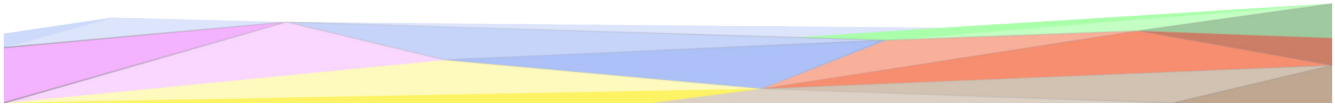
UNIVERSITY
OF TRENTO - Italy

Dipartimento di Ingegneria e Scienza dell'Informazione

– KnowDive Group –

Food and Accommodation

Document Data:	Reference Persons:
February 1, 2022	Byambakhishig Ganbold, Mithat Sinan Ergen



Index:

1 Introduction

Reusability is one of the main principles in the Data Integration (DI) process defined by iTelos. The data integration project documentation plays an important role in order to enhance the reusabiltiy of the resources handled during the methodology, as well as for the resources produced by the data integration process. A clear description of the resources and the process that has to manage them, provides a clear understanding of the information handled in the DI project, allowing external readers to exploit the same resources in different projects.

The current document aims to provide a detailed report of the DI project developed following the iTelos methodology. The report is structured, on top, to describe:

- Section 1: The project's purpose and the resources involved (both schema and data resources) in the integration process.
- Section 2, 3, 4, 5: The integration process along the iTelos phases.
- Section 6: How the result of the integration process (KGs) can be exploited.

2 Purpose and project's resources

This section reports and describes:

- The project's domain and purpose, by reporting the purpose itself and the definition of the project's domain, personas, and scenarios.
- Knowledge resources: The collected reference teleologies that satisfy the purpose along the integration process.
- Data resources: The data collected to satisfy the purpose along the integration process.
- Metadata: The metadata defined for the knowledge and data resources.

2.1 Project Domain and Purpose

The main purpose of this project is to provide users with information regarding food and accommodation in the Trentino region. Users will be able to plan their vacations by checking the information in the application. They will be able to find places to stay according to their personal preferences. They will be able to discover restaurants serving food they enjoy, suitable for their budget.

The accommodation information will include hotels, apartments, private rooms and camping with detailed information about price, facilities offered, number of stars, type of rooms in the Trentino region. The food information will include restaurants, bars, pizzerias, ice-cream shops and their detailed information such as phone number, email and website.

2.2 Scenario and Personas

1. **Sofia**
   Sofia is a 23 year old, studying law at the University of Milan. She loves to watch sports because she is also a professional gymnast. During her high school years, she participated in a lot of competitions very successfully. The beginning of the semester is a good time for travelling since school recently started and there is not much workload. She likes to choose the cheapest option for a trip, eats vegan food. She wants to use the system for organizing trips.

   Scenario A: She is going to visit Trentino during the **sport festival** with her best friend Annia. They want to watch a lot of sports games, have fun and meet new people. During the festival days, as Trento is super crowded, restaurants and hotels are very busy. Girls need the system to find cheaper vegan restaurants, bars, pubs and hotels.

2. **Nara**
   Nara is a 35 year old businesswoman, living in France. The best suitable season is summer for her new trip, which includes visiting beautiful places and feeling that atmosphere. She loves luxury hotels, restaurants and bars, she doesn't mind the prices. Nara intends to use the system to look for well-located hotels and restaurants with good views. The system could be an optimal solution for her.

   Scenario B: Nara is planning to travel Trentino this **summer vacation** for two weeks. Alps mountains and Lake Garda are her main destinations. She wants to book 4 or 5 stars hotels with swimming pools, spas and car parking located near mountains and lake Garda, also restaurants with terrace and good service. She doesn't like having unnecessary issues, therefore she uses the system.

3. **Marco**
   Marco is 46 years old, has a big family, boys and a wife. He loves to spend time with family on vacation. When they go on a trip, they are looking for family hotels, apartments and guest houses with lifts in good locations. Because his wife had a car accident, she cannot climb stairs. The boys very much enjoy hiking and skiing with dad. They want to use the system to find accommodation, close restaurants and pubs. They usually travel with their own vehicle.

   Scenario C: Marco's family is going to spend the **winter vacation** in Italy. This winter is the most beautiful moment for boys, they are super delighted to travel to the Alps for skiing in Trentino. They are looking for family hotels, guest houses and apartments, if the accommodation has a kitchen it is the best choice for them. Also boys love to watch TV and use the Internet, but the first necessary requirement is very good heating, lift and closeness to ski resorts. They are going to choose restaurants and pubs which are located close to ski resorts.

4. **Francesco**
   Francesco is a 28 year old person who is disorganized, he does not like to plan everything.

He likes to travel for extreme sports, which are motorcycling, skiing and snowboarding. He doesn't mind staying in uncomfortable places, he just prefers hostels and junk food. He enjoys short vacations.

Scenario D: Francesco is travelling in Trentino for the **weekend**. This time he wants to try rafting and canyoning, so he found one place in Croviana. During this trip he doesn't want to spend a lot of money, books a hostel and eats fast food. If he has time, he would like to go to a pub and bar.

5. **Steve**
   Steve is a 65 year old married man living with his wife. He has a CEO role in a big iron and steel company. Therefore, he is wealthy and he does not mind paying high prices when travelling. He is not very good with technology so he does not spend a lot of time to look for hotel details, he needs fast and efficient results. The only property he looks for in a hotel is silence and closeness to restaurants serving healthy food. He likes to reserve seats by phone so he is always looking for phone numbers of the restaurants.

   Scenario E: Steve recently came to Trento for a **business meeting**, he wanted to use this opportunity to have a vacation with his family close to the mountains with fresh air. Because of the meetings, he does not have too much time to organize hotels and restaurants. He occasionally enjoys having drinks with his colleagues, so he is looking for places suitable for a business concept where he can book a seat by phone.

6. **Mary**
   Mary is young 18 year old girl with full of energy. She loves to go on vacation in the mountains. She wants to use the system to find hotels close to ski centers. She also enjoys discovering new places and she does not mind staying in cheap hostels so she is also intending to use the system to find hostels in original places. She mainly eats junk food so it is important for her to find fast food restaurants around the city.

   Scenario F: Mary is quite poor and she spends her allowance from her family for travelling. She is visiting Trentino region to look for **adventure** with her friends so she needs high energy food, cheap hostels and hikes to the mountains. She is a smoker so she will look for places where she can smoke and she wants to check the web site of restaurants before visiting them.

7. **Brian**
   Brian is a 50 year old gourmet who likes to eat quality food, he mainly enjoys exquisite fish and desserts. He is intending to use the system to find fish restaurants and pasticcerie. He is also a man of luxury but he likes to look for the best price, he wants to use the system to find 5 star hotels that offer a decent promotion.

   Scenario G: Brian is **touring the world** and looking for original food in every place he visits. When he cannot find original food, he wants to enjoy the best food possible. He is in Trento for now but has a plan to visit all the Trentino region and then continue

north. So he needs the application to find restaurants and hotels fitting his needs in all the Trentino region.

8. **Angel**
   Angel is a 25 year old sportive woman, she likes to play tennis even when she is on vacation. She counts her daily protein intake so she prefers eating high protein food. Despite being careful to what she is eating, she is quite addicted to ice-cream. She is intending to use the system to find hotels with tennis courts and find steak houses to boost up her protein level. Besides playing tennis, she is also a competitive sail racer.

   Scenario H: Angel is in Trento for a couple of days, then she is planning to attend **sail races** in Riva del Garda. She wants to play tennis while she is in Trento and she is looking for hotels either with a tennis court or close to tennis courts. She will also use the app to find hotels in Riva del Garda and find ice-cream shops to eat plenty of ice-cream.

2.3 Knowledge Resources

A Knowledge resource is a dataset which consists of a KB encoding information about schemas, that is entity types and properties. KBs of high quality are usually called ontologies. We call them teleologies, that is meaning by this, ontologies with metadata which empower their practical use in knowledge and data integration.

- We have picked Schema.org as our guide for the vocabulary we are using in our project. Schema.org is a collaborative, community activity with a mission to create, maintain, and promote schemas for structured data on the Internet, on web pages, in email messages, and beyond.

  Schema.org vocabulary can be used with many different encodings, including RDFa, Microdata and JSON-LD. These vocabularies cover entities, relationships between entities and actions, and can easily be extended through a well-documented extension model. Over 10 million sites use Schema.org to markup their web pages and email messages. Many applications from Google, Microsoft, Pinterest, Yandex and others already use these vocabularies to power rich, extensible experiences. A shared vocabulary makes it easier for webmasters and developers to decide on a schema and get the maximum benefit for their efforts. It is in this spirit that the founders, together with the larger community have come together - to provide a shared collection of schemas.

2.4 Data Resources

A data resource is a data set which consists of data in some format,that is tabular, unstructured, entities and property values. This project includes three kinds of data sources:

- Open Data Trentino: it is an open data source providing data for the Trentino region for the purpose of stimulating growth, efficiency and participation.

Open Data Trentino is promoted by the Autonomous Province of Trento with the involvement of Trentino Digitale, the Consortium of Municipalities, the Bruno Kessler Foundation and the University of Trento. Now that system has 6,306 datasets.

It aims to open its data to the digital world, not only as a support to the transparency of public service operations, but above all as a contribution that the process of enhancing public information capital can lead to economic growth through the development of new services.

The Trentino Open Data project is a tool for the benefit of companies and professionals interested in using the information assets of the Public Administration in new services and new applications subsidiary to the services of the Province. The datasets are also published systematically on the national portal Dati.Gov.it and on the European one Europeandata-portal.eu

- Inside Airbnb: Inside Airbnb is an independent, non-commercial set of tools and data that allows you to explore how Airbnb is really being used in cities around the world. By analyzing publicly available information about a city's Airbnb's listings, Inside Airbnb provides filters and key metrics so you can see how Airbnb is being used to compete with the residential housing market.

- Open Street Map: it is a free collaborative project providing data of the world map.

  OpenStreetMap was born in the UK 2004, it is built by a community of mappers that contribute and maintain data about roads, trails, cafes, restaurants, pubs, bars, railway stations, and much more, all over the world.

  The fundamental characteristic of the geographic data present in OSM is that they are distributed with a free license, the Open Database License: it is possible to use them freely for any purpose, including commercial, with the only constraint of citing the source and using the same license for any works derived from OSM data. Everyone can contribute by enriching or correcting the data. A proprietary alternative that took up the crowd-sourcing practices introduced by the OpenStreetMap community was Google Map maker, in which user contributions could be entered on Google Maps against validation, preventing their reuse by third parties.

  Maps are created using data recorded by portable GPS devices , aerial photographs and other free sources as a reference . Both images rendered the vector data, as well as the same database of geodata are widespread licensed under the Open Database License.

  OpenStreetMap was inspired by sites such as Wikipedia: the page where the map is available displays an "Edit" label to proceed with the modification of the data and the project is accompanied by a historical archive of the modifications (history and log). Registered users can upload GPS tracks to the project databases and edit vector data using the editors provided.

  Features of the Open street Map :

  - Browse map content in certain categories
  - Wikipedia integration

- It uses the Overpass API to access the map data
- Download selected object or visible map features in various formats (GeoJSON, OSM JSON, OSM XML formats)
- You can create categories yourself
- Available in several languages

OSM is a database product. There are, however, a number of open source software projects and proprietary products built specifically to edit the database. They are built on the OSM API, that is fetching and saving raw geodata from/to the OpenStreetMap database.

The map database is enhanced by individuals, that is kind of volunteers around the world. Some find it fun and rewarding and consider it a hobby. Others simply want a great map of an area that is important to them. Still others offer help in times of crisis, such as earthquakes and tsunamis.

We gathered hotel and restaurant data for the Trentino region from Open Data Trentino. Data was available in CSV and JSON format so it was straight-forward to extract it from the website.

Secondly, we looked for a resource to find Airbnb data and we've found that Open Data Soft contains a good amount of data on apartments, rooms available on Airbnb. Their data was based on InsideAirbnb website so we decided to get the data from the original source. It was also possible to filter the data according to different parameters such as property type and host response time. We did not filter the data and exported it as a whole in CSV and JSON formats.

Lastly, we used Open Street Map to get more detailed restaurant, bar, ice-cream shop, cafe, fast-food restaurant data in the Trentino region. We had to select small sections on the map and filter them according to different amenities. Then we exported the data in JSON format.

3 Inception

This section reported the integration sub process performed during the inception phase, by describing each activities both in schema and data layer.

Inception sub activities:

- Reusability Categories
- Purpose Formalization and Inception Sheet
- Metadata
- Inception phase evaluation

3.1 Reusability Categories

In order to identify the level of reusability of the resources collected and handled along the methodology, all the resource elements, which are entities, entity properties, e-types, and others are classified into one of three reusability categories, defined as follow:

- Common: this category involves resources associated to aspects which are common to all domains, also outside the project's.
- Core: this category involves information about the most important aspects considered by the purpose, information without which it would be impossible to develop the EG.
- Contextual: this category involves resources which carry specific, possibly unique, information from the domain of interest. These are the resources whose main goal is to create added value. If core resources are necessary for a meaningful application, contextual resources are the ones which can make the difference with respect to the competitors.

3.2 Purpose Formalization and Inception Sheet

We have 8 personas each one has around 7,8 queries, we try to create competency queries(CQs) based on scenarios and datasets which consist of e-types and their properties. According to reusability category, we've divided competency queries to three kernel concepts. Common kernel concept involves location which are reusable in multiple occassions.

To explain better, we had to use location in every single CQ since our application involves the location for every query. Contextual kernel concepts include specific unique information which add value to main concept, we considered them as extra amenities that are offered in hotels or restaurants such as smoking places, swimming pools. Core kernel concept includes the elements that contribute to the main goal. Since we are dealing with accommodation and food our core kernel concepts involve everything related to these concepts such as hotels, restaurants, apartments.

Persona	**NUM**	**Competency Question**	**Common Kernel Concepts**	**Core Kernel Concepts**	**Contextual Kernel Concepts**
Sofia	1.1	Give the list of hotels in city center	location	Hotel	
Sofia	1.2	Give the list of hotels with lower price	location	Hotel	price
Sofia	1.3	Give the email of vegan restaurants in Trento	location	Restaurant	vegan , email
Sofia	1.4	Give the list vegan restaurant with internet access in Rovereto	location	Restaurant	vegan , internet access

Sofia	1.5	Give the phone numbers of restaurants with outdoor seating	location	Restaurant	outdoor seating, phone number
Sofia	1.6	Give the list of hotels with phone number information near train station	location	Hotel	phone number
Sofia	1.7	Give the comfortable hotels under 60€ near mountain and good view	location	Hotel	price
Sofia	1.8	Give the list of the pub in city center open 9 pm	location	Pub	open hours
Sofia	1.9	Give the list of cafe with internet access and takeaway	location	Cafe	internet access, takeaway
Nara	2.1	Give the list of 4 or 5 stars hotels with swimming pool in Trentino	location	Hotel	stars, swimming pool and spas
Nara	2.2	Give the list of hotel with terrace?	location	Hotel	amenity, terrace
Nara	2.3	Give the list of Airbnb private room near Garda Lake with car parking	location	Apartment	car parking, private room
Nara	2.4	Give the list of restaurant with outdoor seating in Garda Lake	location	Restaurant	outdoor seating

Nara	2.5	Give the list of restaurants with brewery and outdoor seating	location	Restaurant	brewery, outdoor seating
Nara	2.6	Give the address of bars in Trento	location	bar	address
Nara	2.7	Give the list of phone number of pizzas in Trento	location	fast_food_restaurant	phone number
Nara	2.8	Give the list of ice cream shops with frozen_yogurt and outdoor_seating	location	Ice_cream-shop	frozen_yogurt ,outdoor_seating
Marco	3.1	Give the list of Airbnb houses with lift for 4 people in Rovereto	location	Apartment	lift
Marco	3.2	Give the list of Airbnb houses with middle level price	location	Apartment	price
Marco	3.3	Give the list of Airbnb apartment with kitchen and TV	location	Apartment	kitchen and TV
Marco	3.4	Give the list of Airbnb houses with heating near ski resorts	location	Apartment	heating, near ski resort
Marco	3.5	Give the list of restaurant with delivery near Andalo	location	Restaurant	delivery
Marco	3.6	Give the list of restaurant with wheelchair	location	Restaurant	wheelchair

Marco	3.7	Give the cafe with wheelchair and vegan food	location	cafe	wheelchair, vegan food
Marco	3.8	Give the list of pubs with outdoor seating	location	pub	outdoor seating
Francesco	4.1	Give the list of camping with ping pong	location	Camping	ping pong
Francesco	4.2	Give the list of hotel with breakfast in city center	location	Hotel	amenity, breakfast
Francesco	4.3	Give the campings with internet close to mountains	location	Camping	amenity, internet
Francesco	4.4	Give the cheapest accommodation with phone number information near mountains	location	Camping	phone number
Francesco	4.5	Give the list of fast food restaurant with takeaway near mountains	location	fast_food_restaurant	takeaway
Francesco	4.6	Give the list of the cheapest pizzas places	location	fast_food_restaurant	cheapest
Francesco	4.7	Give the phone number of fast food restaurant with delivery	location	fast_food_restaurant	phone number, delivery
Francesco	4.8	Give the list of bars with smoking place	location	bar	smoking
Steve	5.1	Give the list of hotels in Trento	location	Hotel	

Steve	5.2	Give the list of hotels that are not in the heart of town	location	Hotel	
Steve	5.3	Give the list of restaurants serving healthy food	location	Restaurant	serving healthy food
Steve	5.4	Give the phone number of Italian restaurants close to city center	location	Restaurant	phone number, Italian restaurants
Steve	5.5	Give the location of pubs close to Trento city center	location	pub	
Steve	5.6	Give the list of restaurants serving multiple food types	location	Restaurant	food type
Mary	6.1	Give list of hotels close to the mountains	location	Hotel	
Mary	6.2	Give list of campings situated outside Trento	location	Camping	
Mary	6.3	Give list of hotels close to train stations	location	Hotel	
Mary	6.4	Give the list of fast food restaurants open in the morning	location	fast_food_restaurant	opening hours
Mary	6.5	Give the list of restaurants outside Trento with a smoking space	location	Restaurant	smoking

Mary	6.6	Give the web site information of restaurants outside Trento	location	Restaurant	website
Mary	6.7	Give the list of Airbnb houses for maximum 2 people	location	Apartment	num-ber_of_guests
Mary	6.8	Give the list of Airbnb private rooms	location	Apartment	private rooms
Brian	7.1	Give list of 5 star hotels	location	Hotel	stars
Brian	7.2	Give list of 4 or 5 star hotels with price lower than 150 euros	location	Hotel	stars, price
Brian	7.3	Find fish restau-rant close to the center of Trento	location	Restaurant	fish restaurant
Brian	7.4	Find pasticcerie close to the cen-ter of Trento	location	cafe	
Brian	7.5	Find restau-rants that are open everyday of the week	location	Restaurant	opening hours
Brian	7.6	Find bars close to the hotels	location	bar	
Brian	7.7	Give the list of Airbnb houses for a family with multiple reviews	location	Apartment	review
Brian	7.8	Give the list of Airbnb houses that are expen-sive and have many reviews	location	Apartment	review, price
Angel	8.1	Find hotels with tennis courts	location	Hotel	tennis court

Angel	8.2	Find steak-houses that are open everyday	location	Restaurant	opening hours
Angel	8.3	Find hotels close to the center	location	Hotel	
Angel	8.4	Find hotels with mid-level price	location	Hotel	price
Angel	8.5	Find ice-cream shops close to Trento city cen-ter	location	Ice_cream-shop	
Angel	8.6	Find Airbnb houses where you can stay short	location	Apartment	minimum_stay

Table 1: Competency Queries

3.3 Metadata and Data Content Information

In this section, metadata and content information on datasets are presented. The metadata information for the datasets is given in the following tables:

- Osteria Metadata from OpenData Trentino in Table 2.
- First Hotel Metadata from OpenData Trentino is shown in Table 3
- Second Hotel Metadata from OpenData Trentino is shown in Table 4
- Airbnb metadata from InsideAirbnb is shown in Table 5
- Restaurant metadata from Open Street Map is shown in Table 6

The content information for the datasets is given in the following tables:

- Restaurant data content from Open Street Map is shown in Table 7
- Osterie of Trentino data content from Open Data Trentino is shown in Table 8
- Accommodation data content from Open Data Trentino is shown in Table 9
- Extra accommodation data content including non-hotel organizations from Open Data Trentino is shown in Table 10
- Airbnb data content from Inside Airbnb is shown in Table 11

Dataset identifier	p_TN:6fbd1580-2264-43df-a91d-9e809893cbb0
Other identifier	N/A
Themes	Istruzione, cultura e sport 2826 vita sociale
Dataset Publisher Name	Servizio artigianato e commercio IPA/IVA Code: 8VFPQI
Issued	16-12-2013
Date modified	03-06-2019
Geographic Coverage	Trento
GeoNames URI	http://www.geonames.org/3165241
Language	italian
Time Extension	N/A
Holder Name	Provincia Autonoma di Trento Codice IPA/IVA: p_TN
Update Frequency	annual
Version	N/A
Compliant with	Standard: utf-8
Author	Servizio artigianato e commercio IPA/IVA: 8VFPQI

Table 2: Osteria Metadata from OpenData Trentino

Dataset identifier	p_TN:d8948251-9e9b-4aa5-9ab1-d13e37cd4320
Other identifier	N/A
Themes	Economy, finance, regions, cities
Dataset Publisher Name	Servizio Turismo e Sport
Issued	28-02-2014
Date modified	25-10-2017
Geographic Coverage	Trento
GeoNames URI	https://www.geonames.org/3165241
Language	italian
Time Extension	N/A
Holder Name	Provincia Autonoma di Trento
Update Frequency	weekly
Version	N/A
Compliant with	Standard: utf-8
Author	Servizio Turismo e Sport IPA/IVA:294WVJ

Table 3: Hotel Metadata #1 from OpenData Trentino

Dataset identifier	p_TN:bf341ac4-aa6e-4c7c-9bef-90d11a70e6ac
Other identifier	N/A
Themes	Economy, finance, population, society
Dataset Publisher Name	Servizio Turismo e Sport
Issued	28-02-2014
Date modified	28-02-2014
Geographic Coverage	Trento
GeoNames URI	https://www.geonames.org/3165241
Language	italian
Time Extension	N/A
Holder Name	Provincia Autonoma di Trento
Update Frequency	daily
Version	N/A
Author	Servizio Turismo e Sport

Table 4: Hotel Metadata #2 from OpenData Trentino

Date Introduced	August, 2020
Version	4
Dataset Publisher Name	InsideAirbnb
Author	InsideAirbnb
Geographic Coverage	Trento
Compliant with	Standard: utf-8
Language	en

Table 5: Airbnb Metadata from InsideAirbnb

Date Introduced	September 2012
Dataset Publisher Name	OpenStreetMap
Author	OpenStreetMap
License	OpenStreetMap License
Geographic Coverage	Trento
Update schedule	Weekly
Compliant with	Standard: utf-8
Language	en

Table 6: Restaurant Metadata from Open Street Map

Attribute	**Explanation**
ID	Unique numbers representing restaurant

Name	Corresponding name of the restaurant
Address for street	corresponding name of street for the restaurant
Address for house number	House number of the restaurant
Postcode	Postcode of the restaurant
City	city of the restaurant
Country	Country of the restaurant
Amenity	Which type of service, for example pub, fast-food, restaurant
Brand	Which company of restaurant, like McDonald's, Burger king
Contact email	Email of restaurant
Contact Phone	Phone number of corresponding restaurant
Contact website	Website of corresponding restaurant
Contact Facebook	Facebook of corresponding restaurant
Outdoor seating	Some restaurant has a square for outside seating
Takeaway	Extra service
Opening hours	Open hours of restaurant
Cuisine	It shows which kind of foods are served, for example fish, steak, pizza
Ref - Vatin	Country code and Vat number
Operator	Owner of restaurant
Wheelchair	For disabled people
Surveillance	Security camera
Brewery	Makes and sells beer
Diet - vegan	Shows if menu includes foods for vegan people or not
Diet - vegetarian	Shows if menu includes foods for vegetarian people or not
Diet - gluten free	Shows if menu includes gluten free foods or not
Internet access	Wifi
Smoking	Some restaurant has a place for smoking
latitude	Latitude of corresponding restaurant
longitude	Longitude of corresponding restaurant
Source	About restaurant data source, for example local knowledge
stars	star and rating

Table 7: Food data from Open Street Map

Attribute	Explanation
Nr	progressive identifier
Comune	municipality where the business is located
Insegna	name
Tipo	type of business (restaurant; Restaurant-Bar; Hotel-Restaurant-Bar; Restaurant-Pizzeria; Restaurant-Pizzeria-Bar; Hotel-Restaurant-Pizzeria-Bar)
Frazione	fraction where the business is located
Indirizzo	address of the business
Civico	house number

Table 8: Osterie Data from Open Data Trentino

Attribute	Explanation
data-inizio-validita	Starting date of the published price
data-fine-validita	Expiry date of the published price
denominazione	Name of the municipality
denominazione-ente	Alternative name of the municipality
denominazione-ente-annuario	Responsible touristic agency
altitudine	Altitude
cap	Postal code
id-localita	Internal ID of the STU system
numero-unita	Number of room available
numero-posti-letto	Number of bed available
tipologia-alberghiera	Type of facility
tipologia-servizio	Type of boarding (e.g.full board, only breakfast etc.)
recapito-telefono	Phone number
recapito-fax	Fax number
recapito-email	email
recapito-www	website
id-EsercizioRicettivo	Internal ID of the STU system
p-iva	VAT number
comune	Municipality
frazione	where the business is located
indirizzo	Address
livello-classifica	Ranking (when available)
prezzo-max-letto-aggiunto	Maximum price for extra bed
struttura-chiusa	Status, open=false/seasonal closure=true

tipologia-unita	Type of accommodation (e.g. room with a private bathroom)
posti-letto	Bed place available for each type of accomodation
prezzo	Price
codice	Internal ID of the STU system
descrizione	Description of the available service
prezzo-max-consumazione-pasto	Maximum price for the meal

Table 9: Accommodation data from OpenData Trentino

Attribute	**Explanation**
data-inizio-validita	Starting date of the published price
data-ultimo-aggiornamento	Expiry date of the published price
denominazione-annuario-Localita-Turistica	Alternative name of the municipality
denominazione-ente-annuario-Localita-Turistica	Responsible touristic agency
denominazione-Localita-Turistica	Name of the municipality
altitudine-Localita	Altitude
cap-Localita	Postal code
id-localita	Internal ID of the STU system
tipologia-extraalberghiera	Type of facility
tipo-servizio	Type of boarding (e.g.full board, only breakfast etc.)
recapito-telefono	Phone number
recapito-fax	Fax number
recapito-email	email
recapito-www	website
numero-posti-letto	Number of bed available
id-EsercizioRicettivo	Internal ID of the STU system
p-iva	VAT number
comune	Municipality
frazione	where the business is located
indirizzo	Address
livello-classifica	Ranking (when available)
struttura-chiusa	Status, open=false/seasonal closure=true
tipologia-unita	Type of accommodation (e.g. apartment, camera-con-bagno
prezzo	Price
codice	Internal ID of the STU system
descrizione	Description of the available service

Table 10: Non-hotel Accommodation data from OpenData Trentino

Attribute	Explanation
id	Airbnb's unique identifier for the listing
listing_url	URL to the Airbnb listing
scrape_id	Inside Airbnb "Scrape" this was part of
last_scraped	UTC. The date and time this listing was "scraped".
name	Name of the listing
description	Detailed description of the listing
neighborhood_overview	Host's description of the neighbourhood
picture_url	URL to the Airbnb hosted regular sized image for the listing
host_id	Airbnb's unique identifier for the host/user
host_url	The Airbnb page for the host
host_name	Name of the host. Usually just the first name(s).
host_since	The date the host/user was created. For hosts that are Airbnb guests this could be the date they registered as a guest.
host_location	The host's self reported location
host_about	Description about the host
host_response_time	Average response time of the host
host_response_rate	Average response rate of the host
host_acceptance_rate	That rate at which a host accepts booking requests.
host_is_superhost	If the host is a superhost
host_thumbnail_url	URL to the thumbnail of the host
host_picture_url	URL to the picture of the host
host_neighbourhood	Neighbourhood of the host
host_listings_count	The number of listings the host has (per Airbnb calculations)
host_total_listings_count	The number of listings the host has (per Airbnb calculations)
host_verifications	
host_has_profile_pic	
host_identity_verified	If the identity of the host is verified
neighbourhood	

neighbourhood_cleansed	The neighbourhood as geocoded using the latitude and longitude against neighborhoods as defined by open or public digital shapefiles.
neighbourhood_group_cleansed	The neighbourhood group as geocoded using the latitude and longitude against neighborhoods as defined by open or public digital shapefiles.
latitude	Uses the World Geodetic System (WGS84) projection for latitude and longitude.
longitude	Uses the World Geodetic System (WGS84) projection for latitude and longitude.
property_type	Self selected property type. Hotels and Bed and Breakfasts are described as such by their hosts in this field
room_type	Entire home/apt|Private room|Shared room|Hotel
accommodates	The maximum capacity of the listing
bathrooms	The number of bathrooms in the listing
bathrooms_text	The number of bathrooms in the listing. On the Airbnb web-site, the bathrooms field has evolved from a number to a textual description. For older scrapes, bathrooms is used.
bedrooms	The number of bedrooms
beds	The number of bed(s)
amenities	
price	daily price in local currency
minimum_nights	minimum number of night stay for the listing (calendar rules may be different)
maximum_nights	maximum number of night stay for the listing (calendar rules may be different)
minimum_minimum_nights	the smallest minimum_night value from the calendar (looking 365 nights in the future)
maximum_minimum_nights	the largest minimum_night value from the calendar (looking 365 nights in the future)
minimum_maximum_nights	the smallest maximum_night value from the calendar (looking 365 nights in the future)
maximum_maximum_nights	the largest maximum_night value from the calendar (looking 365 nights in the future)
minimum_nights_avg_ntm	the average minimum_night value from the calendar (looking 365 nights in the future)

maximum_nights_avg_ntm	the average maximum_night value from the calendar (looking 365 nights in the future)
calendar_updated	
has_availability	[t=true; f=false]
availability_30	availability_x. The availability of the listing x days in the future as determined by the calendar. Note a listing may not be available because it has been booked by a guest or blocked by the host.
availability_60	availability_x. The availability of the listing x days in the future as determined by the calendar. Note a listing may not be available because it has been booked by a guest or blocked by the host.
availability_90	availability_x. The availability of the listing x days in the future as determined by the calendar. Note a listing may not be available because it has been booked by a guest or blocked by the host.
availability_365	availability_x. The availability of the listing x days in the future as determined by the calendar. Note a listing may not be available because it has been booked by a guest or blocked by the host.
calendar_last_scraped	
number_of_reviews	The number of reviews the listing has
number_of_reviews_ltm	The number of reviews the listing has (in the last 12 months)
number_of_reviews_l30d	The number of reviews the listing has (in the last 30 days)
first_review	The date of the first/oldest review
last_review	The date of the last/newest review
review_scores_rating	
review_scores_accuracy	
review_scores_cleanliness	
review_scores_checkin	
review_scores_communication	
review_scores_location	
review_scores_value	
license	The licence/permit/registration number

instant_bookable	[t=true; f=false]. Whether the guest can automatically book the listing without the host requiring to accept their booking request. An indicator of a commercial listing.
calculated_host_listings_count	The number of listings the host has in the current scrape, in the city/region geography.
calculated_host_listings_count_entire_homes	The number of Entire home/apt listings the host has in the current scrape, in the city/region geography
calculated_host_listings_count_private_rooms	The number of Private room listings the host has in the current scrape, in the city/region geography
calculated_host_listings_count_shared_rooms	The number of Shared room listings the host has in the current scrape, in the city/region geography
reviews_per_month	The number of reviews the listing has over the lifetime of the listing

Table 11: Airbnb Accommodation Data

3.4 Inception level evaluation

In this section we reported evaluation phase which is calculated for schema level and data level. We compare e-types and data properties in CQs with reference teleology and datasets. CQs has 11 e-types and 50 properties. Ontology reference has 23 e-types, they have 250 properties.

Schema level: aim to measure coverage and extensiveness metric between CQs and Ontology for e-types and data properties. 31 data properties and 10 e-types overlapped on ontology reference and CQs.

e-type coverage:

$$Cov(CQ_c) = \frac{|CQ_c \cap Ont_c|}{|CQ_c|} = \frac{10}{11} = 0.9 \tag{1}$$

Property coverage:

$$Cov(CQ_p) = \frac{|CQ_p \cap Ont_p|}{|CQ_p|} = \frac{50 - 19}{50} = \frac{31}{50} = 0.62 \tag{2}$$

e-type Extensiveness:

$$Ext(CQ_c) = \frac{|Ont_c - CQ_c|}{|CQ_c \cup Ont_c|} = \frac{12}{24} = 0.5 \tag{3}$$

Properties Extensiveness:

$$Ext(CQ_p) = \frac{|Ont_p - CQ_p|}{|CQ_p \cup Ont_p|} = \frac{250 - 50}{250 + (50 - 31)} = \frac{200}{269} = 0.74 \quad (4)$$

Data level: in this part aim to measure coverage and sparsity metric between CQs and datasets. Dataset include 20 e-types and 208 properties. From the result, we can notice that the coverage metric is very good.

e-type coverage:

$$Cov(D_c) = \frac{|CQ_c \cap D_c|}{|CQ_c|} = \frac{10}{11} = 0.9 \quad (5)$$

Property coverage:

$$Cov(D_p) = \frac{|CQ_p \cap D_p|}{|CQ_p|} = \frac{49}{50} = 0.98 \quad (6)$$

e-type sparsity:

$$Spr(CQ_c) = 1 - \frac{|CQ_c \cap D_c|}{|CQ_c \cup D_c|} = 1 - \frac{10}{29} = 0.65 \quad (7)$$

Property sparsity:

$$Spr(CQ_p) = 1 - \frac{|CQ_p \cap D_p|}{|CQ_p \cup D_p|} = 1 - \frac{49}{209} = 0.76 \quad (8)$$

4 Informal Modeling

This section is dedicated to the description of the informal modeling phase which is a significant phase of iTelos.

- Purpose formalization (informal modeling part) and Modeling sheet description
- ER model description
- Informal Modeling evaluation

Five different data sources are being used for food and accommodation, data sources of Open Data Trentino's attribute names are in Italian. We need to translate them to English according to schema.org. Table 12 shows, how we translate and give data properties names on Protégé for hotel dataset. Table13 shows, attribute name of Osterie Data from Open Data Trentino and food dataset of OpenstreetMap, in this part we need to translate attribute names and match to OpenstreetMap datasets attribute names and give data properties names for ETG in Protégé.

Attribute name of hotel from OpenData Trentino	**Attribute name on English version**
data-inizio-validita	has_starting_date

data-fine-validita	has_expiry_date
denominazione denominazione-ente	has_city_and_town
denominazione-ente-annuario	has_agency
altitudine	has_altitude
cap	has_postal_code
id-localita	has_Internal_ID
tipologia-alberghiera	has_type_of_facility
tipologia-servizio	has_type_of_boarding
recapito-telefono	has_phone_number
recapito-fax	has_fax_number
recapito-email	has_email
recapito-www	has_website
numero-unita	has_number_of_rooms
numero-posti-letto	has_number_of_beds
id-EsercizioRicettivo	has_ID
p-iva	has_ID
comune	has_municipality
frazione indirizzo	has_address
livello-classifica	has_star
prezzo-max-letto-aggiunto	has_max_price_for_extra_bed
prezzo-max-consumazione-pasto	has_max_price_for_meal
struttura-chiusa	has_status
tipologia-unita	has_accommodation_category
posti-letto	has _number_of_beds
prezzo	has_price
codice	has_code_for_service
descrizione	has_description_service

Table 12: Attribute name for Hotel

Attribute name restaurants from OpenStreetMap	**Attribute name**	**Attribute name of Osterie Data from Open Data Trentino**
ID	has_idRestaurant	Nr
Name Amenity	has_name has_amenityFeature	Insegna
Address for street Address for house number	has_address_for_street has_ad- dress_for_house_number	Indirizzo

Postcode	has_postcode	
City	has_city	Comune
Country	has_country	
Brand	has_brand	
Contact email	has_contactEmail	
Contact Phone	has_contactPhone	
Contact website	has_contactWebsite	
Contact Facebook	has_contactFacebook	
Outdoor seating	has_outdoorSeating	
Takeaway	has_takeaway	
Opening hours	has_openingHours	
Cuisine	has_servesCuisine	
Ref - Vatin	has_Ref_Vat	
Operator	has_operator	
Wheelchair	has_wheelchair	
Surveillance	has_surveillance	
Brewery	has_brewery	
Diet - vegan	has_dietVegan	
Diet - vegetarian	has_dietVegetarian	
Diet - gluten free	has_glutenFree	
Internet access	has_internetAccess	
Smoking	has_smoking	
latitude	has_lattitude	
longitude	has_longitude	
Source	has_source	
stars	has_starRating	

Table 13: Attribute name for Foo

4.1 Purpose formalization (informal modeling part) and Modeling sheet description

In this phase reported CQs classified three kind of e-types. Core e-type is defined a specific e-type that includes particular attributes important to the competency queries given as input from the previous phase. The common e-types involves the classes for those entities that are used to represent common aspect of the world to represent, such as space and time aspects. The contextual e-types involves the classes for those entities that are used to represent specific aspects of the problem to solve.

We are describing essential e-types and their corresponding properties from CQs. The extracted e-types and properties are basic of ER model. In this phase has a two main part, data layer and knowledge layer which is designed develop CQs to classified and attributed CQs. Classified CQs have objects, functions, and actions, which are the components of the foundational teleology (FT). The accomplishment of the FT means it's closer to Entity Type Graph(ETG)

since ETG is domain- specific top-level categories FT.

Hence our topic is food and accommodation, CQs based on information of locations we considered common e-type is location.

For Food, restaurant, pub,ice-cream shop, cafe, fast food and bar are core e-types which is involving the main important information. Table 14 shows which is object and function.

For Accommodation, there are three datasets that are hotels, non-hotel from OpenData Trentino and Airbnb from Inside Airbnb, dataset of hotel and non-hotel have similar attribute names but non-hotel data is for private rooms and apartments and camping. Thus we matched non-hotel data with the Airbnb data. Since apartment and private room data properties similar we compressed to one e-type which is apartment.

	Common Kernel Concepts	**Core Kernel Concepts**		
NUM of persona	Function	Object	Function	Action
1.1	location	Accommodation	Hotel	
1.2	location	Accommodation	Hotel	
1.3	location	Food	Restaurant	
1.4	location	Food	Restaurant	
1.5	location	Food	Restaurant	
1.6	location	Accommodation	Hotel	
1.7	location	Accommodation	Hotel	
1.8	location	Food	Pub	
1.9	location	Food	Cafe	
2.1	location	Accommodation	Hotel	
2.2	location	Accommodation	Hotel	
2.3	location	Accommodation	Bed and breakfast	
2.4	location	Food	Restaurant	
2.5	location	Food	Restaurant	
2.6	location	Food	bar	
2.7	location	Food	fast_food_restaurant	
2.8	location	Food	Ice_cream-shop	
3.1	location	Accommodation	Apartment	
3.2	location	Accommodation	Apartment	
3.3	location	Accommodation	Apartment	
3.4	location	Accommodation	Apartment	
3.5	location	Food	Restaurant	
3.6	location	Food	Restaurant	
3.7	location	Food	cafe	
3.8	location	Food	pub	
4.1	location	Accommodation	Camping	
4.2	location	Accommodation	Hotel	

4.3	location	Accommodation	Camping	
4.4	location	Accommodation	Camping	
4.5	location	Food	fast_food_restau-rant	
f	location	Food	fast_food_restau-rant	
4.7	location	Food	fast_food_restau-rant	
4.8	location	Food	bar	
5.1	location	Accommodation	Hotel	
5.2	location	Accommodation	Hotel	
5.3	location	Food	Restaurant	
5.4	location	Food	Restaurant	
5.5	location	Food	pub	
5.6	location	Food	Restaurant	
6.1	location	Accommodation	Hotel	
6.2	location	Accommodation	Camping	
6.3	location	Accommodation	Hotel	
6.4	location	Food	fast_food_restau-rant	
6.5	location	Food	Restaurant	
6.6	location	Food	Restaurant	
6.7	location	Accommodation	Apartment	
6.8	location	Accommodation	Bed and break-fast	
7.1	location	Accommodation	Hotel	
7.2	location	Accommodation	Hotel	
7.3	location	Food	Restaurant	
7.4	location	Food	cafe	
7.5	location	Food	Restaurant	
7.6	location	Food	bar	
7.7	location	Accommodation	Apartment	
7.8	location	Accommodation	Apartment	
8.1	location	Accommodation	Hotel	
8.2	location	Food	Restaurant	
8.3	location	Accommodation	Hotel	
8.4	location	Accommodation	Hotel	
8.5	location	Food	Ice_cream-shop	
8.6	location	Accommodation	Apartment	

Table 14: Informal Modeling sheet

Table 15 shows the E-types, data properties and object properties which show the relations

between entity types. In our case Accommodation is super class that connects to e-types Apartment, Camping and Hotel with has_hotel and has_ apartment and has_camping. Another super class is Establishment that connects to sub e-types(restaurant, bar ..etc) by has_restaurant, has_bar. Each e-type has data properties which involves relation between entity and their data properties value. According to CQs, we shows data properties for example has_diet_vegan is data property of restaurant e-type. That extracted table is base of ER model.

NUM	**e-type**	**Object Propertiies**	**Data Properties**
1.1	Hotel	has_hotel	
1.2	Hotel	has_hotel	has_price
1.3	Restaurant	has_restaurant	has_diet_vegan, has_email
1.4	Restaurant	has_restaurant	has_diet_vegan, has_internet_access, has_city
1.5	Restaurant	has_restaurant	has_outdoor_seating, has_phone_number
1.6	Hotel	has_hotel	has_phone_number
1.7	Hotel	has_hotel	has_price
1.8	Pub	has_pub	has_open_hours
1.9	Cafe	has_cafe	has_internet access, has_takeaway
2.1	Hotel	has_hotel	has_description_service, has_starsRating
2.2	Hotel	has_hotel	has_description_service
2.3	Apartment	has_apartment	has_Description
2.4	Restaurant	has_restaurant	has_outdoor_seating, has_brewery, has_city
2.5	Restaurant	has_restaurant	has_outdoor_seating, has_brewery
2.6	bar	has_bar	has_street_bar, has_house_number_bar, has_postcode_bar, has_city_bar
2.7	fast_food_restaurant	has_fast_food_restaurant	has_cuisine_fb, has_city, has_phone_number_fb
2.8	Ice_cream-shop	has_Ice_cream-shop	has_frozen_yougurt, has_outdoor_seating_IC
3.1	Apartment	has_apartment	has_Amenities
3.2	Apartment	has_apartment	has_price
3.3	Apartment	has_apartment	has_Amenities
3.4	Apartment	has_apartment	has_Amenities, has_ Neighborhood Overview
3.5	Restaurant	has_restaurant	has_delivery, has_city

3.6	Restaurant	has_restaurant	has_wheel_chair
3.7	cafe	has_cafe	has_wheel_chair_cafe, has_vegan_cafe
3.8	pub	has_pub	has_outdoor-seating_pub
4.1	Camping	has_camping	has_service
4.2	Hotel	has_hotel	has_type_boarding
4.3	Camping	has_camping	has_service
4.4	Camping	has_camping	has_phone_number
4.5	fast_food_restaurant	has_fast_food_restaurant	has_takeaway_fb
4.6	fast_food_restaurant	has_fast_food_restaurant	has_cuisine_fb
4.7	fast_food_restaurant	has_fast_food_restaurant	has_delivery_fb, has_phone_number_fb
4.8	bar	has_bar	has_smoking_bar
5.1	Hotel	has_hotel	
5.2	Hotel	has_hotel	
5.3	Restaurant	has_restaurant	has_cuisine
5.4	Restaurant	has_restaurant	has_cuisine
5.5	pub	has_pub	
5.6	Restaurant	has_restaurant	has_cuisine
6.1	Hotel	has_hotel	
6.2	Camping	has_camping	
6.3	Hotel	has_hotel	
6.4	fast_food_restaurant	has_fast_food_restaurant	has_opening_hours_fb
6.5	Restaurant	has_restaurant	has_outdoor_seating, has_smoking
6.6	Restaurant	has_restaurant	has_contact_website
6.7	Apartment	has_apartment	has_accommodates
6.8	Apartment	has_apartment	
7.1	Hotel	has_hotel	
7.2	Hotel	has_hotel	
7.3	Restaurant	has_restaurant	has_cuisine
7.4	cafe	has_cafe	
7.5	Restaurant	has_restaurant	has_opening_hours
7.6	bar	has_bar	
7.7	Apartment	has_apartment	has_number_of_reviews
7.8	Apartment	has_apartment	has_number_of_reviews, has_price
8.1	Hotel	has_hotel	has_description_service
8.2	Restaurant	has_restaurant	has_opening_hours
8.3	Hotel	has_hotel	
8.4	Hotel	has_hotel	has_price_hotels
8.5	Ice_cream-shop	has_Ice_cream-shop	

8.6	Apartment	has_apartment	has_minimum_nights

Table 15: Attributed CQs

4.2 ER Model

We are using yED tool for making ER Model. yED tool is available to show deeply structure of objects based on schema level. According to Table 15 three components which are e-type, data properties and object properties in the domain of food and accommodation can be assembled in to ER model. Figure 1 is illustrating common and core e-types. Common e-type is location it has data properties latitude and longitude with their respective data type, object property has_location is connecting Location to FoodEstablishment, Bed_and_breakfast and Apartment e-types.

Core e-types, Accommodation is super class (e-types) which has a sub classes that are apartment, camping and hotel. Food is another super class(e-type) which sub classes that are pub, bar, cafe, restaurant, ice-cream shop, fast food. Each e-type include attribute names(data properties) with data type and object properties that connect between classes. For instance cafe is e-type, data properties is outdoor_seating, object properties has_cafe.

4.3 Informal Modeling evaluation

In this section reported evaluation phase which calculated for schema level and data level. CQs has 11 e-types and 50 properties. ER model include 13 e-types and 75 data properties.

Schema level: aim to measure coverage and extensiveness metric between CQs and ER model for e-types and data properties.

e-type coverage:

$$Cov(CQ_c) = \frac{|CQ_c \cap ER_c|}{|CQ_c|} = \frac{11}{11} = 1 \tag{9}$$

Property coverage:

$$Cov(CQ_p) = \frac{|CQ_p \cap ER_p|}{|CQ_p|} = \frac{50-1}{50} = \frac{50-1}{50} = 0.98 \tag{10}$$

e-type Extensiveness:

$$Ext(CQ_c) = \frac{|ER_c - CQ_c|}{|CQ_c \cup ER_c|} = \frac{2}{13} = 0.16 \tag{11}$$

Properties Extensiveness:

$$Ext(CQ_p) = \frac{|ER_p - CQ_p|}{|CQ_p \cup ER_p|} = \frac{75-49}{75+1} = \frac{26}{76} = 0.34 \tag{12}$$

Data level: in this part aim to measure coverage and sparsity metric between ER model and Data set.

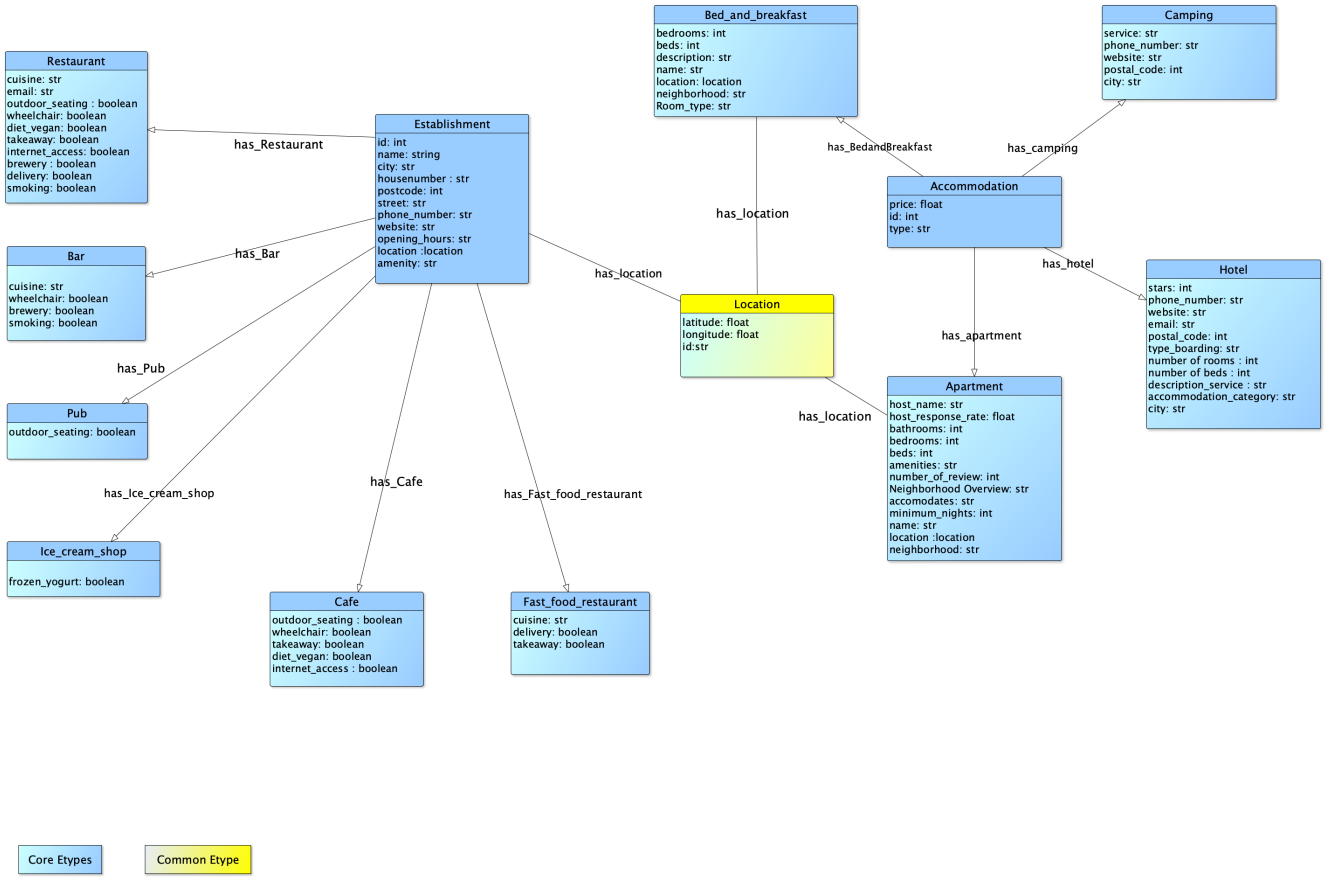


Figure 1: ER model

e-type coverage:

$$Cov(D_c) = \frac{|ER_c \cap D_c|}{|D_c|} = \frac{10}{28} = 0.3 \tag{13}$$

Property coverage:

$$Cov(D_p) = \frac{|ER_p \cap D_p|}{|D_p|} = \frac{71}{208} = 0.3 \tag{14}$$

e-type sparsity:

$$Spr(D_c) = 1 - \frac{|ER_c \cap D_c|}{|ER_c \cup D_c|} = 1 - \frac{10}{31} = 0.67 \tag{15}$$

Property sparsity:

$$Spr(CQ_p) = 1 - \frac{|ER_p \cap D_p|}{||ER_p \cup D_p||} = 1 - \frac{75}{208} = 0.63 \tag{16}$$

5 Formal Modeling

This section is dedicate to the description of the formal modeling phase. According to EER model, we are making ETG on Protégé. Evaluation phase, three metrics of coverage and sparsity and extensiveness measured for e-types and data properties between ETG and Ontology. Basically we have thirteen e-types which are Accommodation, Location, FoodEstablishment, Apartment, Bed_and_breakfast, Camping, Hotel, Bar, Cafe, Fast_food_restaurant, Ice_cream_shop, Pub, Restaurant. Ontology reference has restaurant, hotel and accommodation, we compared data properties.

More in details, this section provides a description of the following activities:

- ETG generation
- Data management (syntactic heterogeneity)
- Formal Modeling evaluation

5.1 ETG generation

The ETG Generation activity is internally defined by three sub activities:

- **Ontology Selection**: As ontologies are meant to specify knowledge about some domain, the process of creating an ontology can be seen as transferring knowledge into a computer-accessible form. Clearly, there are several possible sources of the knowledge to be formalized. These might be categorized with respect to the extent to which they are already accessible to computer systems, and, more precisely, to what extent the structure of the current representation of the provided knowledge can be exploited to facilitate the formalization process (i.e. the process of making the inherent semantics formally explicit). Furthermore, selection of those ontologies which includes appropriate concepts which can be reused to

model the ER.Table 16 shows e-types ontology reference from Schema.org, consumer of the object Establishment having several specialized functions such as Bakery, BarOrPub, Brewery etc, also object of Accommodation has several specialized functions apartment, room. Schema.org is one of the most common reference standard format and it can be regarded as a useful schema for our ER model. According to this table and based on our CQs and datasets, we choose e-types and data properties in ER model and ETG generation.

Accommodation	Lodging Businesses	FoodEstablishment
Apartment	BedAndBreakfast	Bakery
CampingPitch	Campground	BarOrPub
Room	Hostel	Brewery
HotelRoom	Hotel	CafeOrCoffeeShop
MeetingRoom	Motel	Distillery
House	Resort	FastFoodRestaurant
Suite		IceCreamShop

Table 16: Ontology reference from Schema.org

- **Language Alignment**: transforms the informal concepts of the ETG model to formal concepts with the key support of the UKC. UKC is a developing platform offering numerous languages and concepts for use in the field of knowledge integration. Besides, we need to explain conceptual diversity (L1) and language diversity (L2). Conceptual diversity is The concepts are not unique in the model. Language diversity indicates that different languages describe the same concepts. UKC is the specialized tool to address these issues. Furthermore, If ETG model concepts already exist with the UKC concept, KOS application provides a matched unique identifier (GID)from UKC with rendering each concept formal and absorbing L1 diversity. Otherwise we have to create a new concept with new GID absorbing L1 and L2 diversity. KOS performed each term of classes and their attribute names, all terms from all these hierarchies are considered. In our project most of concept is already existed with UKC concept, several concept is we need to create them to get with new GID.
- **Schema Building**: that activity aims to exactly close the gap between the ETG model and the foundational teleology. It formally grounds the ETG model in the foundational teleology, thus producing the final, fully formal ETG ready for the next phase - data integration. Schema alignment involves aligning the objects, functions, actions and relations in the ETG model to their semantically corresponding foundational primitives via intermediate concepts. In our case, we have non living objects that are Location, Food Establishment and Accommodation, producer function apartment, hotel, restaurant, cafe etc...but do not have action.

ETG model

- According to the ER model, we created the ETG model in Protégé. Protégé is a free, open-source platform that provides a growing user community with a suite of tools to construct domain models and knowledge-based applications with ontologies. Using this software

we have been able to highlight different categories for example classes, data types, data properties, object properties. Figure 2 shows e-types(classes) in the ETG model in Protégé. Each class has been set with a global identifier(GID) which are unique numbers illustrating the concepts.

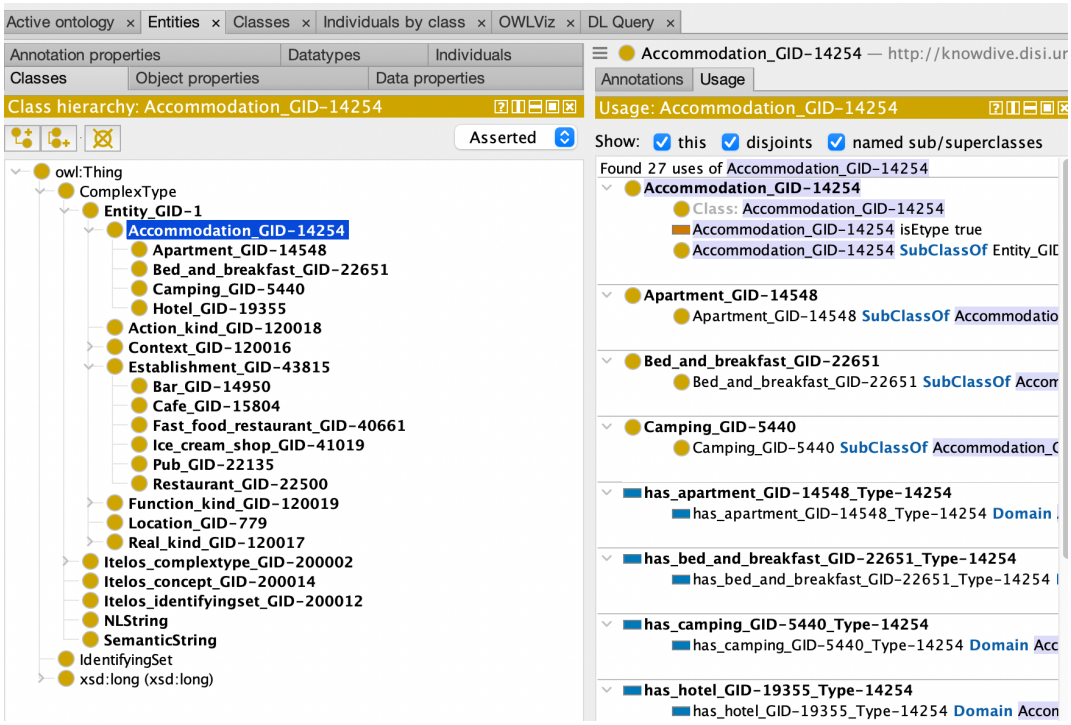


Figure 2: e-types in ETG

- OWL distinguishes between two main categories of properties that an ontology builder may want to define:
  - **Object properties**: link individuals to individuals.
  - **Datatype properties**: link individuals to data values.

  An object property is defined as an instance of the built-in OWL class owl:ObjectProperty. A datatype property is defined as an instance of the built-in OWL class owl:DatatypeProperty. Both owl:ObjectProperty and owl:DatatypeProperty are subclasses of the RDF class rdf:Property.

  Figure 3 shows, object properties linked between two domain and range (e-types). Each object property is set with concept Global IDentifier (GID) and class of GID.

  Domain: for a property one can define (multiple) rdfs:domain axioms. Syntactically, rdfs:domain is a built-in property that links a property (some instance of the class rdf:Property) to a class description. An rdfs:domain axiom asserts that the subjects of such property statements must belong to the class extension of the indicated class description.

  Multiple rdfs:domain axioms are allowed and should be interpreted as a conjunction: these restrict the domain of the property to those individuals that belong to the intersection of the

class descriptions. If one would want to say that multiple classes can act as domain, one should use a class description of the owl:unionOf form.

Range: for a property one can define (multiple) rdfs:range axioms. Syntactically, rdfs:range is a built-in property that links a property (some instance of the class rdf:Property) to to either a class description or a data range. An rdfs:range axiom asserts that the values of this property must belong to the class extension of the class description or to data values in the specified data range.

Multiple range restrictions are interpreted as stating that the range of the property is the intersection of all ranges (i.e., the intersection of the class extension of the class descriptions c.q. the intersection of the data ranges). Similar to rdfs:domain, multiple alternative ranges can be specified by using a class description of the owl:unionOf form .

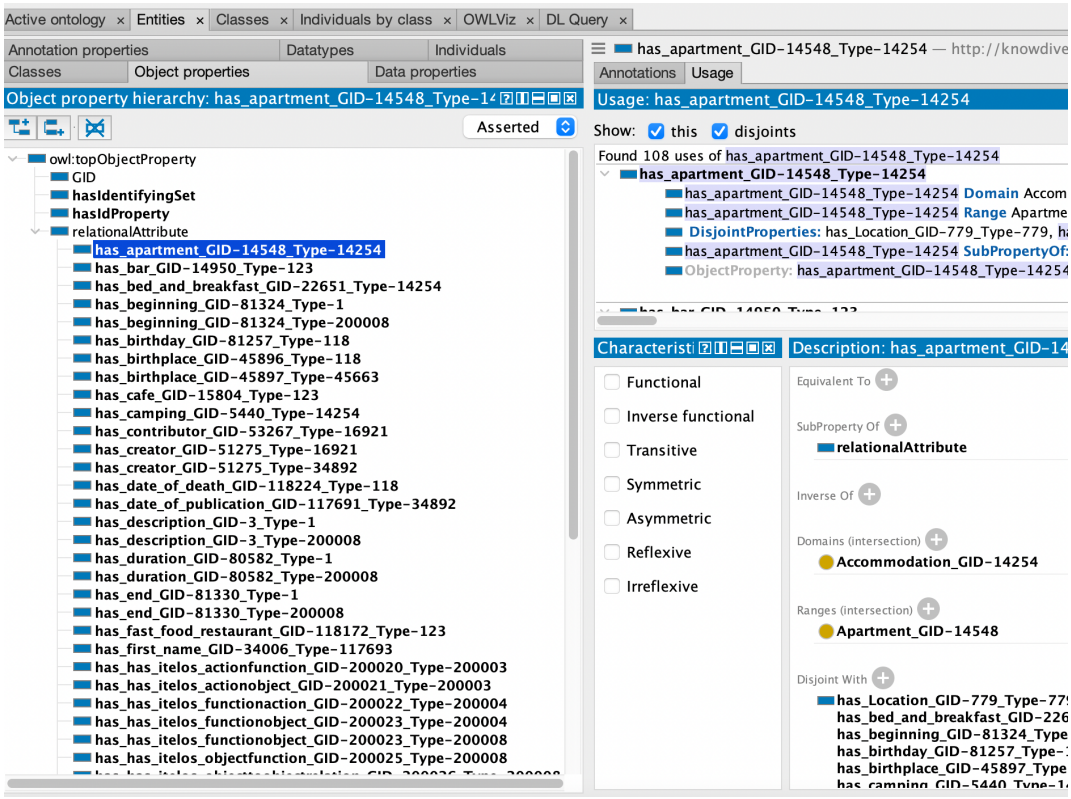


Figure 3: Object properties in ETG

- Figure 4 shows, data properties linked between entity and data value. Domain is e-type and range is data type. Each data properties set with concept Global IDentifier (GID) and GID of class.

5.2 Data management (syntactic heterogeneity)

In terms of data management, we have 5 different datasets. Firstly, we need to clean data and choose data properties which is overlapped with ER model. Hotel and non hotel dataset on the Italian language, need to translate to English, specially description services and attribute

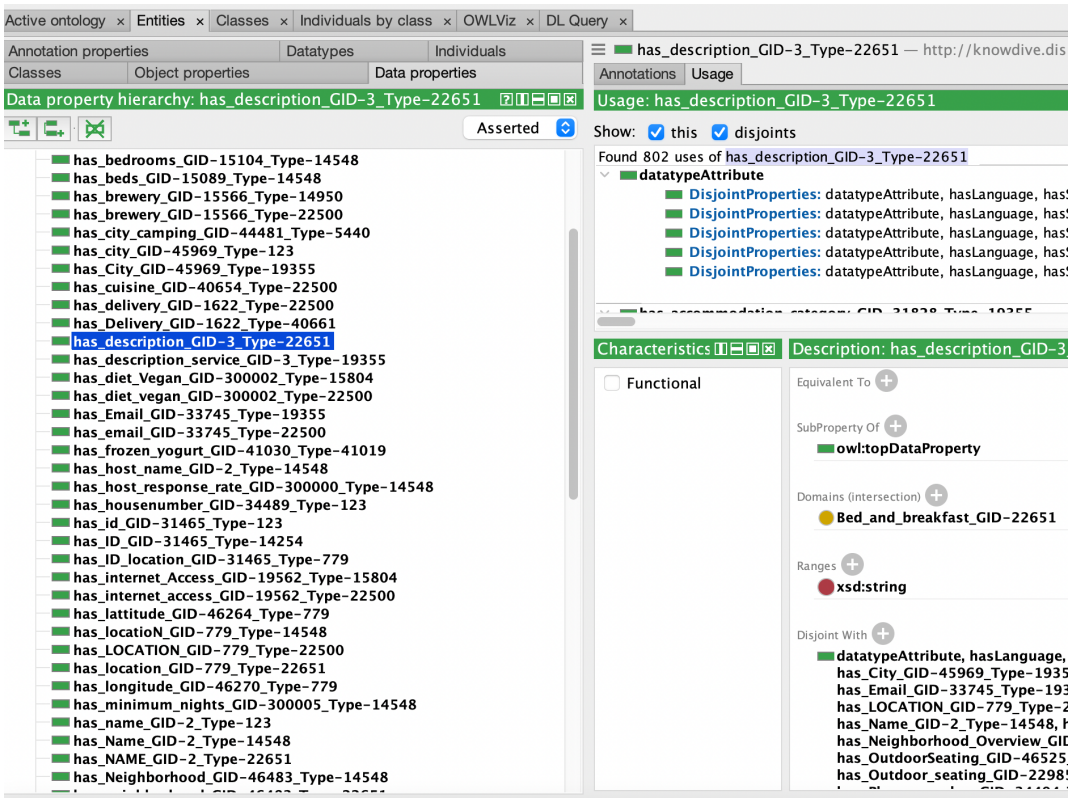


Figure 4: Data properties in ETG

names. When we first download data, there are $XMl, json, csv$ format, initial step was convert to csv for clean and translate data, afterward convert to $json$ file for Karmalinker.

5.3 Formal Modeling evaluation

In this section we aim to measure coverage, sparsity and extensiveness metrics between ETG and Ontology for e-types and data properties. ETG has 13 e-types and 75 properties, Ontology reference consist of 23 e-type and 250 properties. In our case 12 e-types and 43 properties overlapped on ETG and Ontology reference.

e-type coverage:

$$Cov(ETG_c) = \frac{|ETG_c \cap Ont_c|}{|ETG_c|} = \frac{12}{13} = 0.92 \tag{17}$$

Property coverage:

$$Cov(ETG_p) = \frac{|ETG_p \cap Ont_p|}{|ETG_p|} = \frac{43}{75} = 0.57 \tag{18}$$

e-type Extensiveness:

$$Ext(ETG_c) = \frac{|Ont_c - ETG_c|}{|ETG_c \cup Ont_c|} = \frac{11}{24} = 0.45 \tag{19}$$

Properties Extensiveness:

$$Ext(ETG_p) = \frac{|Ont_p - ETG_p|}{|ETG_p \cup Ont_p|} = \frac{250 - 75}{250 + (75 - 43)} = \frac{175}{282} = 0.62 \quad (20)$$

e-type sparsity:

$$Spr(ETG_c) = 1 - \frac{|ETG_c \cap Ont_c|}{|ETG_c \cup Ont_c|} = 1 - \frac{12}{24} = 0.5 \quad (21)$$

Property sparsity:

$$Spr(ETG_p) = 1 - \frac{|ETG_p \cap Ont_p|}{||ETG_p \cup Ont_p||} = 1 - \frac{43}{282} = 0.85 \quad (22)$$

6 Data Integration

This section is dedicated to the description of the data integration phase. Like in the previous section, the current one aims to describe the different sub activities performed by all the team members, as well as the phase outcomes produced.

More in details, this section provides a description of the following activities:

- Data management (semantic heterogeneity)
- Entity matching
- Data integration phase evaluation

6.1 Data management (semantic heterogeneity)

In terms of data management, we have 5 different datasets. 1 dataset represents the hotels, 2 datasets represent the non-hotels and Airbnbs and 2 datasets represent the restaurants, bars and osterie. When we checked the structure of non-hotel and Airbnb data, we observed that there are some data in the non-hotel data we gathered from Open Data Trentino that do not exist in the Airbnb data. This may be because our Airbnb data was scraped earlier than the Trentino data. In this case, we planned to integrate them by adding the non-existent data from the other dataset directly. This would cause us to leave some columns empty since Open Data Trentino does not provide all the detailed information that InsideAirbnb gives. For the intersecting non-hotel properties(apartments) in both datasets, we added the attributes together. To understand that two properties were the same, we checked their names, phone number and address information. Since address might be written in a different format, we only checked the street name.

Secondly, for the food places such as restaurants, bars, pubs, we have two datasets. One of them is from Open Street Map and it includes very detailed attributes. For many of the restaurants, some attributes remain missing. Our other dataset is from Open Data Trentino and it is about the Osterie. So we completed the missing attributes in the Open Street Map data by using

the data in the Open Data Trentino dataset. We compared two records by their names, address and regions to understand if they represent the same place.

6.2 Entity Matching

As a final step in the Itelos methodology, we should perform entity matching which includes identifying if different entities in our datasets are actually the same real world entity and thus they should be merged in our final verison of the entity graph.

Firstly, we should observe the identifying set which is the set of E-type properties. It may not be possible to have a unique identifier all the time so we may have to identify a unique entity by checking the value of other data properties. We performed this operation in Karmalinker which allowed us to use the ontology we've previously defined to integrate the datasets, also solving the semantic heterogeneity.

As a final product, we had the reusable model files from Karmalinker for every E-type. It is possible to observe the results in detail in Figures 5, 6.

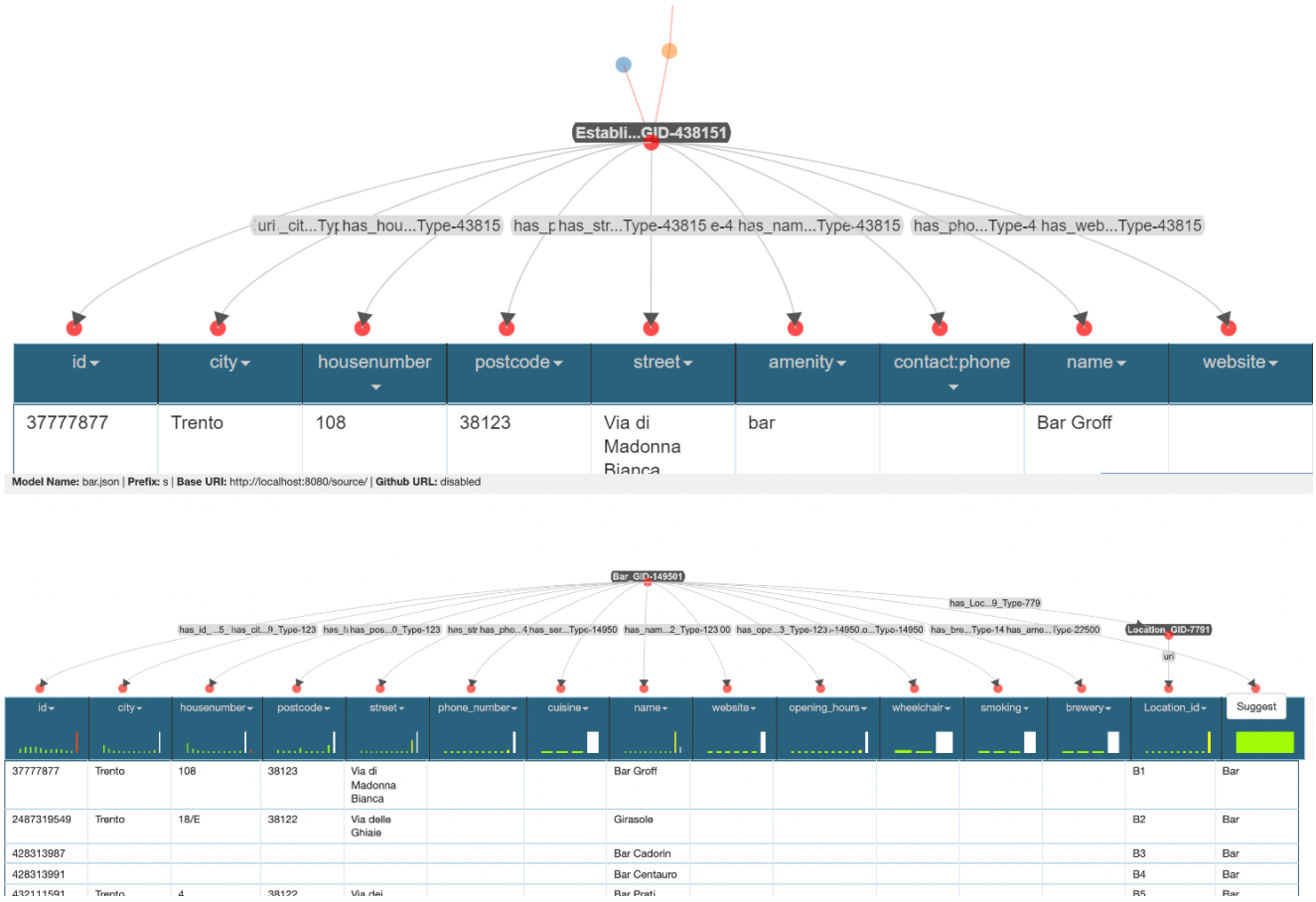


Figure 5: Etype of Bar and Establishment in Karmalinker

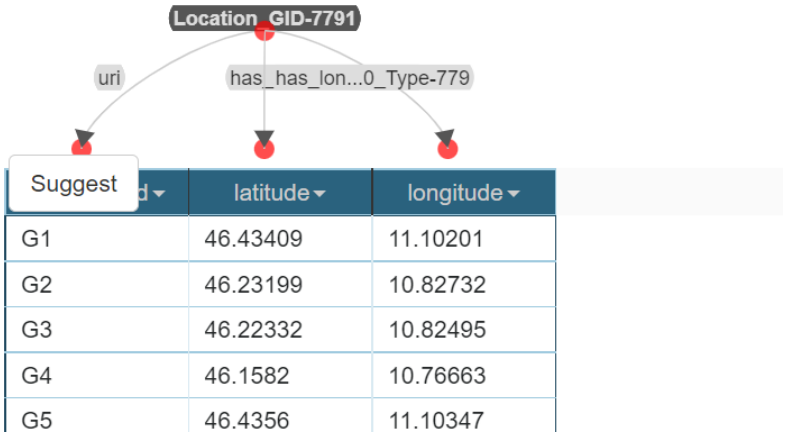


Figure 6: Location in Karmalinker

6.3 Data Integration phase evaluation

During Data Integration phase, we evaluate on data level. We have the proposed final EG. We have two main evaluation part, the first one is if the CQs in inception phase can be answered by our constructed EG. We can do evaluation based on practial applications, like SQL. The model can be answer all CQs,it has been showing more detail in last section.

Another one is if our collected dataset is sufficiently used. By using Sparsity to check if the dataset schema is aligned to ETG properties.

Formal ETG vs Dataset Schema e-type sparsity:

$$Spr(DS_c) = 1 - \frac{|DS_c \cap ETG_c|}{|DS_c \cup ETG_c|} = 1 - \frac{13}{13} = 0 \tag{23}$$

Property sparsity:

$$Spr(DS_p) = 1 - \frac{|DS_p \cap ETG_p|}{||DS_p \cup ETG_p||} = 1 - \frac{75}{75} = 0 \tag{24}$$

7 Outcome exploitation

The final section of the current document aims to provide a description of the data integration process outcome. Here have to be reported the final Knowledge Graph (KB) information (like, number of etypes and properties, number of entities for each etype, and so on). Moreover this section has to provide a description for the KG possible exploitation.

In the end of this section, some general conclusions can be added, describing considerations reached along the integration project, as well as any eventual open issue remained open due to lack of correct solution in the data integration context considered.

7.1 Knowledge Graph Information

In terms of the project, the final step is building the knowledge Graph. We successfully created Knowledge graph for Food and Accommodation, which includes etypes and properties and entities Table17 shows. Unfortunately, we had one problem in Karmalinker, which was limited number of entities were accepted for one etype. Our data for Accommodation is huge, therefore we used for this Knowledge graph smaller partitions of entities.

	Etype	**Property**	**Entity**
	Establishment	12	1461
	Restaurant	10	469
	Bar	4	425
	Café	5	301
	Ice_cream shop	1	79
	Fast food restaurant	3	149
	Pub	1	43
	Location	3	5843
	Accommodation	3	19803
	Apartment	13	4060
	Hotel	11	14803
	Camping	5	569
	Bed and breakfast	7	333
Total	13	75	48338

Table 17: Information of Etypes

7.2 Model Applications

- **GraphDB**: designed as an enterprise-grade semantic repository system, suitable for massive volumes of data, file-based indices (enables it to scale to billions of statements even on desktop machines) and offer inference and query optimizations (ensures fast query evaluations).
- **SPARQL**: RDF is a directed, labeled graph data format for representing information in the Web. This specification defines the syntax and semantics of the SPARQL query language for RDF. SPARQL can be used to express queries across diverse data sources, whether the data is stored natively as RDF or viewed as RDF via middleware. SPARQL contains capabilities for querying required and optional graph patterns along with their conjunctions and disjunctions. SPARQL also supports aggregation, subqueries, negation, creating values by expressions, extensible value testing, and constraining queries by source RDF graph. The results of SPARQL queries can be result sets or RDF graphs.

7.3 Query

The previous stage of Karmalinker makes a rdf(ttl) file for every Etype. In this part, we applied our model to some competency questions using SPARQL in GraphDB based on the rdf file. The result was properly working. More details are shown in the following Figures.

7.3.1 Query 1

Give the list of bars near the city center in Trento? Return: list of bars with name, house number, street, postcode, city, latitude and longitude. In this case we show the just first 4 example of bars.

```
PREFIX iri: <http://knowdive.disi.unitn.it/etype#>
PREFIX omgeo: <http://www.ontotext.com/owlim/geo#>
CONSTRUCT{      ?Establishment a iri:Establishment_GID-43815 ;
                iri:has_Location_GID-779_Type-779 ?Location;
                iri:has_name_GID-2_Type-43815 ?name;
                iri:has_amenity_GID-2973_Type-43815 ?amenity;
                iri:has_housenumber_GID-34489_Type-43815 ?housenumber;
                iri:has_street_GID-24034_Type-43815 ?street;
                iri:has_postcode_GID-34110_Type-43815 ?postcode;
                iri:has_city_GID-45969_Type-43815 ?city.

   ?Establishment a iri:Bar_GID-14950;
                  iri:has_WheelChair_GID-25446_Type-14950?wheelchair;
          iri:has_Smoking_GID-23491_Type-14950?smoking.

?Location  a iri:Location_GID-779;
                iri:has_lattitude_GID-46264_Type-779 ?latitude;
          iri:has_longitude_GID-46270_Type-779 ?longitude.
}
WHERE{
    { ?Establishment a iri:Establishment_GID-43815}
    UNION
    {?Establishment a iri:Bar_GID-14950}
           ?Establishment iri:has_Location_GID-779_Type-779 ?Location;
                iri:has_name_GID-2_Type-43815 ?name;
                iri:has_amenity_GID-2973_Type-43815 ?amenity;
                iri:has_housenumber_GID-34489_Type-43815 ?housenumber;
                iri:has_street_GID-24034_Type-43815 ?street;
                iri:has_postcode_GID-34110_Type-43815 ?postcode;
                iri:has_city_GID-45969_Type-43815 ?city.
                                        filter(?amenity = "Bar")
                        filter(?city="Trento")
    OPTIONAL{
         ?Establishment iri:has_WheelChair_GID-25446_Type-14950?wheelchair;
          iri:has_Smoking_GID-23491_Type-14950?smoking.
             filter(?wheelchair = "yes")
    }
?Location iri:has_lattitude_GID-46264_Type-779 ?latitude;
          iri:has_longitude_GID-46270_Type-779 ?longitude.
     BIND (omgeo:distance(46.0671931, 11.1212042, ?latitude, ?longitude) AS ?distance)

             FILTER (?distance < 5)

}
LIMIT 4
```

Figure 7: Query 1

	name	amenity	housenumber	street	postcode	city	latitude	longitude
1	"Girasole"	"Bar"	"18/E"	"Via delle Ghiaie"	"38122"	"Trento"	"46.0554208"	"11.1161172"
2	"Bar Prati"	"Bar"	"4"	"Via dei Muredei"	"38122"	"Trento"	"46.0572157"	"11.1180954"
3	"pausa kaffè"	"Bar"	"45"	"Via Giuseppe Giu sti"	"38122"	"Trento"	"46.0589545"	"11.1190969"
4	"Bar Florida"	"Bar"	"134"	"Via Vittorio Ven- eto"	"38122"	"Trento"	"46.0589998"	"11.1250541"

Figure 8: Result of query 1

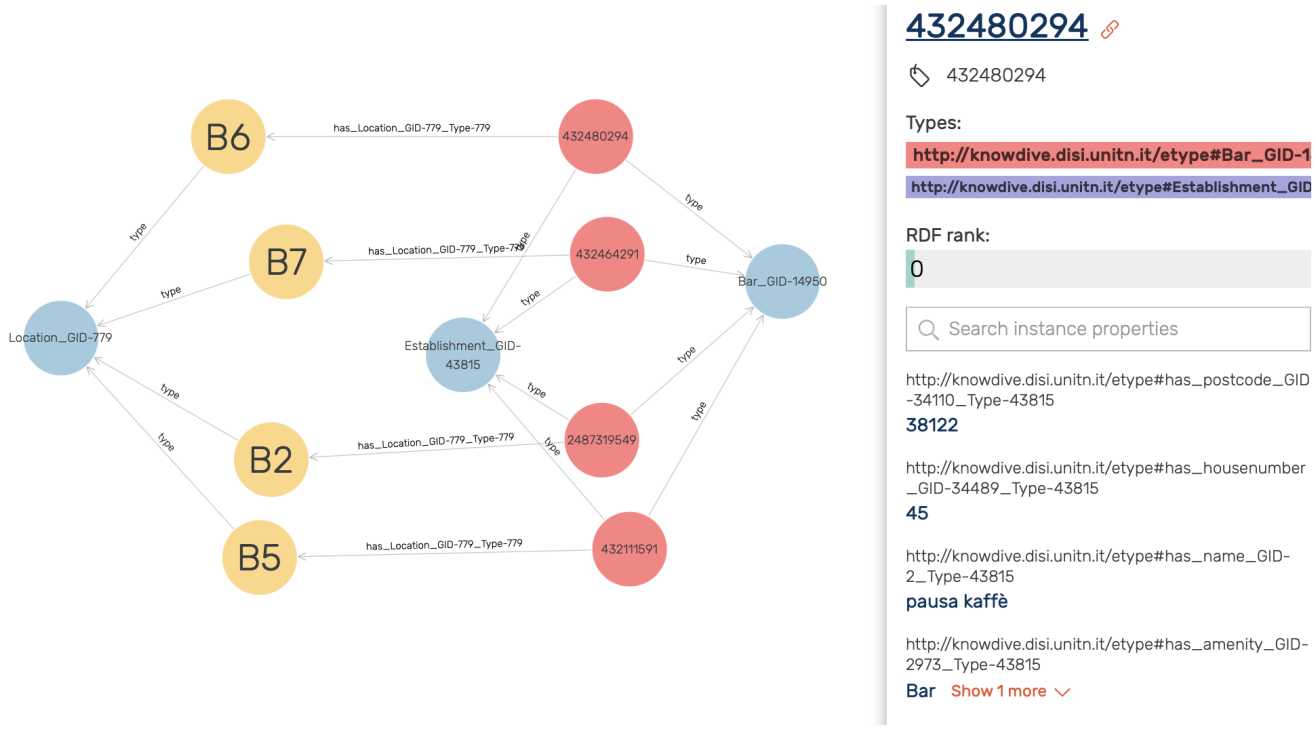


Figure 9: Knowledge graph of query 1

7.3.2 Query 2

Give the list of apartments under 80 dollar in Trentino? Return list of apartments with data properties in Trentino.

```
PREFIX iri: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
CONSTRUCT{
     ?Apartment a iri:Apartment_GID-14548;
        iri:has_Location_GID-779_Type-779 ?Location;
                    iri:has_Location_GID-779_Type-779?Location;
                    iri:has_Name_GID-2_Type-14548?Name;
                        iri:has_price_GID-70571_Type-14254?price;
                        iri:has_beds_GID-15089_Type-14548?beds;
                    iri:has_number_of_review_GID-34489_Type-14548?review;
                            iri:has_Neighborhood_GID-46483_Type-14548?city.

    ?Location a iri:Location_GID-779;
            iri:has_lattitude_GID-46264_Type-779 ?latitude;
            iri:has_longitude_GID-46270_Type-779 ?longitude.
}

where {  {?Apartment a iri:Apartment_GID-14548}
    UNION
    {?Apartment a iri:Accommodation_GID-14254}
         ?Apartment iri:has_Location_GID-779_Type-779 ?Location;
                    iri:has_Location_GID-779_Type-779?Location;
                    iri:has_Name_GID-2_Type-14548?Name;
                        iri:has_price_GID-70571_Type-14254?price;
                        iri:has_beds_GID-15089_Type-14548?beds;
        iri:has_number_of_review_GID-34489_Type-14548?review;
                            iri:has_Location_GID-779_Type-779?Location;
                    iri:has_Neighborhood_GID-46483_Type-14548?city.
                                    FILTER(xsd:integer(?price) < 100)
                        filter contains(?city,"Trentino")

   ?Location iri:has_lattitude_GID-46264_Type-779 ?latitude;
          iri:has_longitude_GID-46270_Type-779 ?longitude.
}
LIMIT 3
```

Figure 10: Query 2

	Name	price	beds	host_response	Neighborhood	latitude	longitude
1	"un posto in paradiso "	"68"	"4"	"100%"	"Vigo di fassa, Trentino, Italy"	"46.42097"	"11.67438"
2	"SKIING AND SPAS I N PEIO"	"60"	"3"	"67%"	"Peio, Trentino-Alto Adige/S√°dtirol, Italy "	"46.36243"	"10.67768"
3	"BILO4 A MARILL-EVA900 SCI E NON S OLO"	"69"	"4"	"100%"	"Marilleva 900, Trentino-Alto Adige/S√°dtirol, Italy"	"46.31381"	"10.80984"

Figure 11: Result of query 2

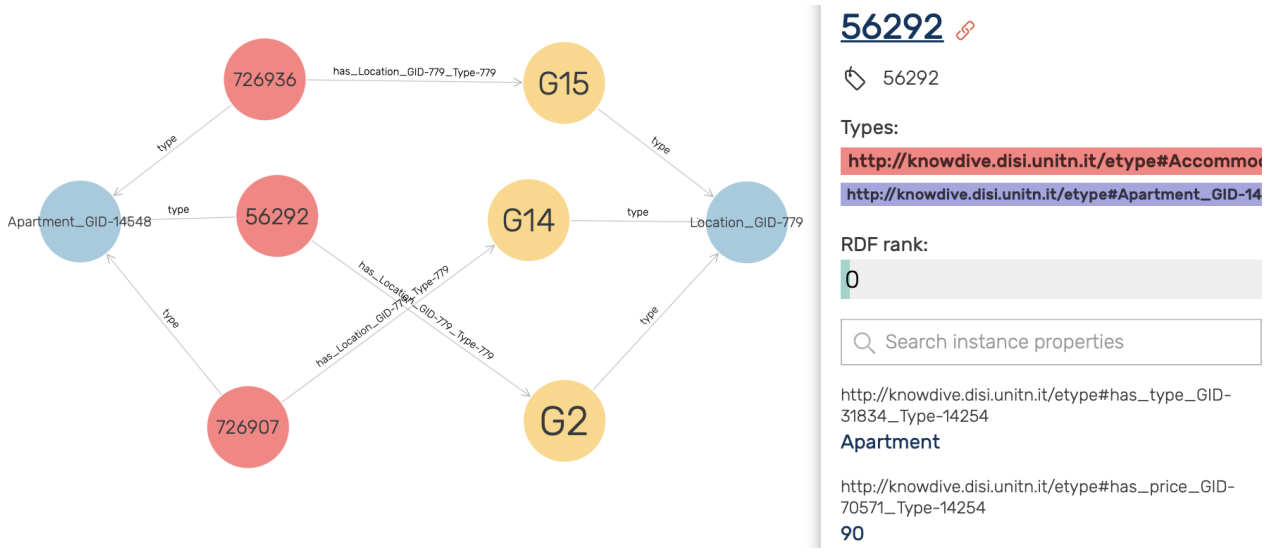


Figure 12: Knowledge graph of query 2

7.3.3 Query 3

Give the list of hotels with full-board and terrace in Andalo? Return list of hotels with data properties in Andalo.

```
PREFIX iri: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
CONSTRUCT{
        ?Accommodation a iri:Accommodation_GID-14254;
      iri:has_type_GID-31834_Type-14254?type.

     ?Accommodation a iri:Hotel_GID-19355;
           iri:has_website_GID-34126_Type-19355?website;
           iri:has_type_boarding_GID-300001_Type-19355?boarding;
           iri:has_number_of_rooms_GID-22651_Type-19355?rooms;
           iri:has_City_GID-45969_Type-19355?city;
           iri:has_Email_GID-33745_Type-19355?email;
           iri:has_phone_number_GID-34494_Type-19355?phone;
           iri:has_stars_GID-31330_Type-19355?stars;

}

where {
{?Accommodation a iri:Accommodation_GID-14254}
    UNION
    {?Accommodation a iri:Hotel_GID-19355}

        ?Accommodation  iri:has_type_GID-31834_Type-14254?type;
            filter(?type ="Hotel")

 OPTIONAL{
 ?Accommodation  iri:has_website_GID-34126_Type-19355?website;
           iri:has_type_boarding_GID-300001_Type-19355?boarding;
           iri:has_number_of_rooms_GID-22651_Type-19355?rooms;
           iri:has_City_GID-45969_Type-19355?city;
           iri:has_Email_GID-33745_Type-19355?email;
           iri:has_phone_number_GID-34494_Type-19355?phone;
           iri:has_stars_GID-31330_Type-19355?stars;
            filter(?stars = "3")
            filter(?boarding = "full-board")
                filter(?city = "ANDALO")
}

} limit 5
```

Figure 13: Query 3

	postcode	boarding	rooms	city	phone	service
1	"38010"	"full-board"	"36"	"ANDALO"	"0461-589127"	"Terrace or garden solarium"
2	"38010"	"full-board"	"77"	"ANDALO"	"0461-585802"	"Terrace or garden solarium"
3	"38010"	"full-board"	"59"	"ANDALO"	"0461-585828"	"Terrace or garden solarium"

Figure 14: Result of query 3

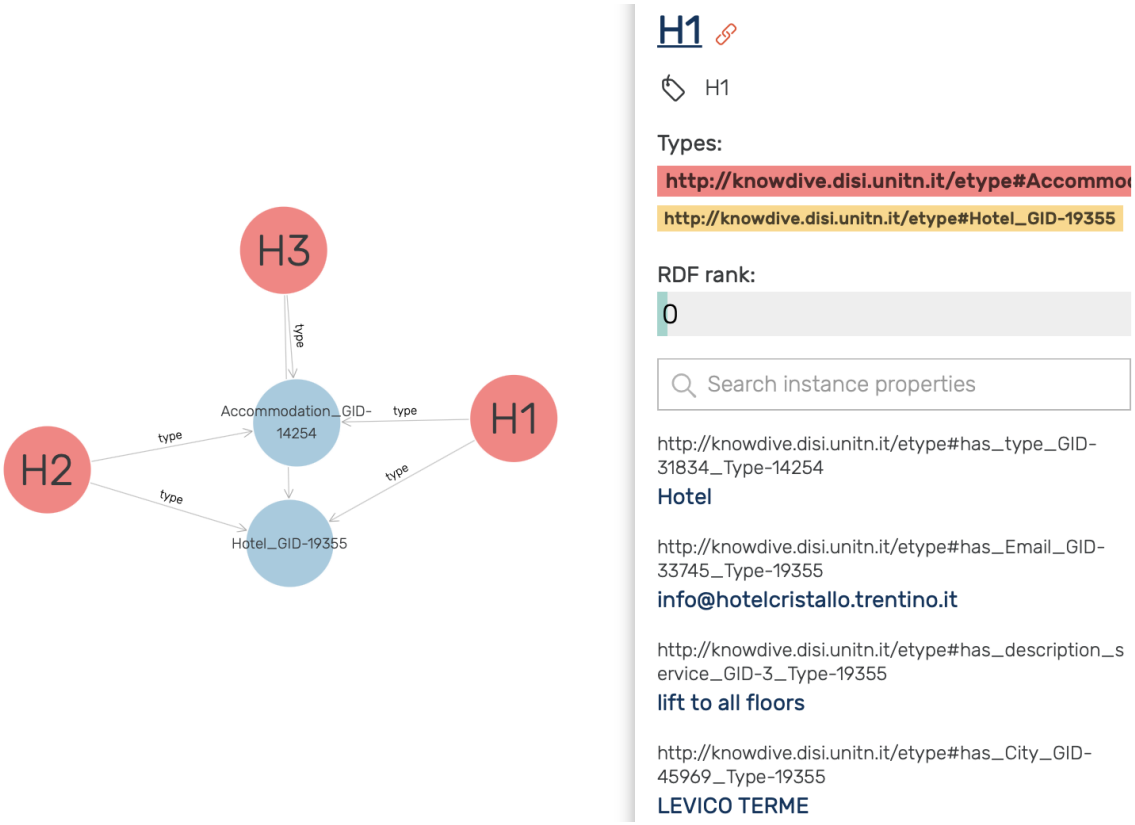


Figure 15: Knowledge graph of query 3

7.3.4 Query 4

Give the list of cafés with outdoor seating in Cles? Return list of cafés with information's.

```
PREFIX iri: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
Construct {
    ?Establishment a iri:Establishment_GID-43815;
                                iri:has_Location_GID-779_Type-779 ?Location;
                iri:has_name_GID-2_Type-43815 ?name;
                iri:has_amenity_GID-2973_Type-43815 ?amenity;
                iri:has_housenumber_GID-34489_Type-43815 ?housenumber;
                iri:has_street_GID-24034_Type-43815 ?street;
                iri:has_postcode_GID-34110_Type-43815 ?postcode;
                iri:has_city_GID-45969_Type-43815 ?city.

    ?Establishment  a iri:Cafe_GID-15804;
                 iri:has_Location_GID-779_Type-779 ?Location;
                 iri:has_OutdoorSeating_GID-46525_Type-15804?OutdoorSeating;
                 iri:has_housenumber_GID-34489_Type-43815 ?housenumber;
                 iri:has_street_GID-24034_Type-43815 ?street;
                 iri:has_postcode_GID-34110_Type-43815 ?postcode;
                 iri:has_city_GID-45969_Type-43815 ?city.

 ?Location  a iri:Location_GID-779;
                iri:has_lattitude_GID-46264_Type-779 ?latitude;
                iri:has_longitude_GID-46270_Type-779 ?longitude.
}


Where
{
    {?Establishment a iri:Establishment_GID-43815}
    UNION
    {?Establishment a iri:Cafe_GID-15804}

   ?Establishment   iri:has_Location_GID-779_Type-779 ?Location;
                iri:has_name_GID-2_Type-43815 ?name;
                iri:has_amenity_GID-2973_Type-43815 ?amenity;
                iri:has_housenumber_GID-34489_Type-43815 ?housenumber;
                iri:has_street_GID-24034_Type-43815 ?street;
                iri:has_postcode_GID-34110_Type-43815 ?postcode;
                iri:has_city_GID-45969_Type-43815 ?city.
                                filter(?amenity = "cafe")
                                        filter(?city ="Trento")

    OPTIONAL {
        ?Establishment iri:has_OutdoorSeating_GID-46525_Type-15804?OutdoorSeating;
                filter(?OutdoorSeating = "yes")
    }

 ?Location iri:has_lattitude_GID-46264_Type-779 ?latitude;
          iri:has_longitude_GID-46270_Type-779 ?longitude.

} ORDER BY  DESC(?OutdoorSeating)
limit 3
```

Figure 16: Query 4

	name	housenumber	street	postcode	city	latitude	longitude	OutdoorSeating
1	"Caffè Grazia"	"44"	"Via Alcide De-gasperi"	"38023"	"Cles"	"46.3628192"	"11.0328503"	
2	"Lanterna Magica"	"20"	"Via Romana"	"38023"	"Cles"	"46.3634483"	"11.0366814"	"yes"
3	"Bar Cles"	"2"	"Piazza Navarrino"	"38023"	"Cles"	"46.3648747"	"11.035348"	

Figure 17: Result of query 4

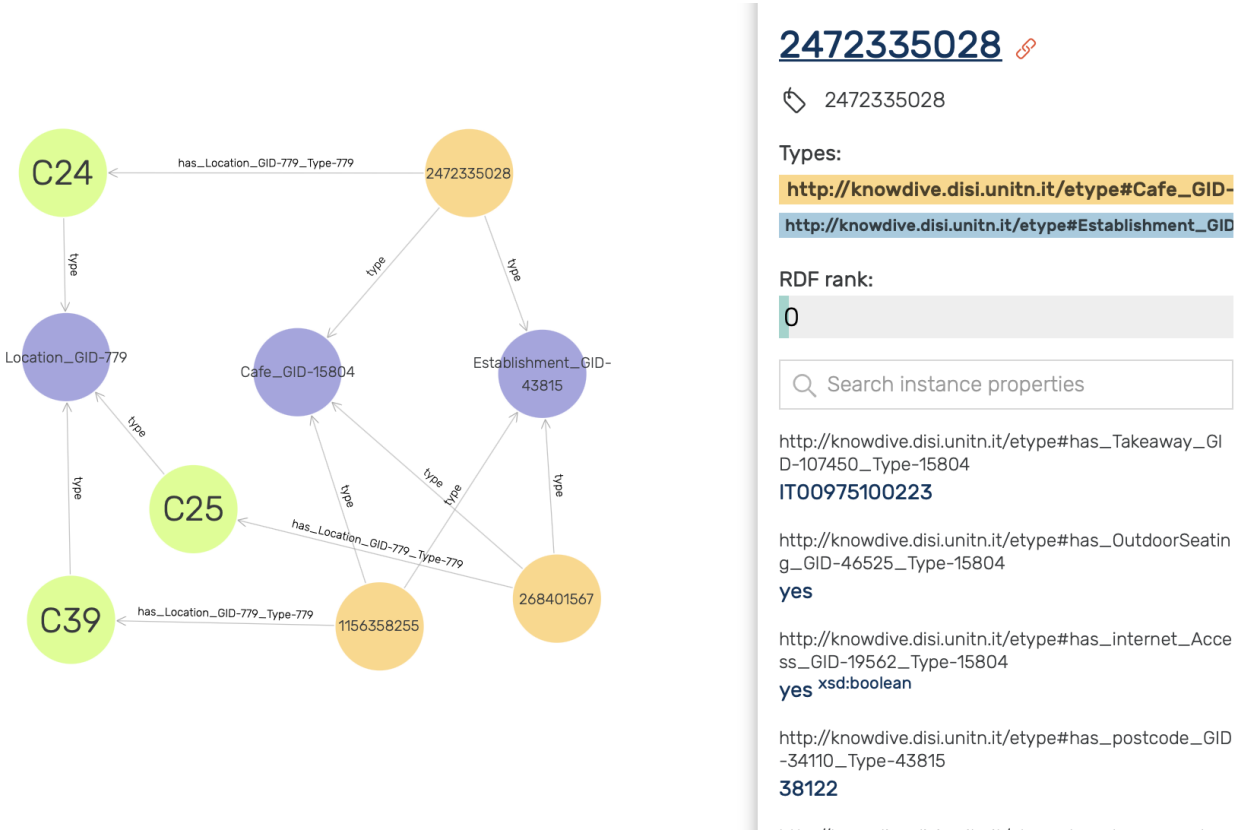


Figure 18: Knowledge graph of query 4

7.3.5 Query 5

Give the list of fast-food restaurants with delivery in Trento? Return list of fast food restaurants with data properties in Trento.

```
PREFIX iri: <http://knowdive.disi.unitn.it/etype#>

CONSTRUCT {
  ?Establishment a iri:Establishment_GID-43815;
                                iri:has_Location_GID-779_Type-779 ?Location;
                iri:has_name_GID-2_Type-43815 ?name;
                iri:has_amenity_GID-2973_Type-43815 ?amenity;
                iri:has_housenumber_GID-34489_Type-43815 ?housenumber;
                iri:has_street_GID-24034_Type-43815 ?street;
                iri:has_postcode_GID-34110_Type-43815 ?postcode;
                iri:has_city_GID-45969_Type-43815 ?city.

   ?Establishment a iri:Fast_food_restaurant_GID-40661;
                  iri:has_Location_GID-779_Type-779 ?Location;
                iri:has_name_GID-2_Type-43815 ?name;
                iri:has_amenity_GID-2973_Type-43815 ?amenity;
                iri:has_housenumber_GID-34489_Type-43815 ?housenumber;
                iri:has_street_GID-24034_Type-43815 ?street;
                iri:has_postcode_GID-34110_Type-43815 ?postcode;
                iri:has_Servescuisine_GID-40725_Type-40661?Servescuisine;
                iri:has_Delivery_GID-1622_Type-40661?delivery;
                iri:has_opening_hours_GID-80563_Type-43815?opening_hours;
                iri:has_city_GID-45969_Type-43815 ?city.

 ?Location  a iri:Location_GID-779;
                iri:has_lattitude_GID-46264_Type-779 ?latitude;
                iri:has_longitude_GID-46270_Type-779 ?longitude.            }

where {
    {?Establishment a iri:Establishment_GID-43815}
    UNION
    {?Establishment a iri:Fast_food_restaurant_GID-40661}

 ?Establishment iri:has_Location_GID-779_Type-779 ?Location;
                iri:has_name_GID-2_Type-43815 ?name;
                iri:has_amenity_GID-2973_Type-43815 ?amenity;
                iri:has_housenumber_GID-34489_Type-43815 ?housenumber;
                iri:has_street_GID-24034_Type-43815 ?street;
                iri:has_postcode_GID-34110_Type-43815 ?postcode;
                iri:has_city_GID-45969_Type-43815 ?city.
                                filter(?amenity = "fast_food")
                                        filter(?city ="Trento")
OPTIONAL {
        ?Establishment iri:has_Servescuisine_GID-40725_Type-40661?Servescuisine;
                  iri:has_Delivery_GID-1622_Type-40661?delivery;
                  iri:has_opening_hours_GID-80563_Type-43815?opening_hours;
                           filter(?delivery = "yes")                                    }


 ?Location iri:has_lattitude_GID-46264_Type-779 ?latitude;
          iri:has_longitude_GID-46270_Type-779 ?longitude.
}
LIMIT 1
```

Figure 19: Query 5

	name	street	postcode	city	Servescuisine	delivery	opening_hours	latitude	longitude
1	"Pedro Pizza"	"Via delle Ghiaie "	"38122"	"Trento"	"pizza"	"yes"	"Mo-Fr 11:00-14:00,18:00-22:00 ; Sa-Su 18:00-22:00"	"46.0556134"	"11.1160854"

Figure 20: Result of query 5

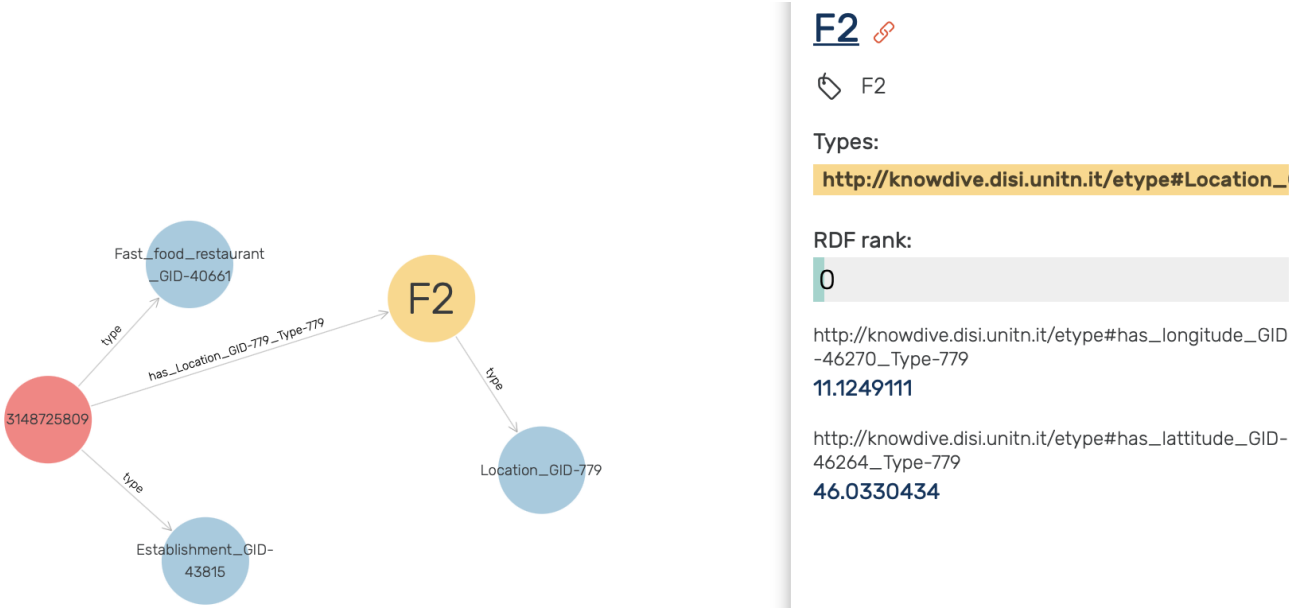


Figure 21: Knowledge graph of query 5

7.4 Conclusion

In this project, through the materials taught in the course, we learned knowledge and data integration practices and applied them in a scenario. We tried to find as much data as possible regarding food and accommodation in the Trentino region and integrated them together. As a result, we've built a base for a possible application to find food and accommodation which can be useful for people who visit Trentino.

During the project, we had certain problems regarding Karmalinker but fortunately we could find a third party Karmalinker application online to fix our problems even after the link shared in the course was offline.

References

[1] Simone Bocca(2021).Purpose Formalization Resource Classification https://unitn-kdi-2021.github.io/unitn-kdi-2021-website/material/slides/W4.L7.M3.T8.1-2.PurposeFormClassification.pdf

[2] Sean Bechhofer, Frank van Harmelenm, Jim Hendler, Ian Horrocks, Deborah L. McGuinness, Peter F. PatelSchneider, Lynn Andrea Stein. OWL Web Ontology Language Reference | Classes. https://www.w3.org/TR/owl-ref/Class, 2004.

[3] Sean Bechhofer, Frank van Harmelenm, Jim Hendler, Ian Horrocks, Deborah L. McGuinness, Peter F. PatelSchneider, Lynn Andrea Stein. SPARQL 1.1 Query Language. https://www.w3.org/TR/sparql11-query/

[4] Dan Brickley, Stéphane Corlosquet, Jason Douglas, R.V.Guha, Sam Goto, Martin Hepp. About Schema.org https://schema.org/docs/about.html

[5] Autonomous Province of Trento with the involvement of Trentino Digitale, the Consortium of Municipalities, the Bruno Kessler Foundation and the University of Trento.The Trentino Open Data project. https://www.trentinodigitale.it/Infrastrutture/Piattaforme-abilitanti/Open-Data