

KNOWLEDGE AND DATA INTEGRATION: FOOD AND ACCOMMODATION

Byambakhishig Ganbold

Mithat Sinan Ergen



OUTLINE

- Project purpose
- Scenario and Personas
- Sources
- Inception phase
- Informal Model
- Formal Model
- Evaluation
- Data Integration
- Queries



PROJECT PURPOSE

- Provide users with information regarding food and accommodation in the Trentino region
- Users will be able to plan their vacations according to this information
- The accommodation information includes 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 includes restaurants, bars, pizzerias, ice-cream shops and their detailed information such as phone number, email and website.

SCENARIO AND PERSONAS

- 8 personas of different ages and background
- Various scenarios such as:
 - sport festival
 - summer vacation
 - winter vacation
 - visit in the weekend.

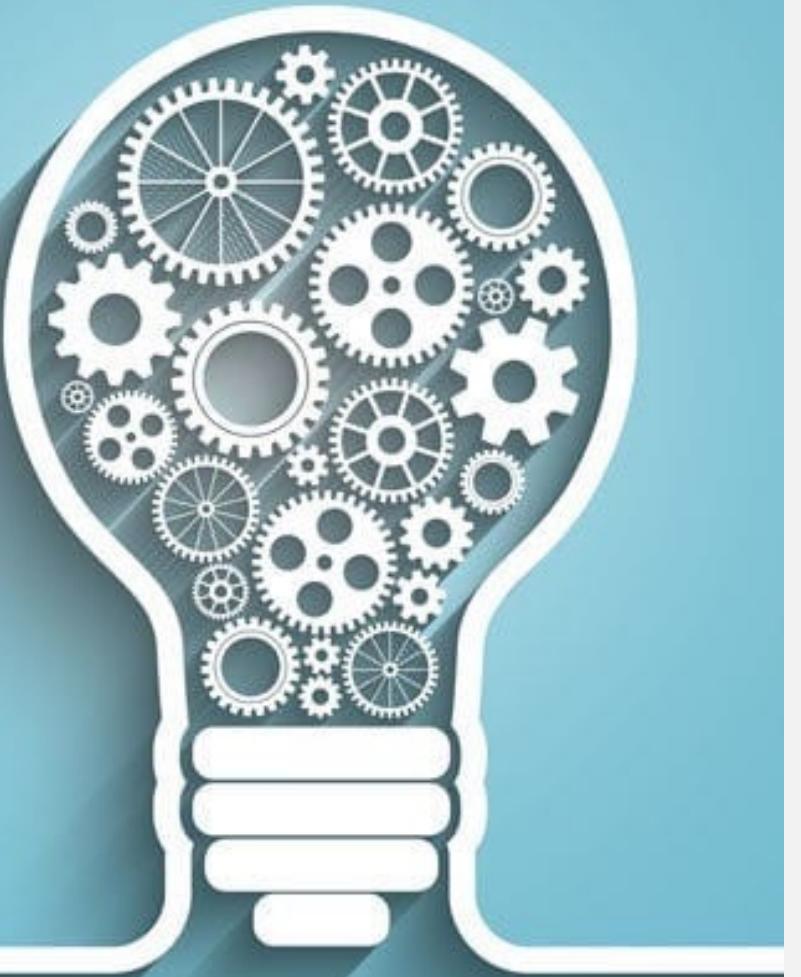




SOURCES

- **Data Sources:**
 - Open Data Trentino: open data source providing data for the Trentino region
 - Inside Airbnb: independent, non-commercial set of tools to explore Airbnb data
 - Open Street Map: a free collaborative project providing data of the world map
- **Knowledge Source:**
 - Schema.org



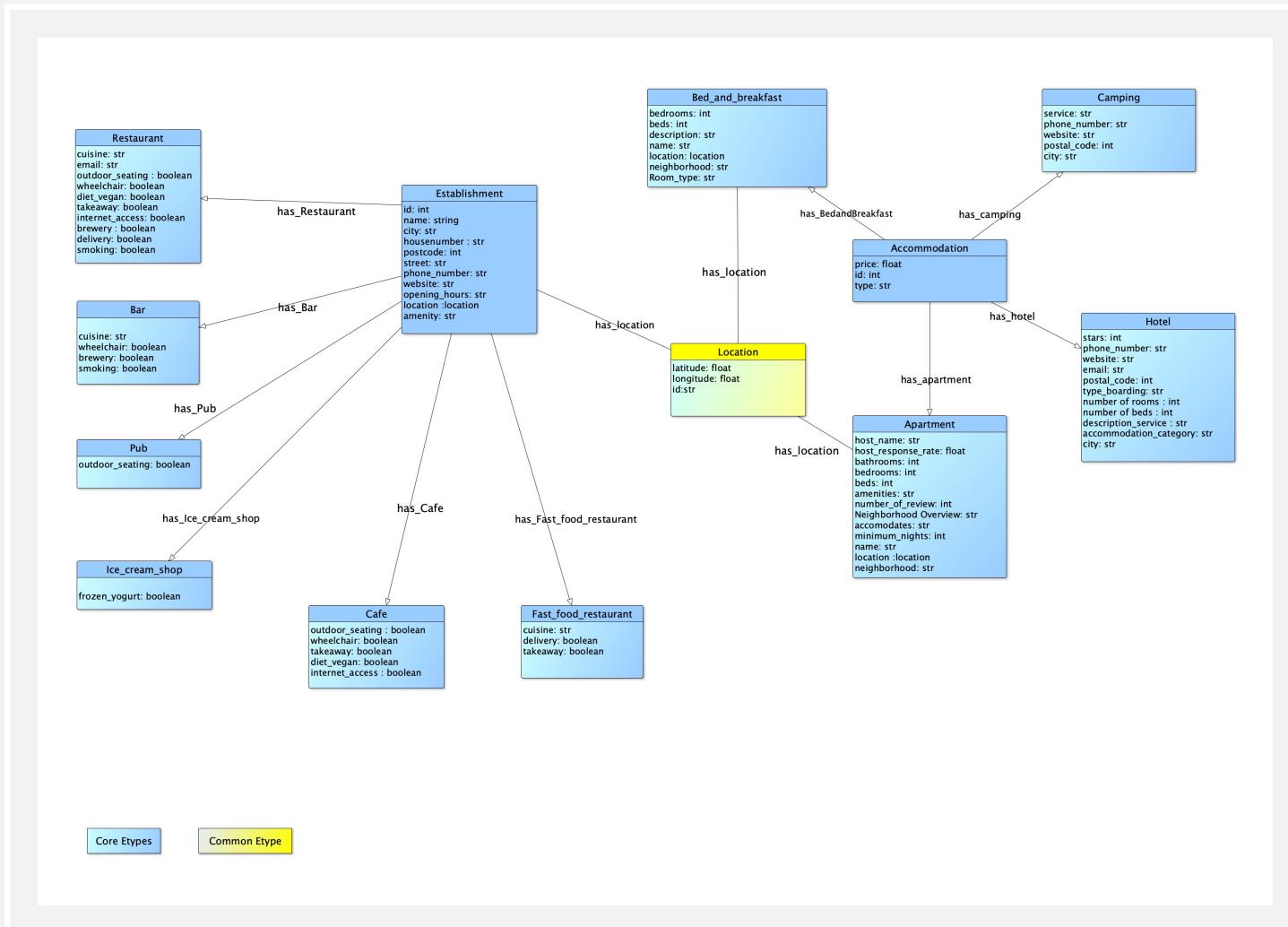


INCEPTION PHASE

- Around 7-8 competency questions(CQ) per persona
- Common, core, contextual kernel concepts are decided
- **Location** as common kernel concept
- **Hotel, Restaurant, Apartment, Bar** as examples of core kernel concepts
- **Amenities, outdoor seating, internet access** as contextual kernel concepts

INFORMAL MODEL

- Common, core, contextual kernel concepts defined at e-type level
- ER Model is formed using yED tool
- Establishment and Accommodation as main e-types
- Location defined as common e-type





FORMAL MODELING

- ETG model

Classes

Active ontology x Entities x Classes x Individuals by class x OWLViz x

Annotation properties Datatypes Individuals

Classes Object properties Data properties

Class hierarchy: xsd:long (xsd:long)

Asserted

owl:Thing

ComplexType

Entity_GID-1

Accommodation_GID-14254

Apartment_GID-14548

Bed_and_breakfast_GID-22651

Camping_GID-5440

Hotel_GID-19355

Action_kind_GID-120018

Context_GID-120016

Establishment_GID-43815

Bar_GID-14950

Cafe_GID-15804

Fast_food_restaurant_GID-40661

Ice_cream_shop_GID-41019

Pub_GID-22135

Restaurant_GID-22500

Function_kind_GID-120019

Location_GID-779

Real_kind_GID-120017

Itelos_complextype_GID-200002

Itelos_concept_GID-200014

Itelos_identifyingset_GID-200012

NLString

SemanticString

IdentifyingSet

xsd:long (xsd:long)

Object properties

Active ontology x Entities x Classes x Individuals by class x OWLViz x

Annotation properties Datatypes Individuals

Classes Object properties Data properties

Object property hierarchy:

Asserted

owl:topObjectProperty

GID

hasIdentifyingSet

hasIdProperty

relationalAttribute

has_apartment_GID-14254_Type-14254

has_bar_GID-14950_Type-43815

has_bed_and_breakfast_GID-22651_Type-14254

has_beginning_GID-81324_Type-1

has_beginning_GID-81324_Type-200008

has_birthday_GID-81257_Type-118

has_birthplace_GID-45896_Type-118

has_birthplace_GID-45897_Type-45663

has_cafe_GID-15804_Type-43815

has_camping_GID-5440_Type-14254

has_contributor_GID-53267_Type-16921

has_creator_GID-51275_Type-16921

has_creator_GID-51275_Type-34892

has_date_of_death_GID-118224_Type-118

has_date_of_publication_GID-117691_Type-34892

has_description_GID-3_Type-1

has_description_GID-3_Type-200008

has_duration_GID-80582_Type-1

has_duration_GID-80582_Type-200008

has_end_GID-81330_Type-1

has_end_GID-81330_Type-200008

has_fast_food_GID-118172_Type-43815

has_first_name_GID-34006_Type-117693

has_has_itelos_actionfunction_GID-200020_Type-200003

has_has_itelos_actionobject_GID-200021_Type-200003

has_has_itelos_functionaction_GID-200022_Type-200004

has_has_itelos_functionobject_GID-200023_Type-200004

Data properties

Active ontology x Entities x Classes x Individuals by class x OWLViz x

Annotation properties Datatypes Individuals

Classes Object properties Data properties

Data property hierarchy:

Asserted

owl:topDataProperty

datatypeAttribute

has_accommodation_category_GID-31828_Type-19355

has_accommodates_GID-300004_Type-14548

has_amenities_GID-26466_Type-14548

has_amenity_GID-2973_Type-43815

has_bathrooms_GID-24688_Type-14548

has_bed_GID-15089_Type-22651

has_bed_type_GID-15089_Type-14548

has_bedroom_GID-15104_Type-22651

has_bedrooms_GID-15104_Type-14548

has_beds_GID-15089_Type-14548

has_brewery_GID-15566_Type-14950

has_brewery_GID-15566_Type-22500

has_city_camping_GID-44481_Type-5440

has_City_GID-45969_Type-19355

has_city_GID-45969_Type-43815

has_cuisine_GID-40654_Type-22500

has_delivery_GID-1622_Type-22500

has_Delivery_GID-1622_Type-40661

has_description_GID-3_Type-22651

has_description_service_GID-3_Type-19355

has_diet_Vegan_GID-300002_Type-15804

has_diet_vegan_GID-300002_Type-22500

has_email_GID-33745_Type-19355

has_email_GID-33745_Type-22500

has_frozen_yogurt_GID-41030_Type-41019

has_host_name_GID-2_Type-14548

has_host_response_rate_GID-300000_Type-14548

has_housenumber_GID-34489_Type-43815

has_id_apartment_GID-34489_Type-14548

has_id_bar_GID-34489_Type-14950

has_id_bed_and_breakfast_GID-34489_Type-22651

has_id_cafe_GID-34489_Type-15804

has_id_campina_GID-34489_Type-5440

DATA MANAGEMENT

- Clear data
- Translate Italian to English
- Convert json file

 Accommodation.json
 apartment.json
 bar.json
 bed_and_breakfast.json
 cafe.json
 camping.json
 Establishment.json
 fast_food.json
 hotel.json
 Ice_cream_shop.json
 location.json
 pub.json
 restaurant.json

EVALUATION

- Schema level

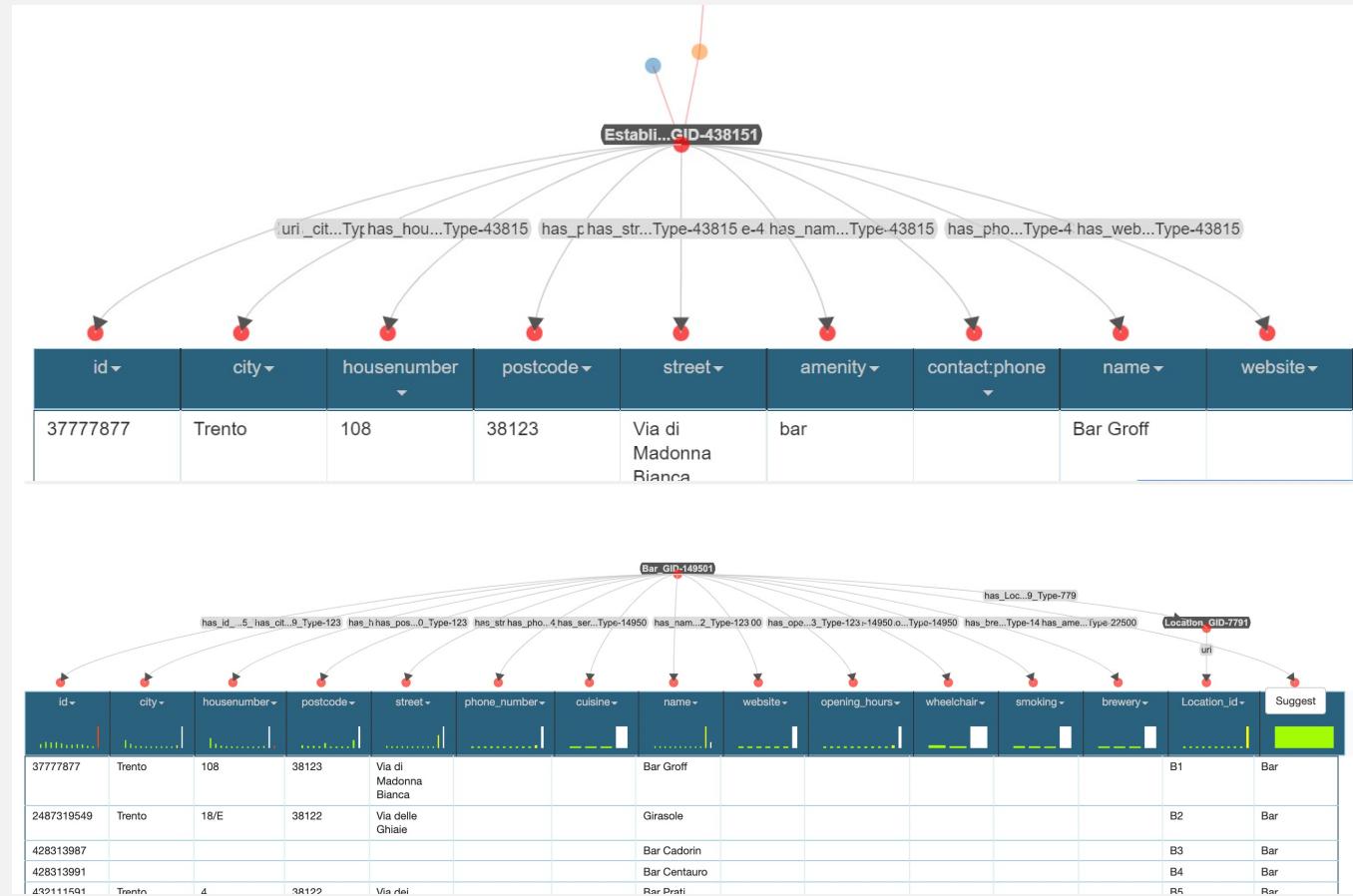
	Inception	Informal Modeling	Formal Modeling
E-type coverage	0.9	1	0.92
Property coverage	0.62	0.98	0.57
E-type extensiveness	0.5	0.16	0.45
Property extensiveness	0.74	0.34	0.62
E-type sparsity			0.5
Property sparsity			0.85

- Data level

	Inception	Informal Modeling
E-type coverage	0.9	0.3
Property coverage	0.98	0.3
E-type sparsity	0.64	0.67
Property sparsity	0.76	0.65

DATA INTEGRATION ~ KARMA LINKER

- Karmalinker allows us to well match the Datasets to the defined Ontology.
- The final step to create the Knowledge Graph of Food and Accommodation





SPARQL QUERY I

```
PREFIX iri: <http://knowdive.disi.unitn.it/etype#>
PREFIX omgeo: <http://www.ontotext.com/owlim/geo#>
select ?name ?amenity ?housenumber ?street ?postcode ?city ?latitude ?longitude

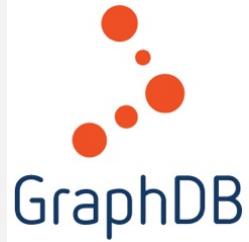
{
  { ?Establishment a iri:Establishment_GID-43815}
    UNION
    {?Establishment a iri:Bar_GID-14950}
      ?Establishment iri:has_Location_GID-779_Type-779 ?Location;
        iri:has_name_GID-2_Type-43815 ?name;
        iri:has_amenity_GID-2973_Type-43815 ?amenity;
        iri:has_housenumber_GID-34489_Type-43815 ?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")

?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)
  # Maximum allowed distance in Km
  FILTER (?distance < 5)

} limit 3
```

- Give the list of bars near the city center in Trento?
- Return list of bars with information's within 5 km from city center.

	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 Giusti"	"38122"	"Trento"	"46.0589545"	"11.1190969"



SPARQL QUERY 2

```
PREFIX iri: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
select ?Name ?price ?beds ?host_response ?Neighborhood ?latitude ?longitude
{
    {?Apartment a iri:Apartment_GID-14548}
    UNION
    {?Apartment a iri:Accommodation_GID-14254}
    ?Apartment 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_host_response_rate_GID-300000_Type-14548?host_response;
        iri:has_Neighborhood_GID-46483_Type-14548?Neighborhood.
        FILTER(xsd:integer(?price) < 80)
        filter contains(?Neighborhood,"Trentino")

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

} limit 3
```

- Give the list of apartment under the 80\$ in Trentino?
- Return list of apartments with data properties in Trentino.

	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/Südtirol, Italy"	"46.36243"	"10.67768"
3	"BIL04 A MARILL-EVA900 SCI E NON S OLO"	"69"	"4"	"100%"	"Marilleva 900, Trentino-Alto Adige/Südtirol, Italy"	"46.31381"	"10.80984"



SPARQL QUERY 3

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

select ?postcode ?boarding ?rooms ?city ?phone ?service
where {
    ?Hotel a iri:Hotel_GID-19355;
    iri:has_postal_code_GID-34110_Type-19355?postcode;
    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_description_service_GID-3_Type-19355?service.
    filter(?boarding = "full-board")
    filter contains(?service, "Terrace")
    filter(?city = "ANDALO")
}
```

LIMIT 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"

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



SPARQL QUERY 4

```
PREFIX iri: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX omgeo: <http://www.ontotext.com/owlim/geo#>
select ?name ?houesnumber ?street ?postcode ?city ?latitude ?longitude ?OutdoorSeating
{
?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_houesnumber_GID-34489_Type-43815 ?houesnumber;
    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 = "Cles")

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.
}
limit 3
```

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

	name	houesnumber	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"	



SPARQL QUERY 5

```
PREFIX iri: <http://knowdive.disi.unitn.it/etype#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX omgeo: <http://www.ontotext.com/owlim/geo#>
select ?name ?street ?postcode ?city ?Servescuisine ?delivery ?opening_hours ?latitude ?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.
}
```

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

	name	street	postcode	city	Servescuisine	delivery	opening_hours	latitude	longitude
1	"Bici Grill Trento"	"Via Lidorno"	"38123"	"Trento"				"46.0330434"	"11.1249111"
2	"Pizzeria Bella Napoli"	"Largo Eugenio Prati"	"38122"	"Trento"				"46.0573733"	"11.1181097"
3	"Pedro Pizza"	"Via delle Ghiae"	"38122"	"Trento"	"pizza"	"yes"	"Mo-Fr 11:00-14:00;18:00-22:00; Sa-Su 18:00-22:00"	"46.0556134"	"11.1160854"
4	"Cavalleggeri"	"Viale Verona"	"38123"	"Trento"	"kebab;pizza"	"yes"	"Mo-Su 10:00-22:00"	"46.0586372"	"11.1275148"

**THANK YOU FOR
LISTENING**

MODELING SHEET

Competency Questions	Common etype			Core Kernel Concepts		
	Object	Function	Action	Object	Function	Action
Give the list of hotels in city center		location		Accommodation	Hotel	
Give the list of hotels with lower price		location		Accommodation	Hotel	
Give the email of vegan restaurants in Trento		location		Food	Restaurant	
Give the list vegan restaurant with internet access in Rovereto		location		Food	Restaurant	
Give the phone numbers of restaurants with outdoor seating		location		Food	Restaurant	
Give the list of hotels with phone number information near train station		location		Accommodation	Hotel	
Give the comfortable hotels under 60€ near mountain and good view		location		Accommodation	Hotel	
Give the list of the pub in city center open 9 pm		location		Food	Pub	
Give the list of cafe with internet access and takeaway		location		Food	Cafe	
Give the list of 4 or 5 stars hotels with swimming pool in Trentino		location		Accommodation	Hotel	
Give the list of hotel with terrace?		location		Accommodation	Hotel	
Give the list of Airbnb private room near Garda Lake with car parking		location		Accommodation	Bed and breakfast	
Give the list of restaurant with outdoor seating in Garda Lake		location		Food	Restaurant	
Give the list of restaurants with brewery and outdoor seating		location		Food	Restaurant	
Give the address of bars in Trento		location		Food	bar	
Give the list of phone number of pizzas in Trento		location		Food	fast_food_restaurant	
Give the list of ice cream shops with frozen_yogurt and outdoor_seating		location		Food	Ice_cream-shop	
Give the list of Airbnb houses with lift for 4 people in Rovereto		location		Accommodation	Apartment	
Give the list of Airbnb houses with middle level price		location		Accommodation	Apartment	
Give the list of Airbnb apartment with kitchen and TV		location		Accommodation	Apartment	
Give the list of Airbnb houses with heating near ski resorts		location		Accommodation	Apartment	
Give the list of restaurant with delivery near Andalo		location		Food	Restaurant	
Give the list of restaurant with wheelchair		location		Food	Restaurant	
Give the cafe with wheelchair and vegan food		location		Food	cafe	
Give the list of pubs with outdoor seating		location		Food	pub	

ATTRIBUTED CQ_S

EType	Object Properties	Data Properties
Hotel	has_hotel	
Hotel	has_hotel	has_price
Restaurant	has_restaurant	has_diet_vegan, has_email
Restaurant	has_restaurant	has_diet_vegan, has_internet_access, has_city
Restaurant	has_restaurant	has_outdoor_seating, has_phone_number
Hotel	has_hotel	has_phone_number
Hotel	has_hotel	has_price
Pub	has_pub	has_open_hours
Cafe	has_cafe	has_internet_access, has_takeaway
Hotel	has_hotel	has_description_service, has_starsRating
Hotel	has_hotel	has_description_service
Bed and breakfast room	has_bed_and_breakfast	has_Description
Restaurant	has_restaurant	has_outdoor_seating, has_brewery, has_city
Restaurant	has_restaurant	has_outdoor_seating, has_brewery
bar	has_bar	has_street_bar, has_house_number
fast_food_restaurant	has_fast_food_restaurant	has_cuisine, has_city, has_phone_number
Ice_cream-shop	has_Ice_cream-shop	has_frozen_yogurt, has_outdoor_seating
Apartment	has_apartment	has_Amenities
Apartment	has_apartment	has_price
Apartment	has_apartment	has_Amenities
Apartment	has_apartment	has_Amenities, has_Neighborhood_Overview
Restaurant	has_restaurant	has_delivery, has_city
Restaurant	has_restaurant	has_wheel_chair
cafe	has_cafe	has_wheel_chair_cafe, has_vegan_cafe
pub	has_pub	has_outdoor_seating_pub
Camping	has_camping	has_service
Hotel	has_hotel	has_type_boarding
Camping	has_camping	has_service
Camping	has_camping	has_phone_number
fast_food_restaurant	has_fast_food_restaurant	has_takeaway
fast_food_restaurant	has_fast_food_restaurant	has_cuisine
fast_food_restaurant	has_fast_food_restaurant	has_delivery, has_phone_number
bar	has_bar	has_smoking_bar
Hotel	has_hotel	
Hotel	has_hotel	

11 classes, 50 data properties
13 object properties