

OPEN DATA AND KNOWLEDGE GRAPHS - HANDS ON

Group 07 : M^a Delfina Ramos Vidal, Thibault Debard, Yuzhi Wu and Yuxue Liang



POLITÉCNICA

UNIVERSIDAD
POLITÉCNICA
DE MADRID

Introduction

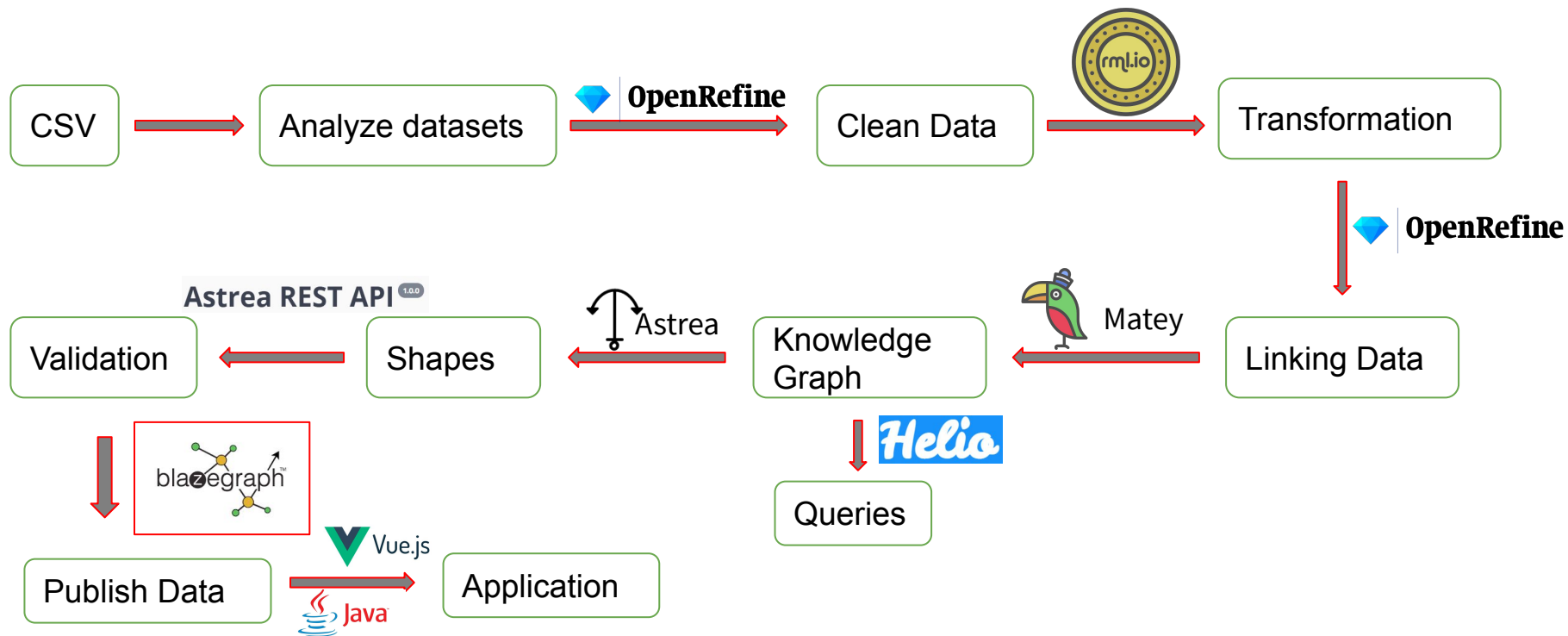
❖ Dataset: Horizon 2020

- Organizations
- Projects
- Fellows

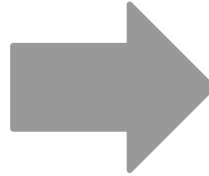
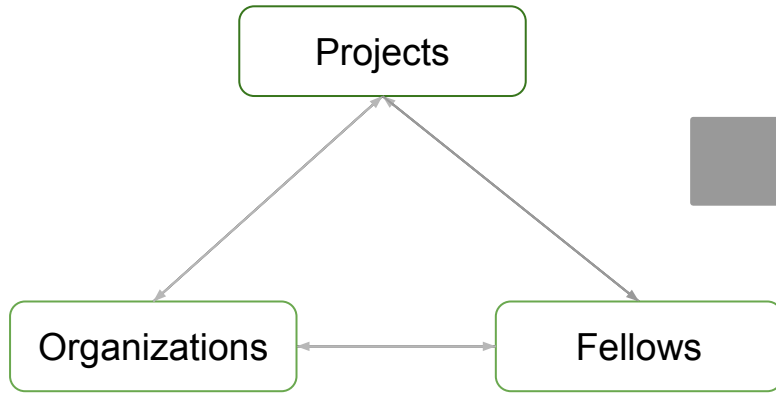


- ## ❖ Idea: design a web page with user interfaces that permits to search with filters and display the information needed in a user friendly way.

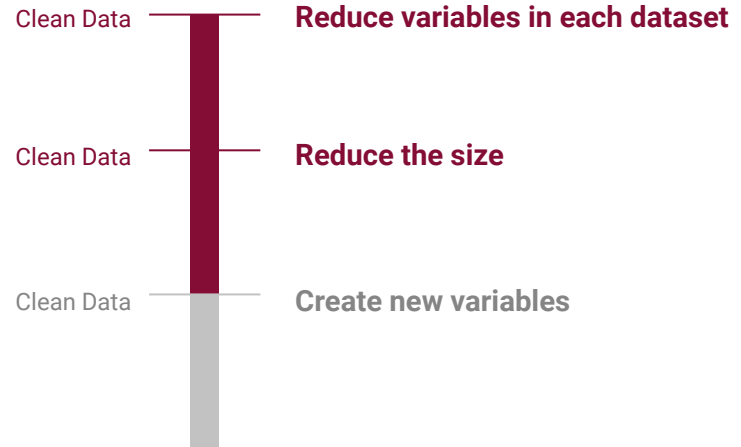
Methodology



Dataset:H2020



Data processing



Dataset Requirements

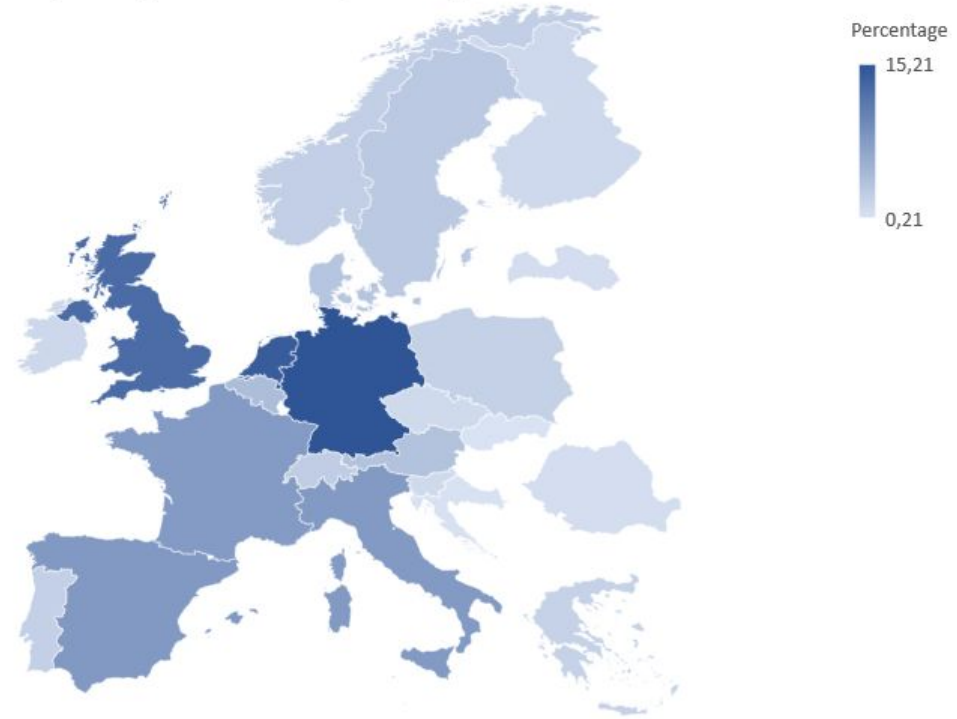
- ❖ Public procurement domain.
- ❖ Open license. —————>
<https://data.europa.eu/euodp/en/copyright>
- ❖ Available in CSV format.
- ❖ Come from multiple data sources.



Data representation

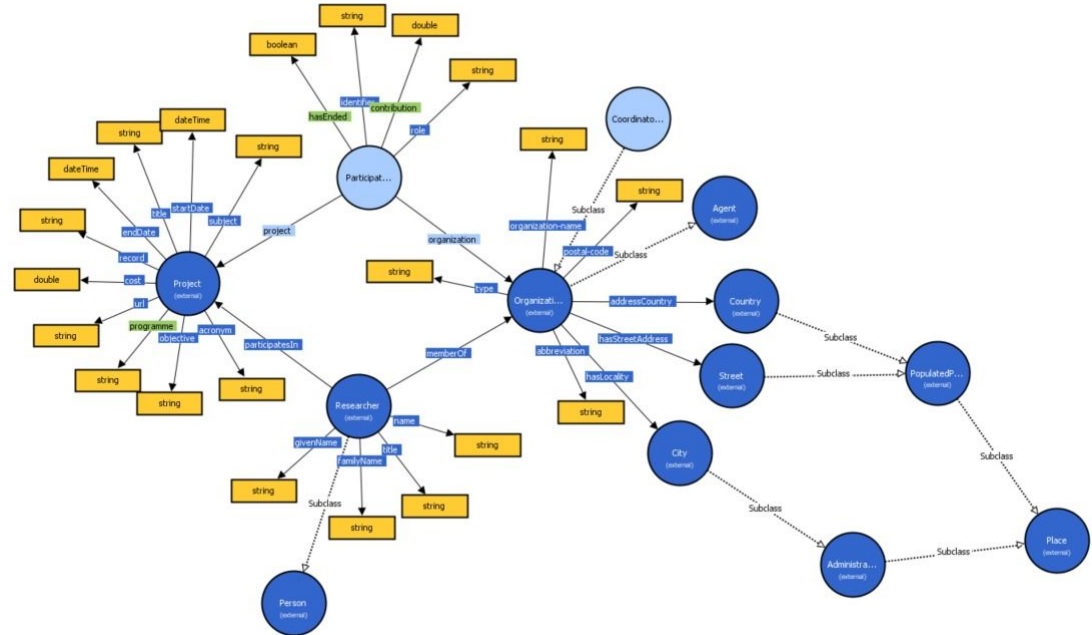
- Different country, different project
- First approach of the project
- Germany invest in various project

Project representation by country



RDF Graph

- 5 main classes
- Data's information
⇒ researcher's information
⇒ University's address
⇒ Project contribution



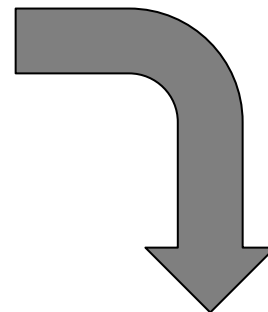
Cleaning and linking datas

Toutes	projectId	projectAcronym	fundingScheme	organizationId	title	firstName	lastName
1.	650216	OntoPerson	MSCA-IF	999598678	DR	Mariusz	Grygianiec
2.	652745	AUTOOLIGO	MSCA-IF	999990267	MS	Petra	Menova
3.	652755	MOLECOPS	MSCA-IF	99998435	DR	Julien	Brachet
4.	652757	ARCGATE	MSCA-IF	999497507	DR	Takamasa	Tsubouchi
5.	652782	FOSICAV	MSCA-IF	999844864	DR	Thomas	Dedieu
6.	652837	ISOTOP	MSCA-IF	999905101	DR	Benno	Rem
7.	652856	CLEANLIB	MSCA-IF	999991722	DR	Hervé	Baudry
8.	652862	IMPRECSIM	MSCA-IF	999988812	DR	Darius	Markauskas
9.	652871	FIBERSTAR	MSCA-IF	999899572	DR	Daniele	Tosi
10.	652890	EMTECS	MSCA-IF	999974262	DR	Yahui	Wang

Creation of a
“fullName” class

Creation of an
organization ID

organizationId	title	firstName	fullName	wikiFullName
999598678	DR	Mariusz	Mariusz Grygianiec	http://www.wikidata.org/entity/Q63123374
999990267	MS	Petra	Petra Ménová	http://www.wikidata.org/entity/Q89870613
99998435	DR	Julien	Julien Brachet	
999497507	DR	Takamasa	Takamasa Tsubouchi	http://www.wikidata.org/entity/Q59833423
999844864	DR	Thomas	Thomas Dedieu	
999905101	DR	Benno	Benno Rem	
999991722	DR	Hervé	Hervé Baudry	http://www.wikidata.org/entity/Q24342393
999988812	DR	Darius	Darius Markauskas	http://www.wikidata.org/entity/Q59913986
999899572	DR	Daniele	Daniele Tosi	http://www.wikidata.org/entity/Q57560967
999974262	DR	Yahui	Yahui Wang	http://www.wikidata.org/entity/Q43137476
999902385	DR	Francesca	Francesca Scrinzi	http://www.wikidata.org/entity/Q89113592
999838074	DR	Eleftherios	Eleftherios Giovanis	http://www.wikidata.org/entity/Q88017445
999642037	DR	Araceli	Araceli Pérez-Sanz	http://www.wikidata.org/entity/Q60655695
909875521	PROF	Atinati	Atinati Mamatsashvili-Kobakhidze	
998096051	PROF	Atinati	Atinati Mamatsashvili-Kobakhidze	
99998435	DR	Christakis	Christakis Constantinides	

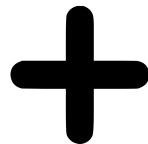


Link with
wikidata



SPARQL Queries

```
<http://www.example.com/group07/ontology/project/822625> a schema:Project;
  rdfs:label "822625";
  dg:objective "ADMIGOV aims to promote an alternative migration governance model"^^xsd:string;
  dbpedia-owl:cost "2788870"^^xsd:double;
  vocab:contribution "2788870"^^xsd:double;
  dcat:record "221876"^^xsd:string;
  cerif:acronym "ADMIGOV"^^xsd:string;
  vocab:programme "H2020-EU.3.6.1.1.+H2020-EU.3.6.1.2."^^xsd:string;
  dct:subject "MIGRATION-02-2018"^^xsd:string;
  dct:title "Advancing Alternative Migration Governance"^^xsd:string;
  schema:startDate "2019-02-01T00:00:00Z"^^xsd:dateTime;
  schema:endDate "2023-01-31T00:00:00Z"^^xsd:dateTime;
  vcard:url "http://admigov.eu~IRI".
```



```
PREFIX country:<http://www.example.com/group07/Country/>

SELECT DISTINCT ?name ?postcode ?locality

WHERE
{
  ?x sch:addressCountry <http://www.example.com/group07/Country/Germany>.
  ?x ns:organization-name ?name.
  ?x ns:hasLocality ?locality.
  ?x ns:postal-code ?postcode
}

ORDER BY asc(?x)

LIMIT 100
```

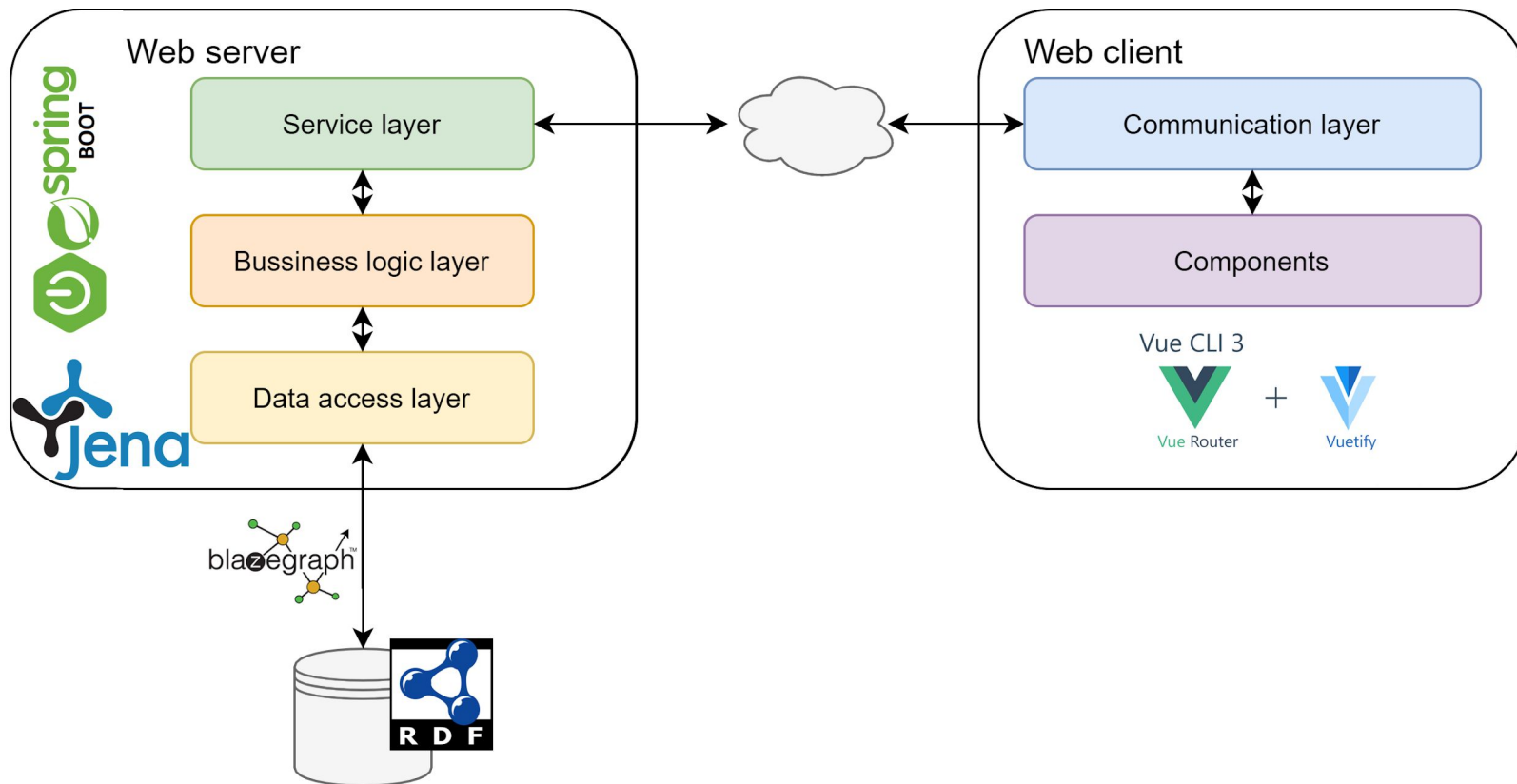
RDF File

SPARQL Query

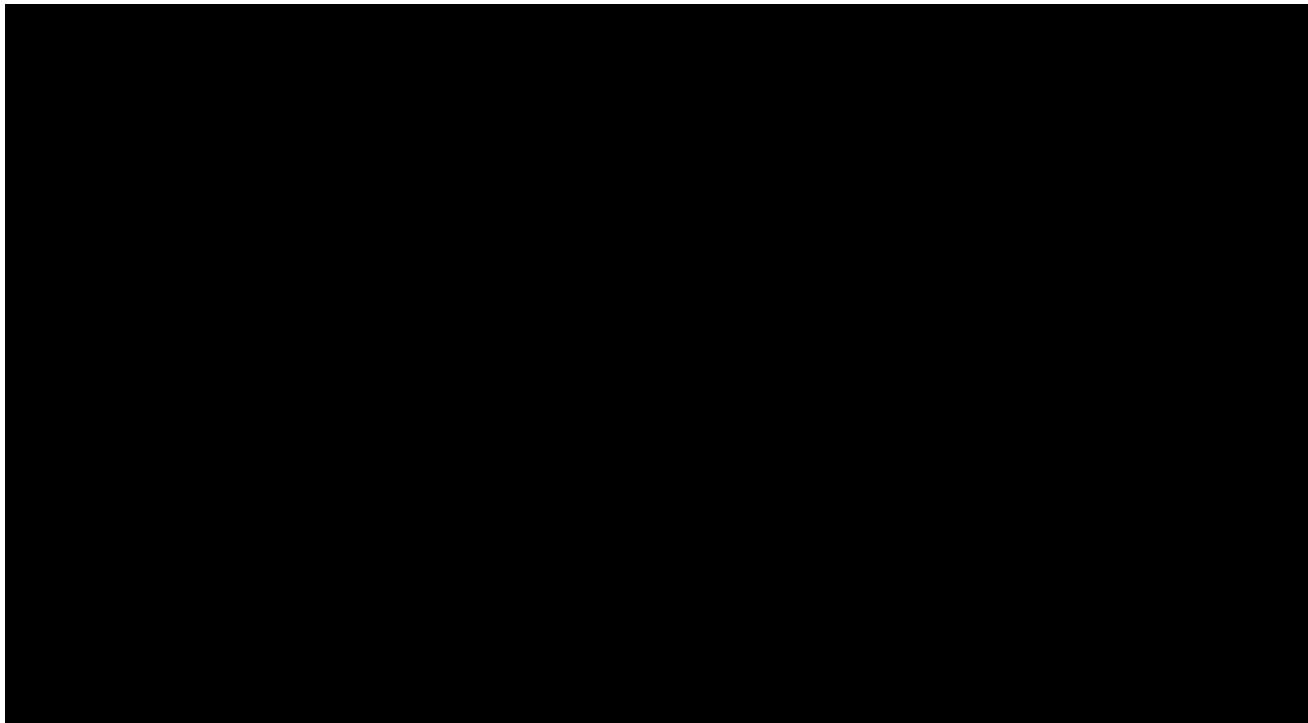
	name	postcode	locality
1	Bayer	51373	http://www.example.com/group07/City/Leverkusen
2	MAGIX Software GmbH	10589	http://www.example.com/group07/City/Berlin
3	University of Applied Sciences Offenburg	77652	http://www.example.com/group07/City/Offenburg
4	Deutscher Wetterdienst	63067	http://www.example.com/group07/City/Offenbach%20am%20Main
5	Jena University Hospital	07743	http://www.example.com/group07/City/Jena
6	Free Hanseatic City of Bremen	28195	http://www.example.com/group07/City/Free%20Hanseatic%20City%20of%20Bremen
7	University of Mannheim	68161	http://www.example.com/group07/City/Mannheim
8	TU Dresden	01069	http://www.example.com/group07/City/Dresden
9	University of Stuttgart	70174	http://www.example.com/group07/City/Stuttgart

Result of the
query

System architecture



Demo video



Conclusions

- Improved skills:
 - Data refining
 - SPARQL Queries
 - Data linking
 - RDF generation
- Learned the impact of Open Data



Future work:

- Implement detailed pages for individuals where we can see other elements related to them.
- Implement the whole datasets

THANK YOU FOR YOUR ATTENTION

Group 07 : M^a Delfina Ramos Vidal, Thibault Debard, Yuzhi Wu
and Yuxue Liang



POLITÉCNICA

UNIVERSIDAD
POLITÉCNICA
DE MADRID