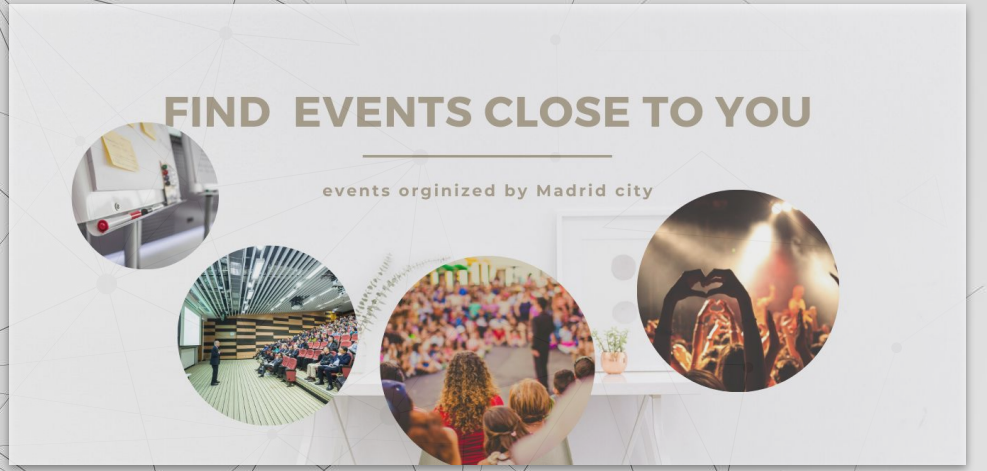


Project ODKG

Open Data and Knowledge graph

Group 5

Aïcha CHERIDI, Alexandre MOREL, Adrien SIMARD

POLITÉCNICA

Project Presentation 01

Data preprocessing 02

Ontology 03

Reconciliation 04

Materials

05 RDF Generation

06 Sparql queries

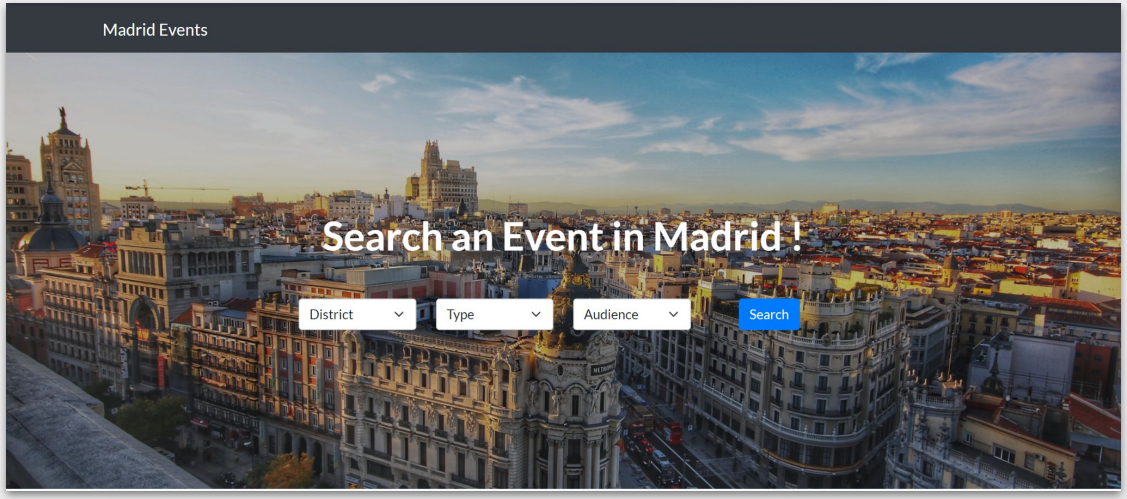
07 Application

08 Validation

09 Lesson Learned

01 Project Presentation

- The app will allow the users to **find** interesting madrid city **events** based on some criteria like district and targeted audience.
- The **dataset** is about **activities** of different types that are going to be held or that are being held.



- **datos.madrid.es**
- **Smart city domain**
- Project license: CC BY-SA 4.0
- Free to share and adapt under the terms of attribution and shareAlike.

- https://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=57be24206a91b510VgnVCM2000001f4a900aRCRD&vgnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnextfmt=default

02 Data preprocessing

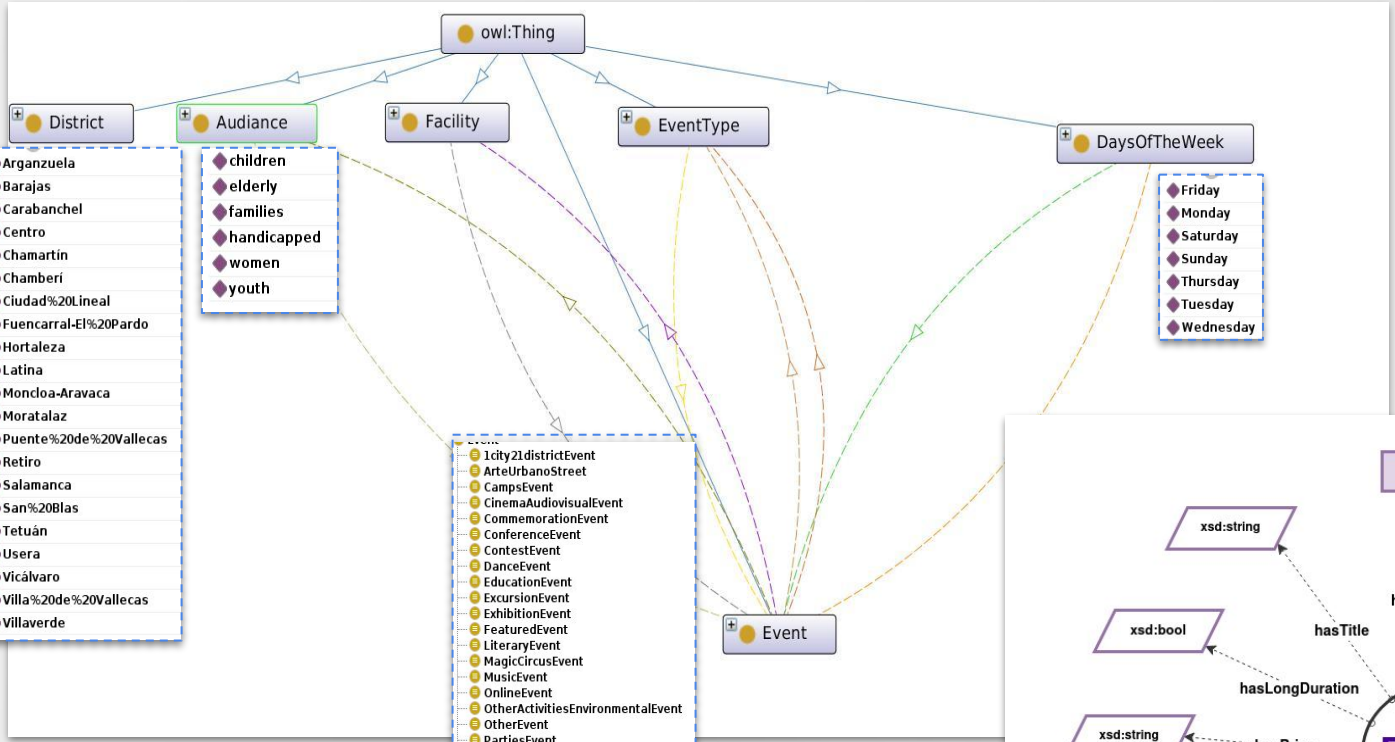
32. Rename column AUDIENCIA to audience
33. Rename column DISTRITO-INSTALACION to district
34. Remove column COORDENADA-X
35. Remove column COORDENADA-Y
36. Text transform on 17 cells in column accessibility: grel:value.replace('0','not_accessible')
37. Text transform on 69 cells in column accessibility: grel:value.replace('2','partially installed accessible for people with reduced mobility')
38. Text transform on 825 cells in column accessibility: grel:value.replace('1','accessible')
39. Text transform on 3 cells in column accessibility: grel:value.replace('3',' no information on accessibility')
40. Text transform on 5 cells in column accessibility: grel:value.replace('4','sign language')
41. Text transform on 10 cells in column accessibility: grel:value.replace('5','Tactile signage')
42. Text transform on 99 cells in column accessibility: grel:value.replace('6','magnetic induction')
43. Text transform on 10 cells in column accessibility: grel:value.replace('Tactile signage','tactile signage')
44. Split 961 cell(s) in column type into several columns by separator

eventID	EventTitle	price	longDuration	DaysOfTheWeek	Sunday	Saturday	Friday	Monday	Thursday	Wednesday	Tuesday	excludedDays	startingDate	EndingDate	hour	description
11457499	1ª Edición Festival Internacional de Luz de Madrid	free	true	L,M,X,J,V,S,D	Sunday	Saturday	Friday	Monday	Thursday	Wednesday	Tuesday		2021-10-29T00:00:00Z	2021-10-31T23:59:00Z		
11461367	35º Certamen coreográfico de Madrid	Entrada general: 15 euros	true	L,M,X,J,V,S,D	Sunday	Saturday	Friday	Monday	Thursday	Wednesday	Tuesday		2021-12-08T00:00:00Z	2021-12-12T23:59:00Z		
11482271	3º Muestra de cine finlandés en Madrid	free	false										2021-10-09T00:00:00Z	2021-10-09T23:59:00Z	19:00	edit En esta nueva edición presentamos un nuevo formato experimental más dinámico y participativo en la programación de nuestra cita anual con el cine finlandés en Hortaleza, y para todo

- Delete some columns
- Rename columns
- Merge the address (join of columns street, number, neighborhood)
- Coherence of values

03

Ontology

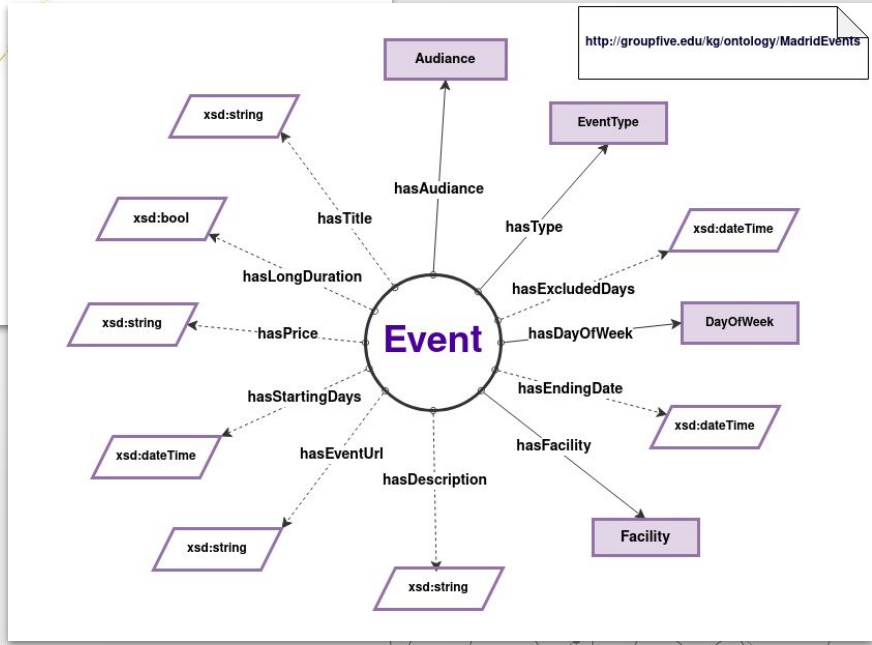


Equivalent To

Event
and (hasType some camps)

Disjoint With

Event, DaysOfTheWeek, District, Audiance



Ontology URI: “http://groupfive.edu/kg/ontology/MadridEvents#hasAudiance”
Resource URI : "http://groupfive.edu/kg/resources/event/{EVENTID}"

04 Reconciliation

- We reconcile our data with the **district** data of wikidata.
- We tried with other columns of our dataset (Facilities, Address) but it was impossible. Just **few data** and line could be reconcile.
- Wikidata doesn't match the **full expression**.

```
@prefix rr: <http://www.w3.org/ns/r2rml#>.
@prefix ont: <http://groupfive.edu/kg/ontology/MadridEvents#>.
@prefix schema: <https://schema.org/>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>.
@prefix vcard:<http://www.w3.org/2006/vcard/ns#>.
@prefix saref:<https://saref.etsi.org/core/>.
@prefix gtfs:<http://vocab.gtfs.org/terms#>.
@prefix bimerr:<http://bimerr.iot.linkeddata.es/def/occupancy-profile#>.
@prefix owl:<htpp://www.w3.org/2002/07/owl#>.

<#EVENT>
a rr:TriplesMAp ;

  rr:logicalTable [rr:tableName "EVENTSUPDATED"];

  rr:subjectMap [
       rr:template  "http://groupfive.edu/kg/resources/event/{EVENTID}";
       rr:class schema:Event;
  ];
  rr:predicateObjectMap [
      rr:predicate ont:hasTitle;
      rr:objectMap [
          rr:column "EVENTTITLE" ;
          rr:language "es";

      ];
  ];

  rr:predicateObjectMap [
      rr:predicate ont:hasDescription;
      rr:objectMap [
          rr:column "DESCRIPTION" ;
          rr:language "es";
      ];
  ];
```

05

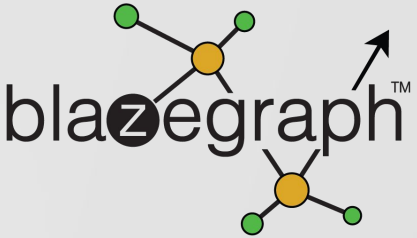
RDF Generation

The RDF generation steps:

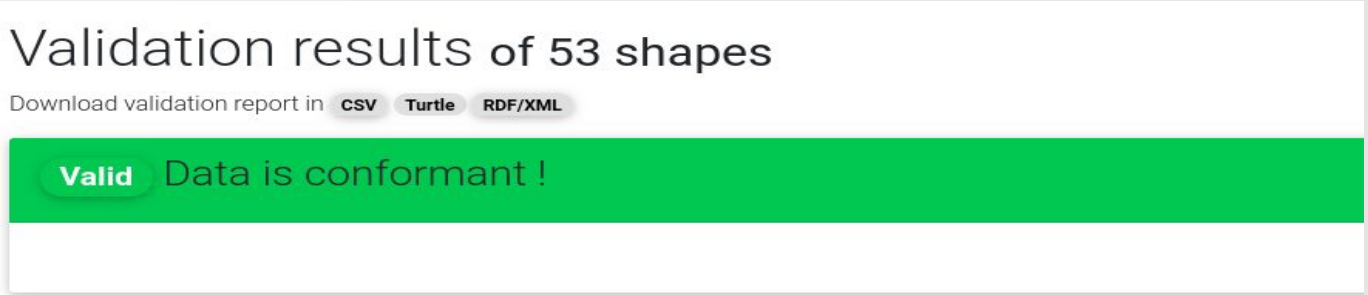
- Creating the mapping csv-rdf using R2RML
- Converting the csv to an RDF

06 Sparql Queries

- We used **Blazegraph** to publish our linked data and to make SPARQL queries.
- Our queries **adapt** to the search criteria selected by the user.
- The **value** of the selected attribute is **replaced** in the query when searching.



Application 07



08
Validation

```
@prefix :        <http://groupfive.edu/kg/ontology/MadridEvents#> .
@prefix afn:     <http://jena.apache.org/ARQ/function#> .
@prefix dash:    <http://datashapes.org/dash#> .
@prefix graphql: <http://datashapes.org/graphql#> .
@prefix owl:     <http://www.w3.org/2002/07/owl#> .
@prefix rdf:     <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs:    <http://www.w3.org/2000/01/rdf-schema#> .
@prefix sh:      <http://www.w3.org/ns/shacl#> .
@prefix swa:     <http://topbraid.org/swa#> .
@prefix tosh:    <http://topbraid.org/tosh#> .
@prefix xsd:     <http://www.w3.org/2001/XMLSchema#> .

[ rdf:type        sh:ValidationReport ;
  sh:conforms  true
] .
```

09 Lesson learned

Lesson learned during this project :

- Data **reconciliation** is more convenient to be executed during the **early stages** probably just after data preparation.
- **Ontology** is the trickiest part **reusing** existing ones it far more **efficient**, but unfortunately not many sources to find complete ontologies.
- Create a **simple** ontology in the beginning and then **refine** it, this approach makes debugging the reasoner much easier.
- Knowledge graph validation before building the application.

Thank you

Any questions ?