Barcelona Environmental Monitoring

Group 15

Zaki Amin, Lucia Fabbri, Paul Faschingbauer, Jonah Wenzel

17 November 2023

Application demo

Our datasets: overview

Measurements and the environment in the city of Barcelona:

- **Temperaturesbarcelonadesde1780.csv:** shows temperatures for each month from year 1780 and onwards

- **qualitat_aire_contaminants.csv**: basic information about contaminants including number and name

- **OD_Arbrat_Zona_BCN.csv**: descriptions of trees, species and locations

- **2023_qualitat_aire_estacions.csv**: information about air quality cabins including the districts they are in

- **2022_tramer_pm10_mapa_qualitat_aire_bcn_csv**: connecting streets with air quality measurements

❏ **Licence**: Creative Commons Attribution 4.0 licence, allowing to share and remix the content even for commercial purposes

❏ **Linkage Property**: neighbour, district and contaminant code, allowing to link elements to each other

❏ **Documentation**: not in-depth documentation but some light documentation

❏ **Data Structure**: The data itself is quite well-structured and self-descriptive and they came from multiple data sources.
  ❏ Contaminant codes are used across two CSV files

Data cleaning, resource naming and ontology

- Deletion of irrelevant columns and cleaning, separation of data with OpenRefine

- Following simple resource naming strategy:
  `https://barca_environmental_monitoring.es/resource/<Class>/<ID>`

- Reuse of existing ontologies such as schema.org

- Adding new properties for reconciled data from Wikidata

- Application queries traverse the ontology:

  air contaminant → air quality cabin → address → neighbourhood
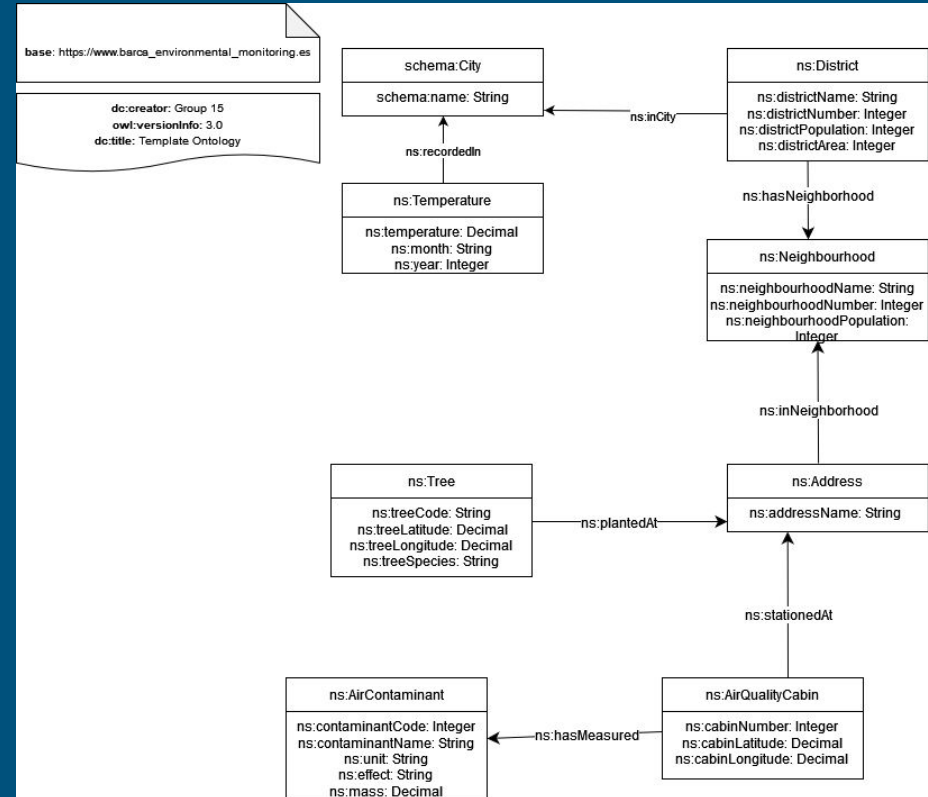
Reconciliation and RDF generation

- Reconciliation with Wikidata
  - Contaminants: effect, mass
  - Populations for neighbourhood and district
- RDF generation: YML mapping and conversion to RML before generating RDF
  - Use of csv, yarrrml-parser, morph-kgc
  - Following the resource naming strategy and ontology
  - Automating the two conversions with a script to make iterations with improved ontologies easier

YML for mapping a district into RDF

```
districts:
  sources:
    - [ ../csv/qualitat_aire_estacions-with-links.csv~csv ]
  s: ns:resource/District/$(District_Number)
  po:
    - [ a, ns:District ]
    - [ ns:districtNumber, $(District_Number), xsd:integer ]
    - [ ns:districtName, $(District_Name) ]
    - [ ns:hasNeighbourhood, ns:resource/Neighbourhood/$(Neighbourhood_Number)~iri ]
    - [ ns:inCity, ns:resource/City/Barcelona~iri ]
    - [ ns:districtPopulation, $(District_Population), xsd:integer ]
    - [ owl:sameAs, $(Wikidata_District)~iri ]
```

Outcome: Generated RDF for a district

```
<https://barca_environmental_monitoring.es/resource/District/10> a ns1:District ;
    owl:sameAs <https://wikidata.org/wiki/Q250935> ;
    ns1:districtName "Sant Martí" ;
    ns1:districtNumber 10 ;
    ns1:districtPopulation 241181 ;
    ns1:hasNeighbourhood <https://barca_environmental_monitoring.es/resource/Neighbourhood/68> ;
    ns1:inCity <https://barca_environmental_monitoring.es/resource/City/Barcelona> .
```

Thank you!

Group 15

Zaki Amin, Lucia Fabbri, Paul Faschingbauer, Jonah Wenzel