/rdf-schema#>
PREFIX xsd:<http://www.w3.org/2001/XMLSchema#>
SELECT ?activityLocation ?propertyType WHERE{
?activityType
<http://www.semanticweb.org/ontologies/2020/10/activityTypeName> ?Activities .
?location
<http://www.semanticweb.org/ontologies/2020/10/activityLocation> ?activityLocation .
?name
<http://www.semanticweb.org/ontologies/2020/10/propertyType> ?propertyType .
?type
<http://www.semanticweb.org/ontologies/2020/10/accommodationBeds> ?Beds
FILTER (?Activities = "Fishing" && ?Beds >= 3)
} GROUP BY ?activityLocation ?propertyType ORDER BY ?activityLocation

Similarly, this query also utilises the group by function to return a unique combination of activities, activityLocation, propertyType and beds. This allows us to query the graph to find the various accommodations with more than three beds in locations where the user can also go fishing.

#Q4. Where can someone stay to go fishing next summer on 2021-01-18?
PREFIX
HostURL:<http://www.semanticweb.org/ontologies/2020/10/Host>
PREFIX Activity:
<http://www.semanticweb.org/ontologies/2020/10/Activity>
PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX rdfs:<http://www.w3.org/2007/01/rdf-schema#>
PREFIX xsd:<http://www.w3.org/2001/XMLSchema#>

```
SELECT DISTINCT ?accommodationId
?accomadationName ?accomodationURL ?activity
?locationPredicate {
    ?locationObject
<http://www.semanticweb.org/ontologies/2020/10/activityL
ocation> ?locationPredicate.
    ?availabilityDate
<http://www.semanticweb.org/ontologies/2020/10/availabili
tyDate> "2021-07-18" .
    ?availableListing
<http://www.semanticweb.org/ontologies/2020/10/hasAcco
mmodationAvailabilities> ?availabilityDate .
    ?availableListing
<http://www.semanticweb.org/ontologies/2020/10/accomm
odationName> ?accomadationName .
    ?availableListing
<http://www.semanticweb.org/ontologies/2020/10/accomm
odationLocatedIn> ?accomadationLocation .
    ?availableListing
<http://www.semanticweb.org/ontologies/2020/10/accomm
odationId> ?accommodationId .
    ?availableListing
<http://www.semanticweb.org/ontologies/2020/10/accomm
odationURL> ?accomodationURL .
    ?activityId
<http://www.semanticweb.org/ontologies/2020/10/hasActivi
tyType> ?activityType.
    ?activityId
<http://www.semanticweb.org/ontologies/2020/10/activityN
ame> ?activity .
    ?activityId
<http://www.semanticweb.org/ontologies/2020/10/activityL
ocatedIn> ?activityLocation .
    FILTER(?locationPredicate = "Dublin" &&
?activityLocation = <http://foo.example/Location/Dublin>
&& ?activityType =
<http://foo.example/ActivitiesTypes/Fishing>)
}
```

This query refers to competency question 4, it allows the
user to see the various locations, accommodations, and
activities relating to fishing on a given day. This query
checks the availability of all the accommodations in the
same area as a fishing activity on the specified date.

#Q5. What are the amenities offered in accommodations for someone that wishes to visit Golf activity?

```
PREFIX
HostURL:<http://www.semanticweb.org/ontologies/2020/1
0/Host> PREFIX Activity:
<http://www.semanticweb.org/ontologies/2020/10/Activity>
PREFIX owl:<http://www.w3.org/2002/07/owl#> PREFIX
rdfs:<http://www.w3.org/2007/01/rdf-schema#> PREFIX
xsd:<http://www.w3.org/2001/XMLSchema#> SELECT
DISTINCT ?amenities ?activity ?activityLocation {
?locationObject
<http://www.semanticweb.org/ontologies/2020/10/activityL
ocation> ?locationPredicate. ?activityId
<http://www.semanticweb.org/ontologies/2020/10/hasActivi
tyType> <http://foo.example/ActivitiesTypes/Golf> .
?activityId
<http://www.semanticweb.org/ontologies/2020/10/activityN
ame> ?activity . ?activityId
<http://www.semanticweb.org/ontologies/2020/10/activityL
ocatedIn> ?activityLocation . ?accomodationId
<http://www.semanticweb.org/ontologies/2020/10/accomm
odationLocatedIn> ?activityLocation . ?accomodationId
<http://www.semanticweb.org/ontologies/2020/10/hasAmen
ities> ?amenitiesId . ?amenitiesId
<http://www.semanticweb.org/ontologies/2020/10/amenity
Name> ?amenities . } ORDER BY ?activityLocation
```

This query returns the amenities, activity and location,
regarding question five. This provides information for the
user to understand the types of amenities offered by
accomodations in the same area as golfing activities.

#Q6. What activities can you do that have a communication rating of 9?

```
PREFIX
AccRating:<http://www.semanticweb.org/ontologies/2020/1
0/AccommodationRating> PREFIX AccName:
<http://www.semanticweb.org/ontologies/2020/10/Accomo
dation> PREFIX
Activity:<http://www.semanticweb.org/ontologies/2020/10/
Activity> PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX rdfs:<http://www.w3.org/2007/01/rdf-schema#>
PREFIX xsd:<http://www.w3.org/2001/XMLSchema#>
SELECT DISTINCT ?activities ?CommunicationRating {
?rating
<http://www.semanticweb.org/ontologies/2020/10/accomm
odationRatingCommunication> ?CommunicationRating .
?activityType
<http://www.semanticweb.org/ontologies/2020/10/activityT
ypeName> ?activities . FILTER (?CommunicationRating
>= 9) }
```

This query uses the group by function to return unique
combinations of communicationRating and activities. This
allows the user to identify hosts with good communication
ratings in an area and what types of activities are available.

#Q7. What properties does host 234243 have in the same county as the Shopping Activity?

```
PREFIX
HostURL:<http://www.semanticweb.org/ontologies/2020/1
0/Host> PREFIX Activity:
<http://www.semanticweb.org/ontologies/2020/10/Activity>
PREFIX owl:<http://www.w3.org/2002/07/owl#> PREFIX
rdfs:<http://www.w3.org/2007/01/rdf-schema#> PREFIX
xsd:<http://www.w3.org/2001/XMLSchema#> SELECT
DISTINCT ?accomadationName ?hostName ?activity
?accommodationLocation { ?activityId
<http://www.semanticweb.org/ontologies/2020/10/hasActivi
tyType> <http://foo.example/ActivitiesTypes/Shopping> .
<http://foo.example/ActivitiesTypes/Shopping>
<http://www.semanticweb.org/ontologies/2020/10/activityT
ypeName> ?activity . ?activityId
<http://www.semanticweb.org/ontologies/2020/10/activityL
ocatedIn> ?activityLocation . ?accomodationId
<http://www.semanticweb.org/ontologies/2020/10/accomm
odationLocatedIn> ?accommodationLocation .
?accomodationId
```

<http://www.semanticweb.org/ontologies/2020/10/accomm odationName> ?accomadationName . ?hostId <http://www.semanticweb.org/ontologies/2020/10/ownsAcc ommodation> ?accomodationId . ?hostId <http://www.semanticweb.org/ontologies/2020/10/hostId> "234243" . ?hostId <http://www.semanticweb.org/ontologies/2020/10/hostNam e> ?hostName } GROUP BY ?accomadationName ?hostName ?activity ?accommodationLocation

This group by query relates to competency question seven. Where the user can see the properties are in the same county as the shopping activity, and, owned by host 234243.

#Q8. Where can someone stay to go on a fishing holiday for under 250 euro? (return available dates)
PREFIX HostURL:<http://www.semanticweb.org/ontologies/2020/1 0/Host> PREFIX Activity: <http://www.semanticweb.org/ontologies/2020/10/Activity> PREFIX owl:<http://www.w3.org/2002/07/owl#> PREFIX rdfs:<http://www.w3.org/2007/01/rdf-schema#> PREFIX xsd:<http://www.w3.org/2001/XMLSchema#> SELECT DISTINCT ?accomadationName ?availabilityDate ?activity { ?activityId <http://www.semanticweb.org/ontologies/2020/10/hasActivi tyType> <http://foo.example/ActivitiesTypes/Fishing> . ?activityId <http://www.semanticweb.org/ontologies/2020/10/activityN ame> ?activity . ?activityId <http://www.semanticweb.org/ontologies/2020/10/activityL ocatedIn> ?activityLocation . ?listingId <http://www.semanticweb.org/ontologies/2020/10/accomm odationLocatedIn> ?activityLocation . ?listingId <http://www.semanticweb.org/ontologies/2020/10/hasAcco mmodationAvailabilities> ?availability . ?listingId <http://www.semanticweb.org/ontologies/2020/10/accomm odationName> ?accomadationName . ?availability <http://www.semanticweb.org/ontologies/2020/10/availabili tyDate> ?availabilityDate . ?availability <http://www.semanticweb.org/ontologies/2020/10/availabili tyPrice> ?price . FILTER(?price <= 250.00) } order by ?availabilityDate

This query relates to question 8. The data returned is accommodations near fishing activities where the price of the accommodation is less than €250.

#Q9. What accommodations are available dates to visit "Ardgillan Demesne" where the host has an acceptance rate of 70%?
PREFIX HostURL:<http://www.semanticweb.org/ontologies/2020/1 0/Host> PREFIX Activity: <http://www.semanticweb.org/ontologies/2020/10/Activity> PREFIX owl:<http://www.w3.org/2002/07/owl#> PREFIX rdfs:<http://www.w3.org/2007/01/rdf-schema#> PREFIX xsd:<http://www.w3.org/2001/XMLSchema#> SELECT DISTINCT ?hostName ?acceptanceRate ?roomType ?activityLocation { ?activityId <http://www.semanticweb.org/ontologies/2020/10/activityN ame> "Ardgillan Demesne" . ?activityId <http://www.semanticweb.org/ontologies/2020/10/activityL ocatedIn> ?activityLocation . ?accomodationId <http://www.semanticweb.org/ontologies/2020/10/accomm odationLocatedIn> ?activityLocation . ?accomodationId <http://www.semanticweb.org/ontologies/2020/10/isRoomT ype> ?roomTypeId . ?roomTypeId <http://www.semanticweb.org/ontologies/2020/10/roomTyp e> ?roomType . ?accomodationId <http://www.semanticweb.org/ontologies/2020/10/accomm odationOwnedBy> ?hostId . ?hostId <http://www.semanticweb.org/ontologies/2020/10/hostAcce ptanceRate> ?acceptanceRate . ?hostId <http://www.semanticweb.org/ontologies/2020/10/hostNam e> ?hostName FILTER (?acceptanceRate >= 70) }

This query relates to question. This provides the user with information about the accommodations where the acceptance rate is higher than 70% in the area of the Ardgillan Demesne.

#Q10. Where can someone go on a fishing holiday for at least 3 nights?
PREFIX HostURL:<http://www.semanticweb.org/ontologies/2020/1 0/Host> PREFIX Activity: <http://www.semanticweb.org/ontologies/2020/10/Activity> PREFIX owl:<http://www.w3.org/2002/07/owl#> PREFIX rdfs:<http://www.w3.org/2007/01/rdf-schema#> PREFIX xsd:<http://www.w3.org/2001/XMLSchema#> SELECT DISTINCT ?accomadationName ?locationName ?activity ?availabilityDate ?availabilityMin{ ?activityId <http://www.semanticweb.org/ontologies/2020/10/hasActivi tyType> <http://foo.example/ActivitiesTypes/Fishing> . ?activityId <http://www.semanticweb.org/ontologies/2020/10/activityN ame> ?activity . ?activityId <http://www.semanticweb.org/ontologies/2020/10/activityL ocatedIn> ?activityLocation . ?locationId <http://www.semanticweb.org/ontologies/2020/10/partOfRe gion>?activityLocation . ?locationId <http://www.semanticweb.org/ontologies/2020/10/countyN ame> ?locationName . ?listingId <http://www.semanticweb.org/ontologies/2020/10/accomm odationLocatedIn> ?activityLocation . ?listingId <http://www.semanticweb.org/ontologies/2020/10/hasAcco mmodationAvailabilities> ?availability . ?availability <http://www.semanticweb.org/ontologies/2020/10/availabili tyDate> ?availabilityDate . ?listingId <http://www.semanticweb.org/ontologies/2020/10/accomm odationName> ?accomadationName . ?availability <http://www.semanticweb.org/ontologies/2020/10/availabili tyMin> ?availabilityMin . FILTER(?availabilityMin <= 3) }

This query filters availability of accommodation where the user can stay for at least 3 nights. This allows the user to understand where they can stay on their fishing trip that lasts longer than three days.

VIII. Challenges Faced

The main challenges that the project team face while creating the Airbnb ontology are the structure of the downloaded data CSV files to be inserted into the ontology and the size of the data files.

A. Preprocessing Data

The data needed to to be preprocessed before inserting the data into the ontology as some of data properties such as amenities and activity types has many values within the one column this mean that the meant that this data had to be preprocessed before inserting it into the ontology as it would make it more difficult to attempt to use the R2RML directly on this data. Another example of the data that needed to be preprocessed was the price data attribute as this was encoded to include the "$" currency symbol.

B. R2RML

The project team also struggled with the R2RML mapping due to the size of the selected datasets.

1. The activities CSV file has 11,537 rows of data (including the header) and is 2.1 Mb in size.
2. The calendar CSV file has 9,904,602 rows of data (including the header) and is 425.5Mb in size.
3. The listings CSV file has 27,132 rows of data (including the header) and is 302.6Mb in size.
4. The reviews CSV file has 1,003,086 rows of data (including the header) and is 64.9Mb in size.

Due to the size of these datasets and the teams limited access to hardware this meant that the team was not able to insert all of the data into the ontology, thus the team took a sample of the data.

IX. Conclusion

This project followed the process from ideation, design to implementation of an ontology-driven application. We firstly created our ontology model, uplifted our data, designed competency questions, and finally created complex queries that used both data sets which are presented in our application. A weakness of our ontology model is that it relies heavily on the refresh of data from both Airbnb and the activities dataset. The strengths of our ontology model are it all allows the user to query the various tourism datasets with ease in order to plan their holidays.

X. References

[1] *Fáilte Ireland, "Activities", 2020 [Online] Available: https://data.gov.ie/dataset/activities [Accessed on: 02/12/2020].*

[2] Inside Airbnb, "Get the Data", 2020 [Online] Available: http://insideairbnb.com/get-the-data.html [Accessed on: 02/12/2020].

[3] D. Allenang, J. Hendler, and F. Gandon, *Semantic Web for the Working Ontologist: Effective Modeling for Linked Data, RDFS, and OWL,* New York, ACM Books, 2020.

[4] DBPedia, "Data Download", 2020 [Online] Available: https://wiki.dbpedia.org/ [Accessed on: 02/12/2020].