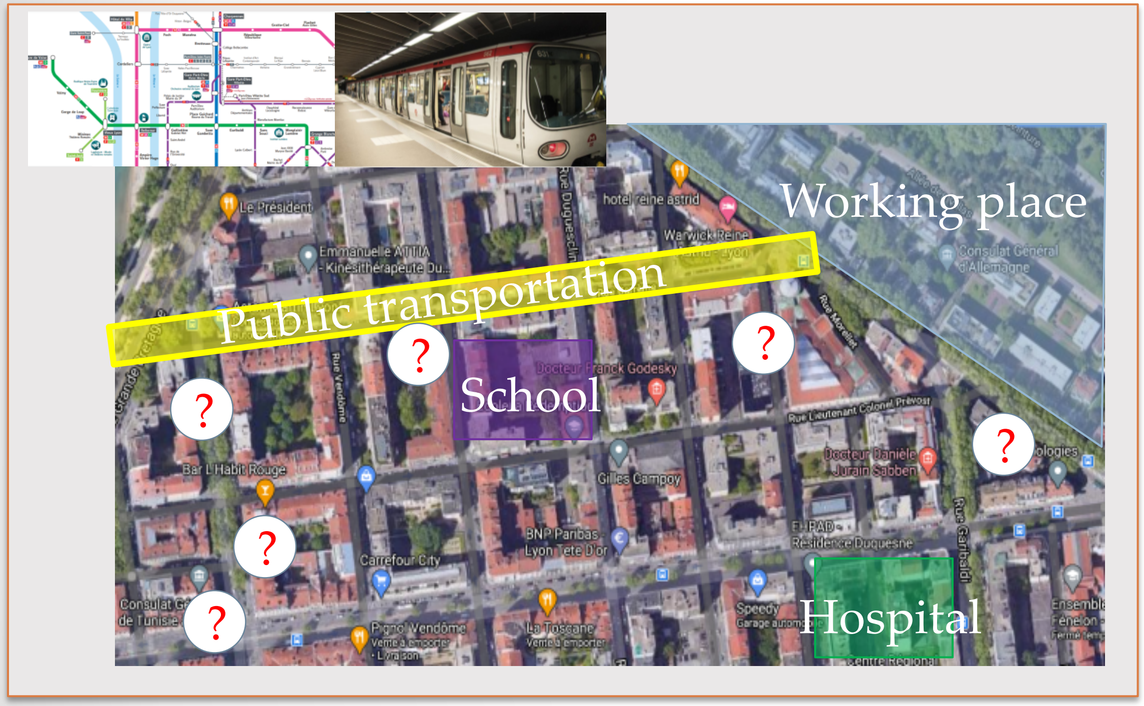
Ontology-Based Approach for Neighborhood and Real Estate Recommendations (NAREO)

Wissame Laddada, Fabien Duchateau, Franck Favetta, Ludovic Moncla

LocalRec'20, November 3, 2020

Context and Motivations



Buying/renting a real estate:

Multiple criteria

Preferences

Macro level: amenities, neighborhood, leisure, shopping, etc.
Micro level: apartment, number of rooms, price, balcony, etc.

Micro level: Real Estate

Macro level: Neighborhood

Within

Context and Motivations

Tiresome task

Recommendation system

Buying/renting a real estate:

Multiple criteria

Preferences

Macro level: amenities, neighborhood, leisure, shopping, etc.
Micro level: apartment, number of rooms, price, balcony, etc.

Micro level: Real Estate

Macro level: Neighborhood

Within

Outline

- Spatial Recommendation Background
- Ontology for Neighborhood and Real estate recommendation
  - Motivations and approach overview
  - Conceptualization/modeling
  - SWRL rules definition for criteria
- Case study
- Perspectives and discussion

Spatial Recommendation Background

User perception is consistent with the quantitative numbers that describe the given neighbor-hood

Algorithms/ Machine learning

Le Falher et al.
Barret et al.

Limited to distance between places

Pattern matching

Li et al.

Ontologies

Malczeweski et al.

SWRL Rules

Heavy task for users when scoring preferences (taking into account by means of rules)

Yuan et al.

SEED Layout

Limited to distances from locations and some real estate constraints

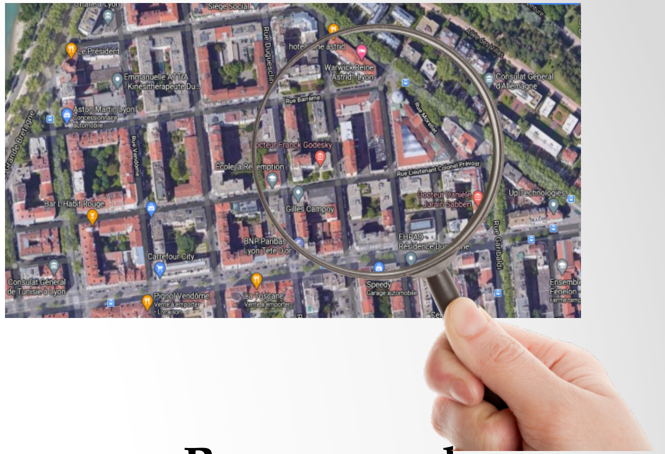
Ontology for Neighborhood and Real estate recommendation NAREO

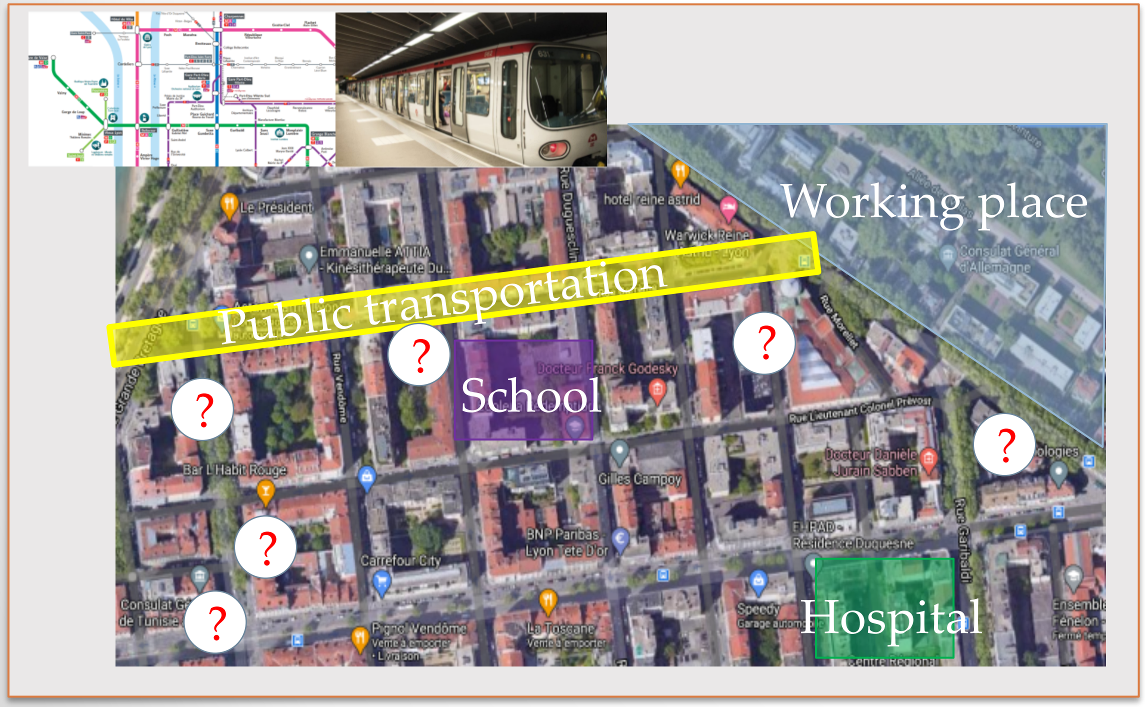
Motivation

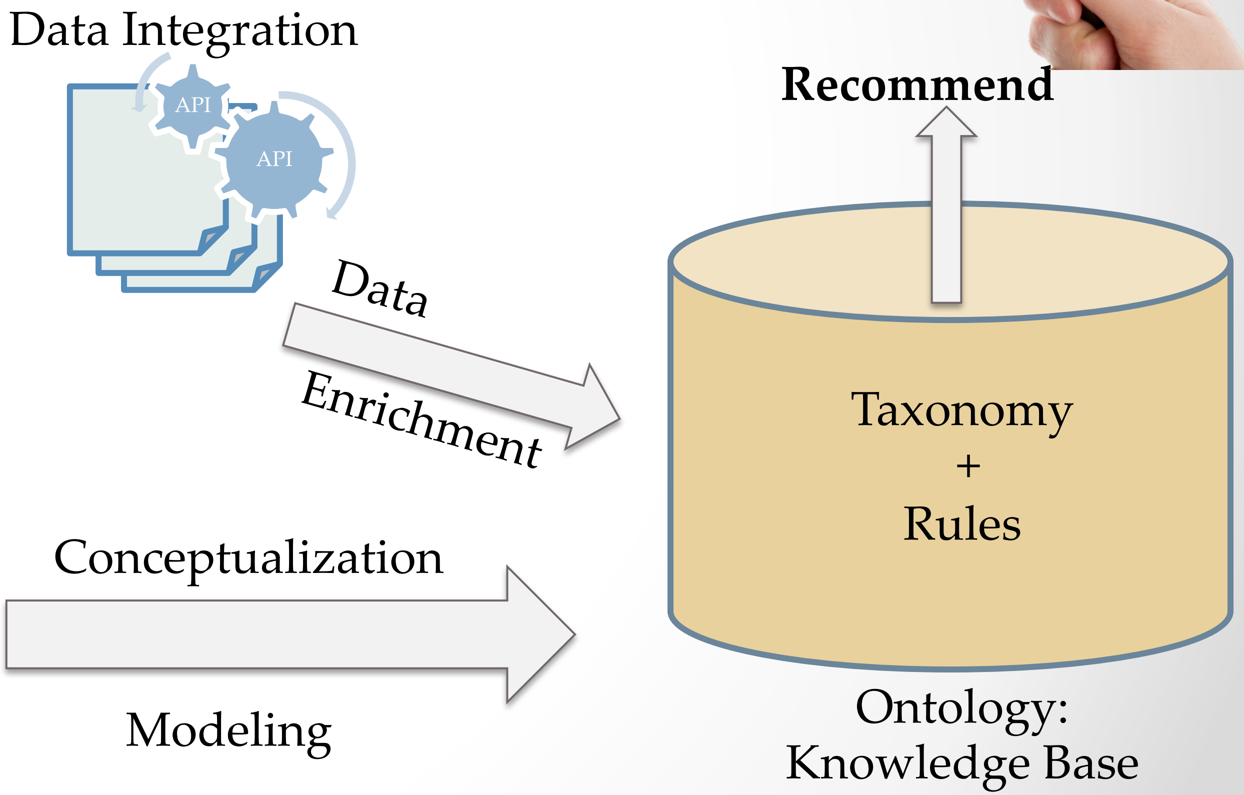
- Modeling qualitative and quantitative preferences
- Extension is possible if needed (adding classes, properties, rules, axioms)
- Integrate spatial information with theories for spatial reasoning
- Relevant semantic expressiveness for queries to get a specific information

Approach Overview

- Conceptualization/modeling: define a taxonomy of concepts/ roles within a knowledge base to model the environment
- Rules definition: define rules for criteria so a user can express preferences
- Data integration: collect data from different resources to enrich the ontology
- Recommendation : suggestions by means of semantic queries and reasoning process

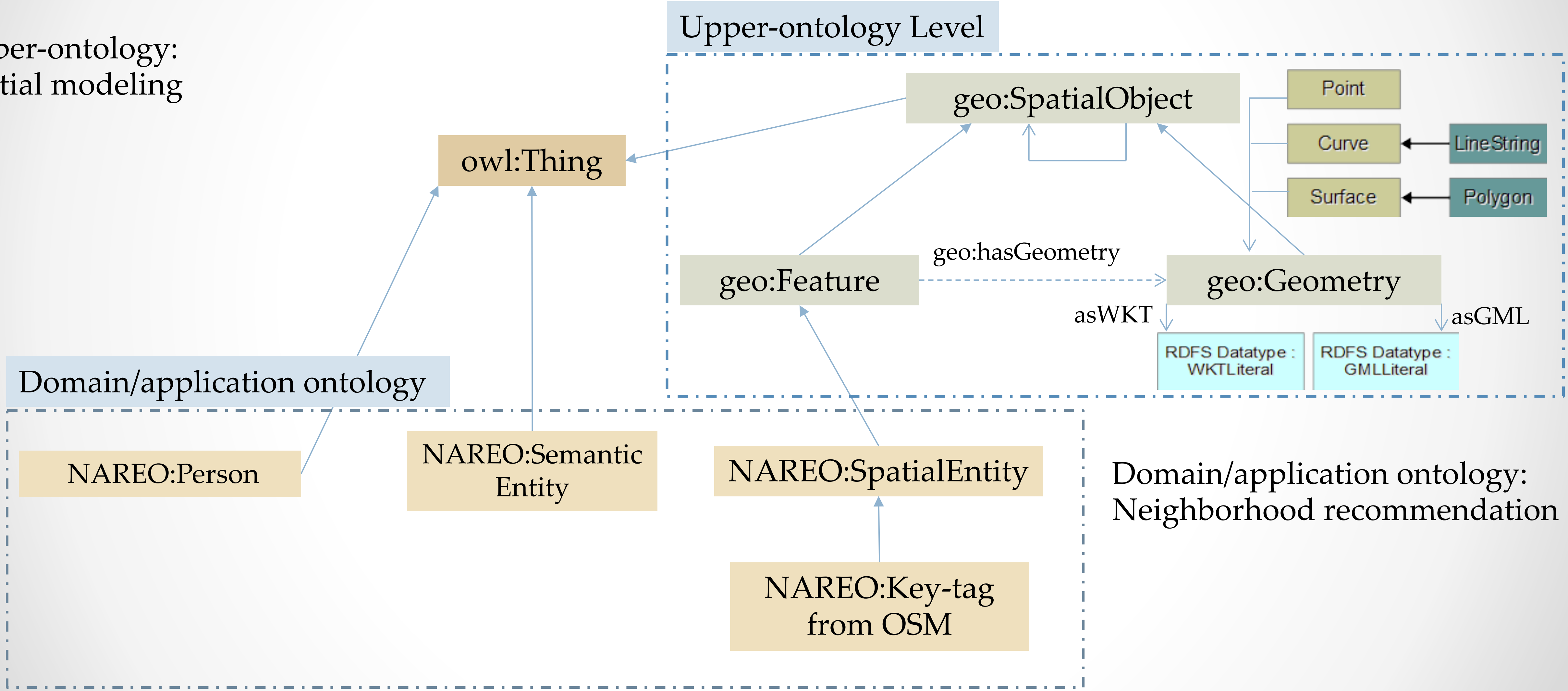






Conceptualization/modeling: Concept Definition (1/3)

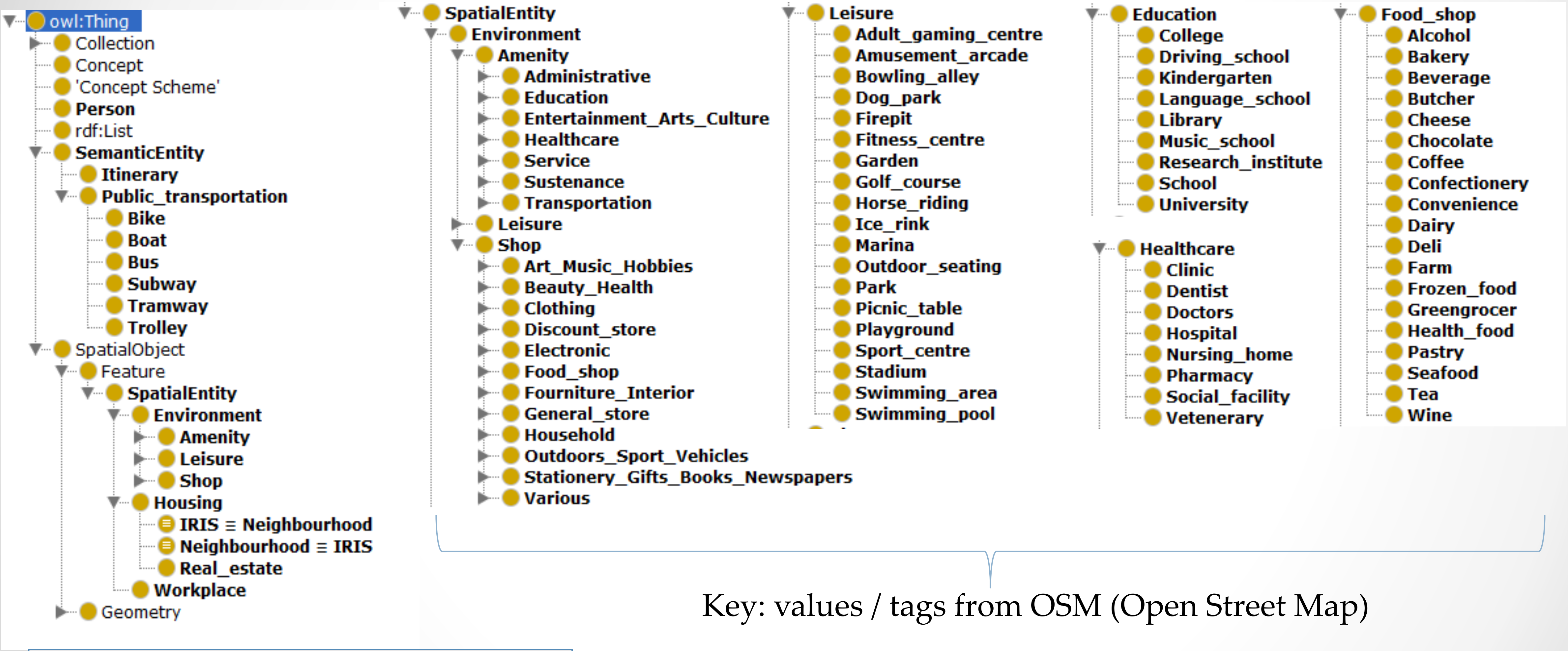
Upper-ontology:
Spatial modeling



Domain/application ontology:
Neighborhood recommendation

geo: http://www.opengis.net/ont/geosparql#

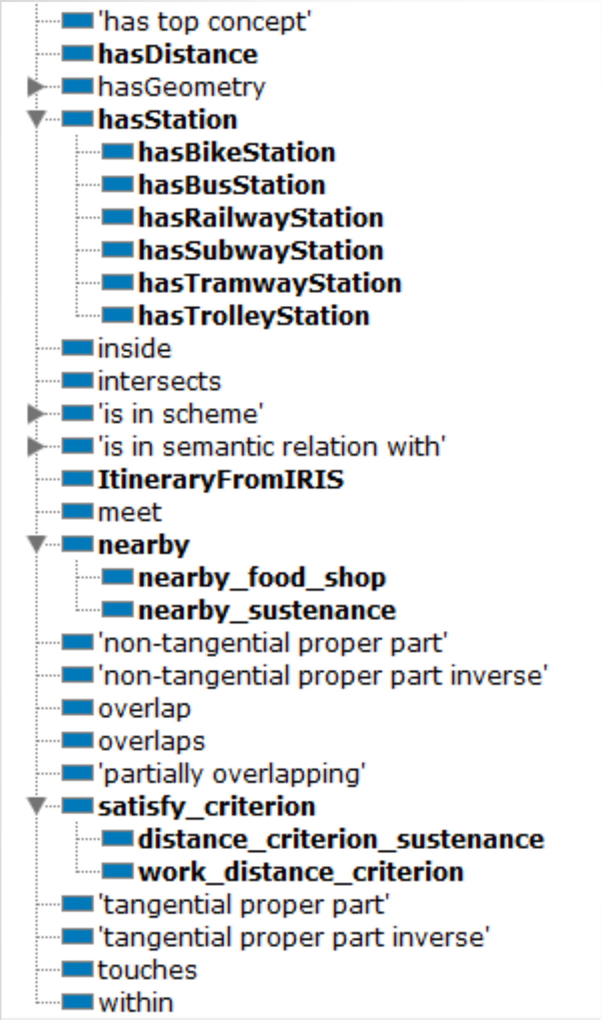
Conceptualization/modeling: Concept Definition (2/3)

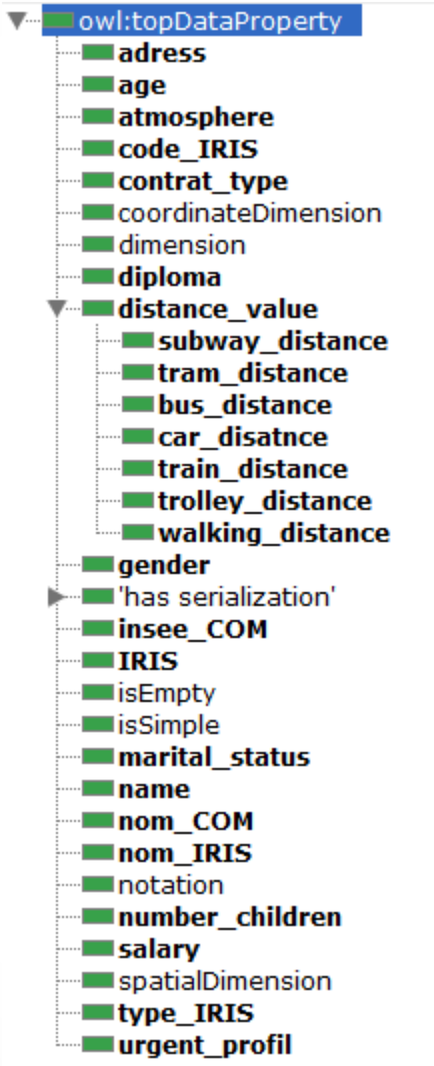


Key: values / tags from OSM (Open Street Map)

NAREO: Partial overview (concepts)

Conceptualization/modeling: Roles Definition (3/3)





Example

hasDistance(SpatialEntity,Itinerary)
hasSubwayStation(Subway,Transport_station)
itineraryFromIRIS(Itineray,IRIS)

IRIS(s)
Intinerary(i)
Workplace(k)
hasDistance(k,i)
itineraryFromIRIS(i,s)
subway_distance(i, ’20’)
tram_distance(i,’15’)

NAREO: Partial overview (Properties)

SWRL rules definition for criteria (1/2)

$Predicate_1 (x_1)$ ^… $Predicate_2 (x_1,y_1)$ … ^$Predicate (x_n)$ => $Consequences(x_i,y_j)$ → DL safe rule

SWRL Rule 1

```
IRIS(?i) ^ hasGeometry(?i,?g1) ^ asWKT(?g1,?w1) ^
Food_shop(?f) ^ hasGeometry(?f,?g2) ^ asWKT(?g2,?w2) ^
distance_criterion_food_shop(?d,?w1,?w2)
=> nearby_food_shop(?i, ?f)
```

→ Proximity to food shops criterion
Built-in: distance_criterion_food_shop

SWRL Rule 2

```
IRIS(?i) ^ hasGeometry(?i,?g1) ^ asWKT(?g1,?w1) ^
Sustenance(?s) ^ hasGeometry(?s,?g2) ^ asWKT(?g2,?w2) ^
distance_criterion_sustenance(?d,?w1,?w2)
=> nearby_sustenance(?i, ?s)
```

SWRL Rule 3

```
IRIS(?i) ^ (nearby_sustenance >= 5)(?i)
=> atmosphere(?i, "true")
```

→ Proximity to sustenance and neighborhood atmosphere criterion
Built-in: distance_criterion_sustenance

SWRL rules definition for criteria (2/2)

SWRL Rule 4

```
IRIS(?i) ^ hasGeometry(?i,?g1) ^ asWKT(?g1,?w1) ^
Workplace(?k) ^ hasGeometry(?k,?g2) ^ asWKT(?g2,?w2) ^
hasDistance(?k, ?iT1) ^...^ hasDistance(?k,iT5) ^
differentFrom(?iT1,?iT2) ^ differentFrom(?iT1,?iT3) ^
...^ differentFrom(?iT4,?iT5) ^
itineraryFromIRIS(?iT1,?i) ^...^
itineraryFromIRIS(?iT5,?i) ^
work_distance_criterion (?w1,?w2,?dv1,...,
?dv5,?v1,?v2,?v3,?v4...)
=> subway_distance(?iT1,?v1 ) ^...^
tram_distance(?iT5,?v35) ^distance_value(?iT1,?dv1)^...^
distance_value(?iT5,?dv5)
```

→

Distances to workplace
Built-in: work_distance_criterion

Inference of different distances (tram_distance, subway_distance, etc.) for each itinerary, from a given neighborhood to a workplace

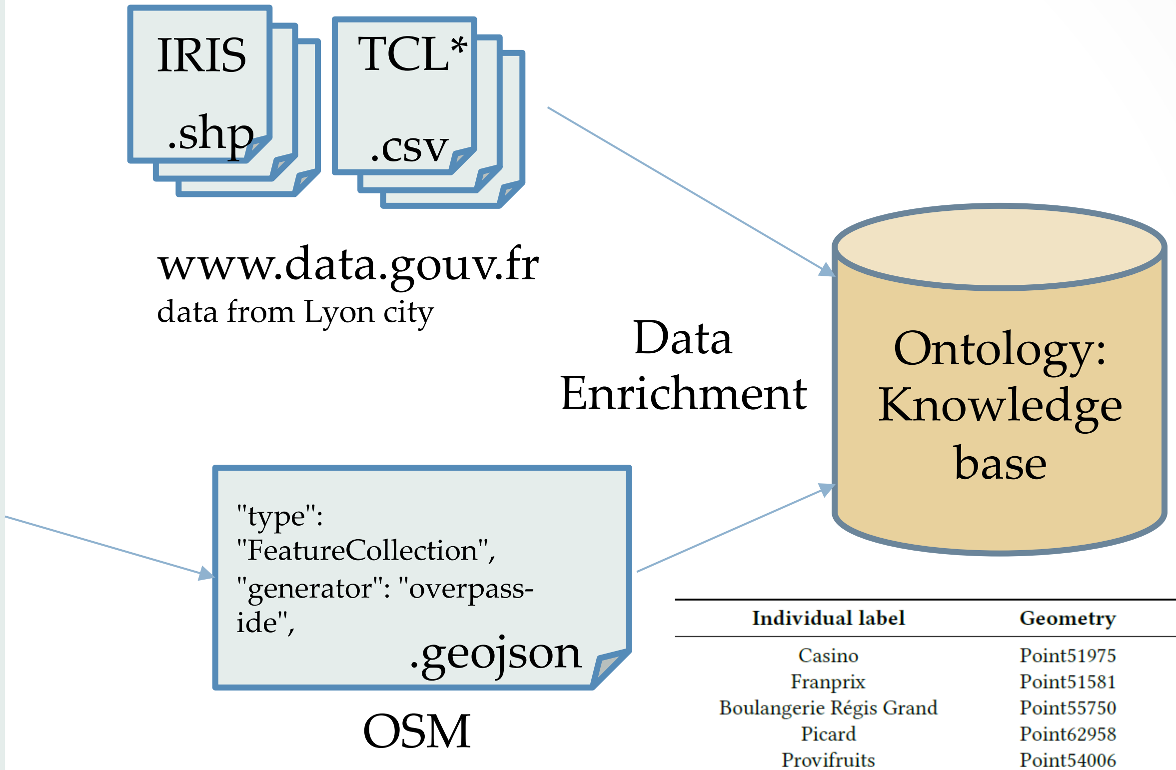
Case study: Data integration (1/3)

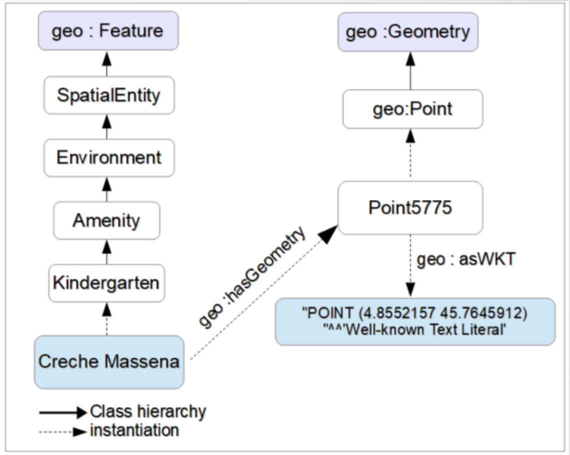
```
[out:json];
// gather results
(
  // query part for: “aminity=*”
  node["amenity"](around:10000,45.75,4.85);
  way["amenity"](around:10000,45.75,4.85);

relation["amenity"](around:10000,45.75,4.85)
;
  nwr["shop"](around:10000,45.75,4.85);
  nwr["leisure"](around:10000,45.75,4.85);
);
// print results

out;
>;
out skel qt;
```

https://overpass-turbo.eu/





Individual label	Geometry	asWKT	Class
Casino	Point51975	POINT (4.8655005 45.7711646)	Supermarket
Franprix	Point51581	POINT (4.867083 45.7737778)	Supermarket
Boulangerie Régis Grand	Point55750	POINT (4.8618856 45.7705183)	Bakery
Picard	Point62958	POINT (4.8618621 45.77022)	Frozen_food
Provifruits	Point54006	POINT (4.8660036 45.7709304)	Greengrocer
Martins Boucher	Point54389	POINT (4.8669555 45.7733211)	Butcher
Leader Price Express	Point56688	POINT (4.8628103 45.7724536)	Convenience
The Brew Brothers	Point50562	POINT (4.856848 45.7697452)	Pub
Le Waldeck Sweet Bar	Point50623	POINT (4.8583287 45.7698096)	Pub
Le Select	Point57491	POINT (4.8667944 45.7731651)	Bar
Le Charpenne	Point57512	POINT (4.8664263 45.7730669)	Bar
Okawali	Point61447	POINT (4.863124 45.769918)	Restaurant
McDonald’s	Point52337	POINT (4.8633017 45.7700648)	Fast_food
Sapori di casa	Point57474	POINT (4.8612948 45.7708478)	Restaurant
Le Béranger	Point57529	POINT (4.8609165 45.7699466)	Restaurant
Le Bistrot du Potager-Stalingrad	Point70582	POINT (4.8586965 45.7706483)	Restaurant
Le Hoggar	Point72993	POINT (4.867677 45.771059)	Restaurant
692660301 (Tonkin-Sud)	MultiPolygon129	MULTIPOLYGON (((4.86594583533633 45.77...	IRIS

*TCL: Transport en Commun de Lyon
(public transportation of Lyon city)

Case study: Inferences by means of SWRL rules (2/3)

Proximity criterion (to food shops) (Rule 1)

IRIS	Inference	Food_shop
692660301 (Tonkin-Sud)	nearby_food_shop	Casino
692660301 (Tonkin-Sud)	nearby_food_shop	Franprix
692660301 (Tonkin-Sud)	nearby_food_shop	Boulangerie Régis Grand
692660301 (Tonkin-Sud)	nearby_food_shop	Picard
692660301 (Tonkin-Sud)	nearby_food_shop	Provifruits
692660301 (Tonkin-Sud)	nearby_food_shop	Martins Boucher
692660301 (Tonkin-Sud)	nearby_food_shop	Leader Price Express

Proximity criterion (to sustenance) (Rule 2)

IRIS	Inference	Sustenance
692660301 (Tonkin-Sud)	nearby_sustenance	The Brew Brothers
692660301 (Tonkin-Sud)	nearby_sustenance	Le Waldeck Sweet Bar
692660301 (Tonkin-Sud)	nearby_sustenance	Le Select
692660301 (Tonkin-Sud)	nearby_sustenance	Le Charpenne
692660301 (Tonkin-Sud)	nearby_sustenance	Okawali
692660301 (Tonkin-Sud)	nearby_sustenance	McDonald's
692660301 (Tonkin-Sud)	nearby_sustenance	Sapori di casa
692660301 (Tonkin-Sud)	nearby_sustenance	Le Béranger
692660301 (Tonkin-Sud)	nearby_sustenance	Le Hoggar

Inference of the triplet about the atmosphere (rule3) :

•"692660301" → atmosphere → "true"

Facts about workplace :

- Workplace(LIRIS)
- hasGeometry(LIRIS, Point9600)
- asWKT(Point9600, POINT(4.865.. 45.78..))
- Itinerary(T1), Itinerary(T2), Itinerary(T3)
- itineraryFromIRIS(T1, 692660301)
- itineraryFromIRIS(T2, 692660301)
- itineraryFromIRIS(T3, 692660301)

Inferences of distances (Rule 4)

- **walking_distance(T1, 11)**, **bus_distance(T1, 2)**
- **distance_value(T1, 13),** train_value(T1, 0)
- subway_distance(T1, 0), car_distance(T1, 0)
- tram_distance(T1, 0), Trolley_distance(T1, 0)
- **walking_distance(T2, 5)**, **tram_distance(T2, 5)**
- **distance_value(T2, 10)**, train_value(T2, 0)
- subway_distance(T2, 0), car_distance(T2, 0)
- bus_distance(T2, 0), Trolley_distance(T2, 0)
- **walking_distance(T3, 19),distance_value(T3, 19)**
- tram_distance(T3, 0), train_value(T3,0)
- subway_distance(T3, 0), car_distance(T3, 0)
- bus_distance(T3, 0), Trolley_distance(T3, 0)

Case study: Semantic queries to express preferences (3/3)

SPARQL Query 1

```
select ?name_IRIS ?itinerary ?temporal_distance
  where{?x a base:IRIS;
        base:nom_IRIS ?name_IRIS.
        ?w a base:Workplace;
        rdfs:label \"Nautibus\";
        base:hasDistance ?itinerary.
        ?itinerary base:ItineraryFromIRIS ?x;
        base:distance_value ?temporal_distance.
        FILTER (?temporal_distance > 0 &&
                ?temporal_distance <= 20)
   }
```

1. Get neighborhoods (IRIS) for an accommodation by taking into account the distance to a workplace

SPARQL Query 2

```
select ?name_IRIS (count(distinct ?shop) as ?count)
  where {?x a base:IRIS;
         base:nom_IRIS ?name_IRIS;
         base:nearby_food_shop ?shop.
         } GROUP BY ?name_IRIS
```

2. Get neighborhoods nearby food shops classified per categories

SPARQL Query 3

```
select ?name_IRIS
   where{?x a base:IRIS;
         base:nom_IRIS ?name_IRIS;
         base:atmosphere true.
         }
```

3. Get neighborhoods that are animated or quiet

SPARQL Query 4

```
select ?name_IRIS ?name_sustenance ?class
  where {?x a base:IRIS;
         base:nom_IRIS ?name_IRIS;
         base:nearby_sustenance ?tag.
         ?tag rdfs:label ?name_sustenance;
         a ?class."
         FILTER( STRSTARTS(STR(?class),str(base:)) )
         }
```

4. Refine recommendation considering an animated neighborhood and add a restriction on sustenance category

Perspectives and discussion

- More semantics expressing criteria, like natural elements in the surrounding area (IRIS), should be added to NAREO to refine neighborhood recommendation
- Extend NAREO with more concepts/roles to recommend a real estate being within the suggested neighborhood
- Handling spatial reasoning by means of geospatial RDF stores (Parliament, Strabon, etc.)

➢ Enhancing NAREO with semantics to express criteria and preferences leads to define other rules which may increase the complexity with more data enrichment

Thank you for your attention.