



From Data Spaces to Decentralized Knowledge Spaces

The Public Procurement use case

David Chaves-Fraga & Oscar Corcho

Ontology Engineering Group
Universidad Politécnica de Madrid, Spain

Postdoctoral research at Ontology Engineering Group (UPM) and DTAI (KULeuven)

PhD in Artificial Intelligence at Universidad Politécnica de Madrid (2021)



POLITÉCNICA

Co-chair W3C Community Group on Knowledge Graph Construction (2019-now)



Main author of the Spanish National guideline for generating RDF graphs from Tabular Data (2021)



Main Coordinator of Open Summer of Code Spain (2018-2021)



Principal Scientist of the Public Procurement Data Space (2022-2023)



European
Commission

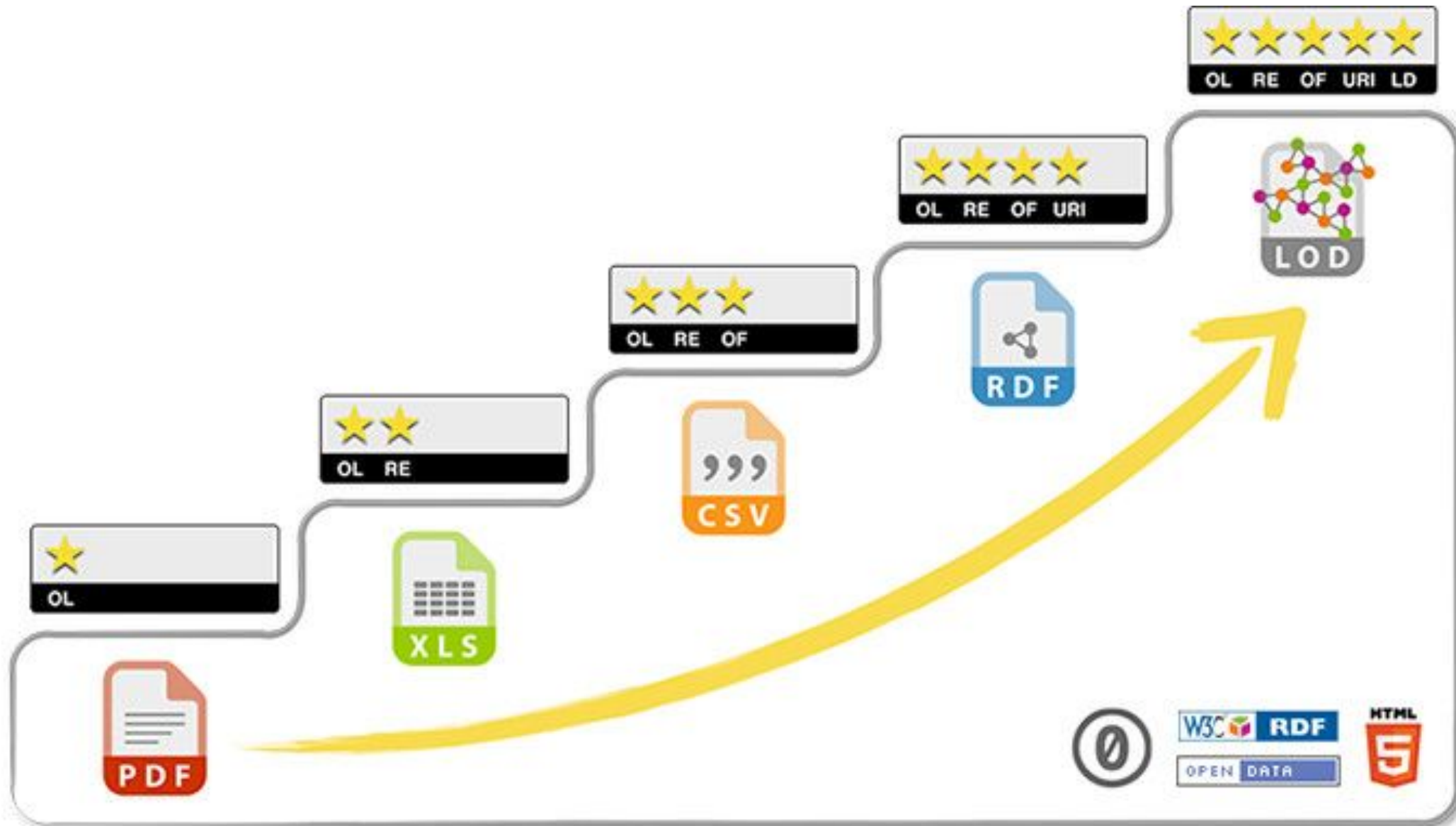
Why should we incorporate **KNOWLEDGE** in Data Spaces?

How can we incorporate **KNOWLEDGE** in Data Spaces?

*“Data spaces are an abstraction in data management that aim to overcome some of the problems encountered in **data integration system**.”**

Data (or Knowledge) integration is complex problem to which we did not pay enough attention

* Geisler, S., Vidal, M. E., Cappiello, C., Lóscio, B. F., Gal, A., Jarke, M., ... & Rehof, J. (2021). Knowledge-driven data ecosystems toward data transparency. ACM Journal of Data and Information Quality (JDIQ), 14(1), 1-12.

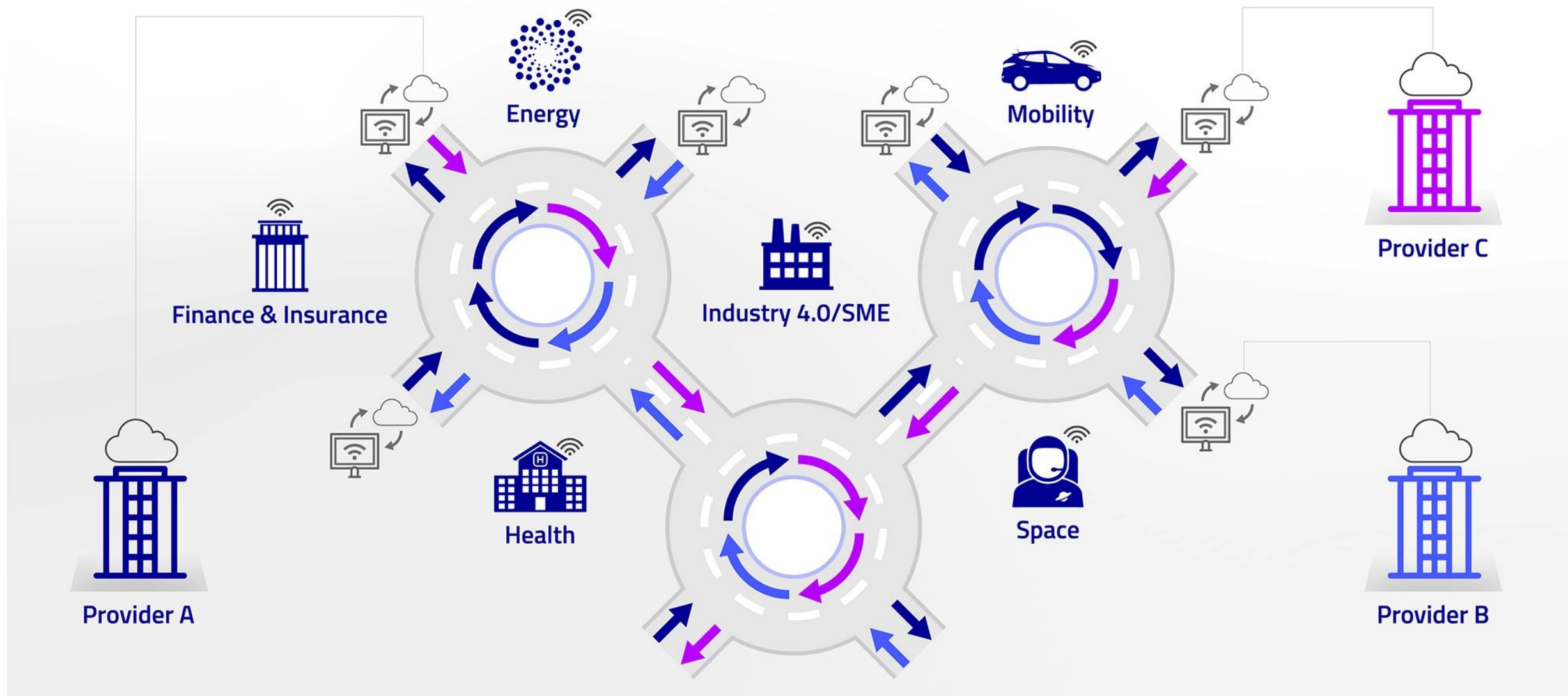


What is the most used format for uploading data on open data portal?

| Formato |
|-------------------|
| CSV (8854) |
| XLS (6339) |
| JSON (4517) |
| HTML (3540) |
| XML-APP (2593) |
| PDF (2094) |
| ASCII (1909) |
| PC-Axis (1302) |
| RDF-Turtle (1192) |
| XLSX (1186) |
| Mostrar más |

| Formats |
|-------------------|
| CSV (86834) |
| TXT (65699) |
| ZIP (51600) |
| JSON (45900) |
| GMZ (44712) |
| HTML (42959) |
| PDF (35421) |
| XLS (24418) |
| WMS (21800) |
| SHP (19507) |
| Show More Formats |

| Formatos |
|------------|
| CSV (371) |
| XLS (169) |
| XML (129) |
| XLSX (122) |
| WMS (29) |
| RDF (21) |
| prj (7) |
| SHP (7) |
| SHX (7) |
| ZIP (7) |
| dBase (6) |
| WMTS (5) |
| JSON (4) |
| KML (3) |
| BDF (2) |
| DAT (2) |



Use **OWL** to create and manage your **vocabularies**

Use **RDF** to **represent** the entities/data

Use **SPARQL** to **query** the data

Use **[R2]RML** to **create** RDF from other sources

Use **SHACL** to **validate** the quality of your data





Previous experience in PP

CONTSEM

Spanish National Project
collaboration (2014)

Collaboration to create the PPROC
(Public Procurement Ontology).



More information:

<http://contsem.iasoft.es/> (project)

<http://contsem.unizar.es/def/sector-publico/pproc.html> (PPROC ontology)

TheyBuyForYou (TBFY)

EU H2020 Project 2018-2021 (3 years)

Innovation Action

Enhancing Public Procurement in the
EU through Constructing and Exploiting
an Integrated Knowledge Graph

More information:

<https://theybuyforyou.eu> (project)



NextProcurement

EU CEF Project 2022-2024 (3 years)

Public Procurement in Spain in collaboration with different stakeholders.

Includes the whole processing (data source processing, storage, natural language processing...)

More information:

<http://nextprocurement-project.com/#> (project)

<https://procurement.linkeddata.es/> (our contributions)



AI4GOV Master

Teaching PP stakeholders (yearly)

Providing education on:

- (i) the management of AI-based public services
- (ii) the usage of AI in public administration's operations
- (iii) the governance of AI in the public sector and for the delivery of public services to society.

More information:

<https://ai4gov-master.eu/> (website)





The Public Procurement Data Space **(Proof of Concept)**

Constructing the EU decentralized Data Space for Public Procurement

- Homogenize the access to all public procurement data across Europe
- First semantic data architecture of this magnitude in the world
- Calculate standard transparency indicators for each member state
- Researching on: vocabularies, resources maintainability, federated query processing, semantic data ingestion, query performance/scalability...

• Partners:

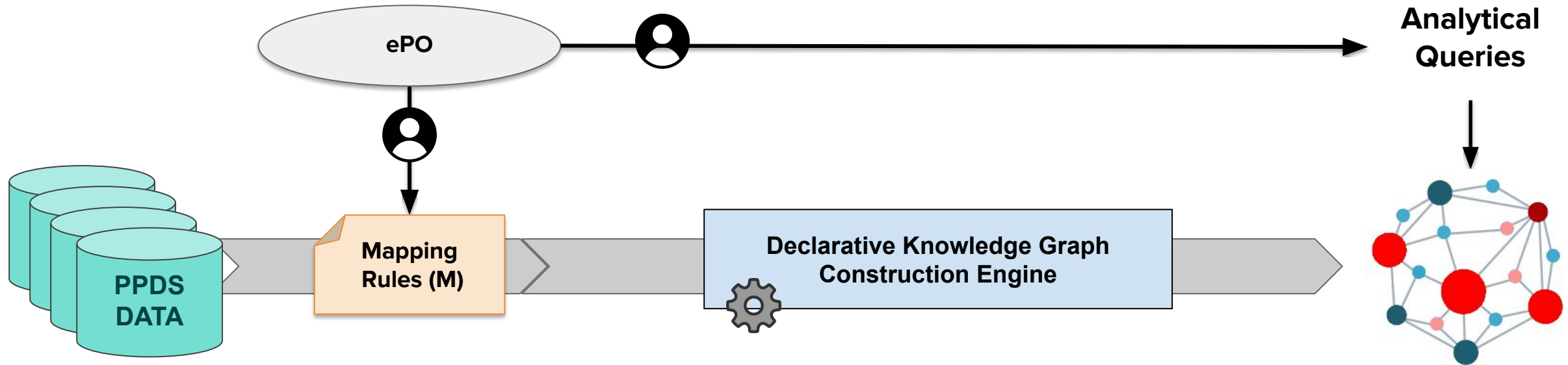


Analyse input data from CODICE and how map the values to ePO (expert knowledge)

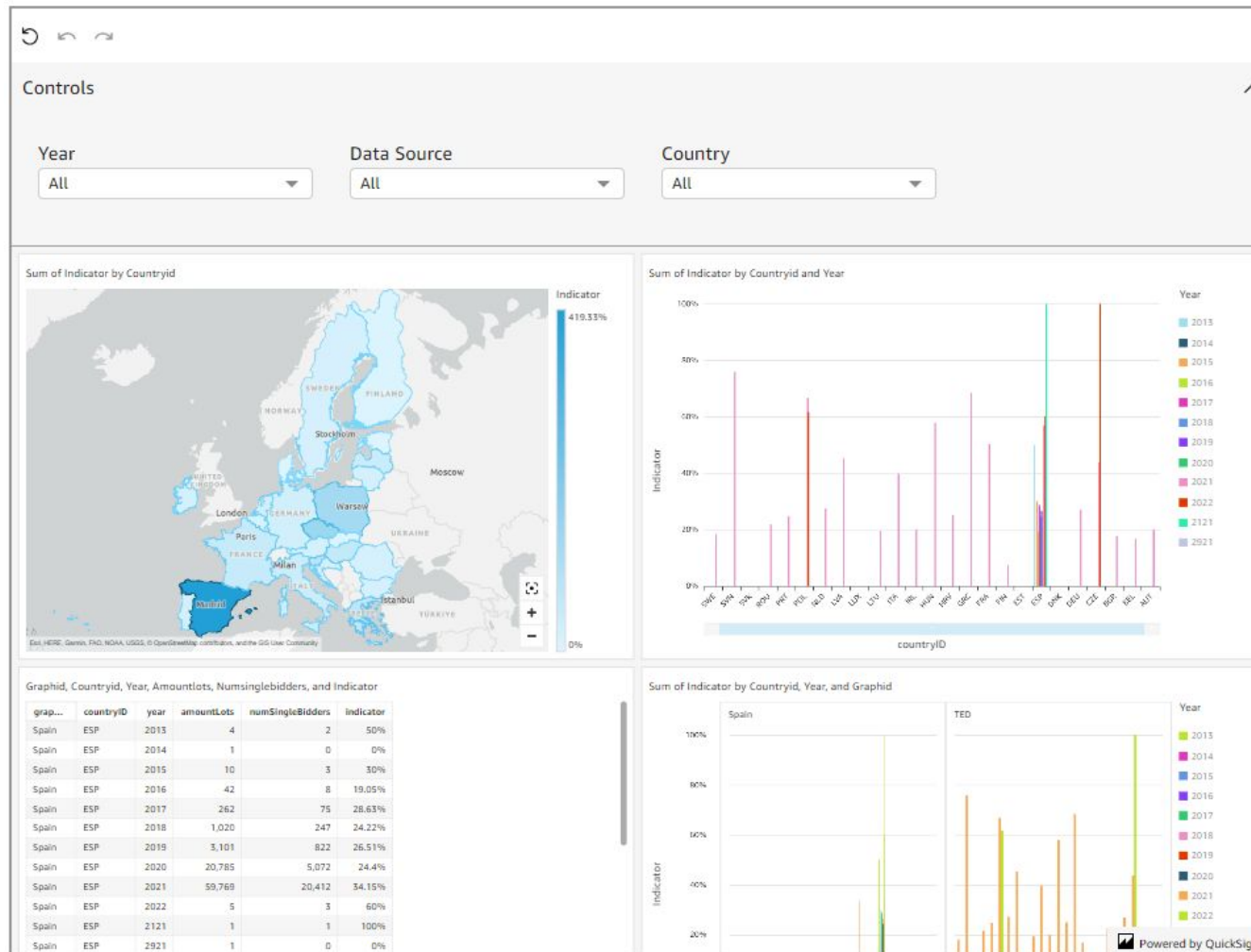
```

2021-11-10T17:54:07.824+01:00 xmlns:at="http://portal.org/atompub/commonschemas/1.0" />
<entry>
  <id>https://contrataciondelestado.es/sindicacion/licitacionesPerfilContratante/8610314</id>
  <link href="https://contrataciondelestado.es/wps/poc?uri=deeplink:detalle_licitacion&idEvI=8Wz%2BZU0V0qgBPRBxZ4nJ%2Fg%3D%3D"/>
  <summary type="text">Id licitación: 73/2021 (GEST 9471/2021); Órgano de Contratación: Pleno del Ayuntamiento de Almuñécar;
  Importe: 4807594.05 EUR; Estado: PUB</summary>
  <title>La ejecución de las obras descritas y recogidas en el documento de Proyecto Básico y de Ejecución redactado por los
  arquitectos Nico Alexander Heinz, Alejandro Sánchez Fernández, Francisco Martínez Manso y Rafael Soler Márquez,
  pertenecientes a la mercantil MERCADO DE ALMUÑÉCAR UTE, adjudicataria del contrato de servicios para la redacción de
  proyecto, dirección de obras y trabajos de seguridad y salud de los nuevos mercado y parking municipales de Almuñécar,
  aprobado por Acuerdo del Ayuntamiento Pleno de fecha 17 de septiembre de 2021.
  n objeto del presente contrato las OBRAS PARA LA EJECUCIÓN DEL NUEVO MERCADO MUNICIPAL DE ALMUÑÉCAR Y APARCAMIENTO PÚBLICO.</
  title>
  <updated>2021-11-11T08:15:16.800+01:00</updated>
  <cac:place-ext:ContractFolderStatus>
    <cbc:ContractFolderID>73/2021 (GEST 9471/2021)</cbc:ContractFolderID>
    <cbc:place-ext:ContractFolderStatusCode languageID="es" listURI="https://contrataciondelestado.es/codice/cl/2.04/SyndicationContractFolderStatusCode-2.04.gc">PUB</cbc:place-ext:ContractFolderStatusCode>
    <cac:place-ext:LocatedContractingParty>
      <cbc:ContractingPartyTypeCode listURI="http://contrataciondelestado.es/codice/cl/1.04/ContractingAuthorityCode-1.04.gc">3</cbc:ContractingPartyTypeCode>
      <cbc:ActivityCode listURI="http://contrataciondelestado.es/codice/cl/2.05/ContractingAuthorityActivityCode-2.05.gc">22</cbc:ActivityCode>
      <cbc:BuyerProfileURIID>https://contrataciondelestado.es/wps/poc?uri=deeplink:perfilContratante&idBp=q33vTay0qWh7h85%2Fpmmfw%3D%3D</cbc:BuyerProfileURIID>
      <cac:Party>
        <cbc:WebsiteURI>http://contratacion.almunecar.es</cbc:WebsiteURI>
        <cac:PartyIdentification>
          <cbc:ID schemeName="DIR3">L01180172</cbc:ID>
        </cac:PartyIdentification>
      </cac:Party>
    </cac:place-ext:LocatedContractingParty>
  </cac:place-ext:ContractFolderStatus>
</entry>

```



Public Procurement Data Space (PPDS) POC Demo - Indicator [1]: Single bidder



The objective was to calculate three different indicators of PP using data from Spain and from TED, but integrated.

| Indicator |
|--|
| [1] Single bidder |
| [2] No calls for bids |
| [3] Publication rate |
| [4] Cooperative procurement |
| [5] Award criteria |
| [6] Decision speed |
| [7] SME contractors |
| [8] SME bids |
| [9] Procedures divided into lots |
| [10] Missing calls for bids |
| [11] Missing seller registration numbers |
| [12] Missing buyer registration numbers |

Indicator 6 query results

| graphID | countryID | year | avgdays |
|---------|-----------|------|---------------------|
| Spain | ESP | 2013 | 21.25 |
| Spain | ESP | 2014 | 627 |
| Spain | ESP | 2015 | 61.6 |
| Spain | ESP | 2016 | 43.88888888888889 |
| Spain | ESP | 2017 | 183.642857142857143 |
| Spain | ESP | 2018 | 196.254681647940075 |
| Spain | ESP | 2019 | 241.175395858708892 |
| Spain | ESP | 2020 | 111.572499279331219 |
| Spain | ESP | 2021 | 49.883622285291214 |
| Spain | FRA | 2020 | 49 |
| Spain | FRA | 2021 | 87.857142857142857 |
| Spain | ITA | 2021 | 19 |
| Spain | PRT | 2021 | 23 |

BUT

| graphID | countryID | year | avgdays |
|---------|-----------|------|---------------------|
| Spain | ESP | 2013 | 21.25 |
| Spain | ESP | 2014 | 627 |
| Spain | ESP | 2015 | 61.6 |
| Spain | ESP | 2016 | 43.88888888888889 |
| Spain | ESP | 2017 | 183.642857142857143 |
| Spain | ESP | 2018 | 191.194139194139194 |
| Spain | ESP | 2019 | 240.273665048543689 |
| Spain | ESP | 2020 | 6.041945666730438 |
| Spain | ESP | 2021 | 27.55762868063634 |
| Spain | ESP | 2022 | -55.5 |
| Spain | ESP | 2121 | -36499 |
| Spain | ESP | 2921 | -328719 |
| Spain | FRA | 2020 | 49 |
| Spain | FRA | 2021 | 87.857142857142857 |
| Spain | ITA | 2021 | 19 |
| Spain | PRT | 2021 | 23 |

Public Procurement:

- Data quality policies are mandatory
- Analytic queries are complex
- Very distributed environment
- Data but also assets governance will be very relevant

General remarks:

- Data without semantics/meaning is not useful
- Reusability needs to be ensured (stop publishing data for the sake of publishing)
- Administration needs to improve its (technical) skills
- Decentralization will be the key for successful Data Spaces
- F.A.I.R principles implemented



From Data Spaces to Knowledge Space

The Public Procurement use case

David Chaves-Fraga & Oscar Corcho

Ontology Engineering Group
Universidad Politécnica de Madrid, Spain



david.chaves@upm.es



dchavesf



02/06/2023



Alicante