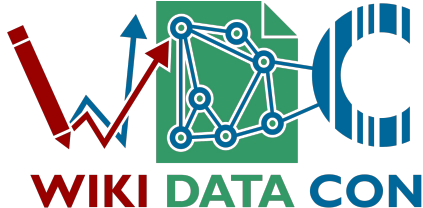


Improving access to Culture Heritage: data.cervantesvirtual.com

Gustavo Candela Romero

gustavo.candela@cervantesvirtual.com



Universitat d'Alacant
Universidad de Alicante



FUNDACIÓN BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES

Index

1. Introduction
2. Keys to success
3. Wikidata and BVMC
4. Future work
5. References



The screenshot shows the FBVMC Labs website. At the top, there is a dark navigation bar with the FBVMC Labs logo on the left and the links 'DATOS ENLAZADOS', 'NOVEDADES', and 'LABS' on the right. Below the navigation bar is a large banner image of an open book with the text 'FBVMC' and a progress indicator. To the right of the banner is a list of news items with dates and titles. Below the banner is a search bar with the placeholder text 'Introduce un término de búsqueda' and a search icon. Below the search bar is the text 'Descubre el catálogo de datos abiertos de la Biblioteca Virtual Miguel de Cervantes'. To the right of the search bar is a map of Europe with a red circular overlay containing the text 'Explora por lugar' and 'Mapa >>'. The news items are:

- 29 / 10 / 2017 Ejemplo de aplicación de ElasticSearch
- 10 / 10 / 2017 La BVMC seleccionada para los premios Aporta 2017 de red.es
- 10 / 10 / 2017 Publicados los artículos del congreso DATECH
- 6 / 10 / 2017 Introducción a Elasticsearch
- 6 / 10 / 2017 Portales de la BVMC de autores que han ganado el premio Cerv...
- 27 / 9 / 2017 Sentencia SPARQL para obtener Autores que han ganado el prem...



¿Sabes cómo usar la base de datos Wikidata?



Datos enlazados
Los datos corresponden al catálogo principal y se...



Labs
Prototipos y nuevos desarrollos.

data.cervantesvirtual.com

1. Introduction

- In February 2015, the first release of the linked open data website based on **RDA and FRBR** data.cervantesvirtual.com was launched. The project intends to promote **data sharing, interoperability, data re-use and dissemination of best practices**.
- Starting from scratch, the project is no longer a demo, but a rich source of lessons learnt in order to stimulate innovative types of projects.

1. Introduction

Step by step



2012



Relational
Database

2015



RDF
Repository

TPDL 2015
IODC 2016
Semantic Web Journal



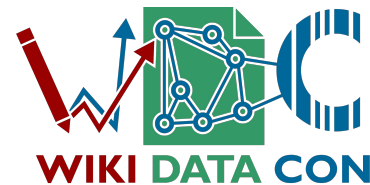
Stanford Prize for Innovation
in Research Libraries



RESOURCE DESCRIPTION & ACCESS

1. Introduction

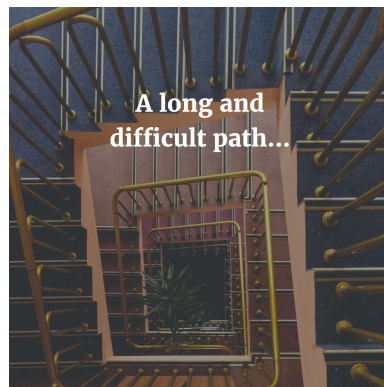
data.cervantesvirtual.com in 2017



1. Introduction

Why we did it?

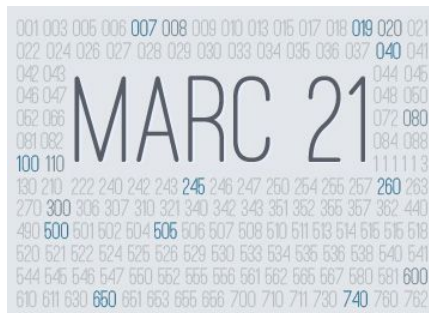
- Open our traditional catalog to the world for human and computers.
- Provide a public interface for querying the dataset according to international recommendations (SPARQL).
- Establish relationships and links with broadly used data sets such as VIAF and Wikidata.
- Improve the catalog and promote reuse.



2. Keys to success

Preprocessing of sources

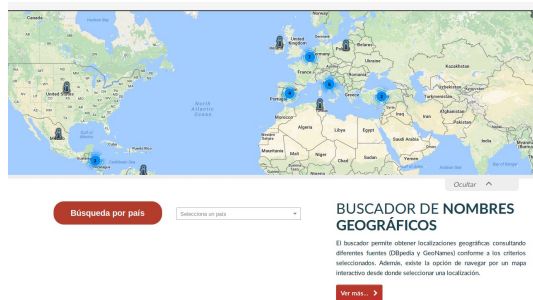
- Since some fields are required (for example, field 245 containing the title) while some others are optional or user-defined, **the homogeneity of the data across libraries cannot be guaranteed**. Furthermore, the content of a field can be expressed with different conventions, in different languages, or it may contain typos.
- These features **represent a challenge** when MARC21 records must be shared between libraries.



2. Keys to success

Preprocessing of sources

- However, further refinements are needed for the recognition and extraction of implicit relationships expressed in natural language, such as geographic locations and dates.

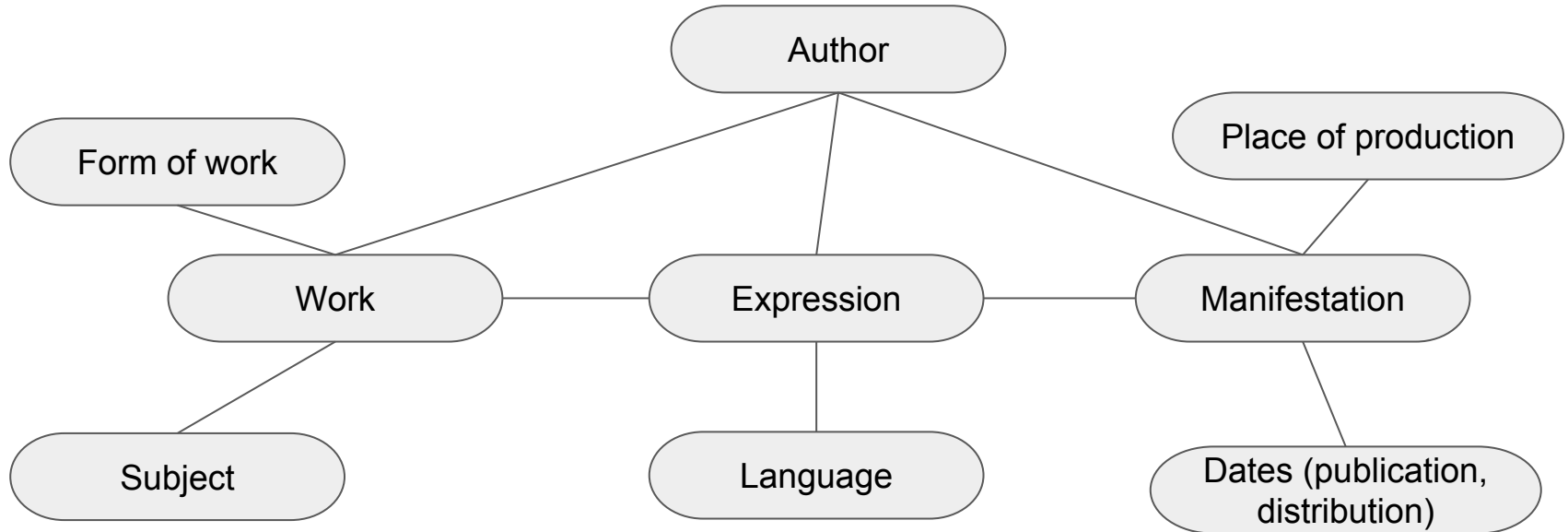


- En **Sevilla**, : en la imprenta de Joseph Padrino ..., [entre 1748 y 1775]
- Sevilla**, : por Thomas Lopez de Haro ..., , 1679
- [**Sevilla** : s.n., 1760]
- Impreso en **Sevilla** : por Juan Francisco de Blas..., 1693
- Hispli** :, Antonius Martinez, Alfonsus de Portu et Bartholomaeus Segura, 1477

2. Keys to success

Reuse of vocabularies (RDA and FRBR)

www.rdaregistry.info/



2. Keys to success

Identify access points

Entity	URI
Person	http://data.cervantesvirtual.com/person/{id}
Corporate Body	http://data.cervantesvirtual.com/corporatebody/{id}
Family	http://data.cervantesvirtual.com/family/{id}
Work	http://data.cervantesvirtual.com/work/{id}
Expression	http://data.cervantesvirtual.com/expression/{id}
Manifestation	http://data.cervantesvirtual.com/manifestation/{id}
Country	http://data.cervantesvirtual.com/country/{id}
Date	http://data.cervantesvirtual.com/date/{id}
Language	http://data.cervantesvirtual.com/language/{id}

2. Keys to success

Metadata enrichment



WIKIPEDIA
The Free Encyclopedia



Latin Place Names

RBMS/BSC Latin Place Names File

res

Research and
Education Space

2. Keys to success

Increase visibility and promote interoperability

- **Social Media (Facebook and Twitter)**
- **Conferences**
- **Sharing data**
- **SEO techniques**
- **Technology blog**
- **Github profile**
- **Encouraging students at the university**



3. Wikidata and BVMC

Wikidata properties

- <https://www.wikidata.org/wiki/Property:P2799> BVMC Person id (~7000 links)
- <https://www.wikidata.org/wiki/Property:P3976> BVMC Work id (~150 links)

Some examples of possible additional properties:

- BVMC Journal id
- BVMC Location id
- BVMC Date id
- BVMC Manuscript id



FUNDACIÓN BIBLIOTECA VIRTUAL
MIGUEL DE CERVANTES



3. Wikidata and BVMC

Reusing Wikidata

- Wordpress plugin for querying wikidata authors of literary movements.

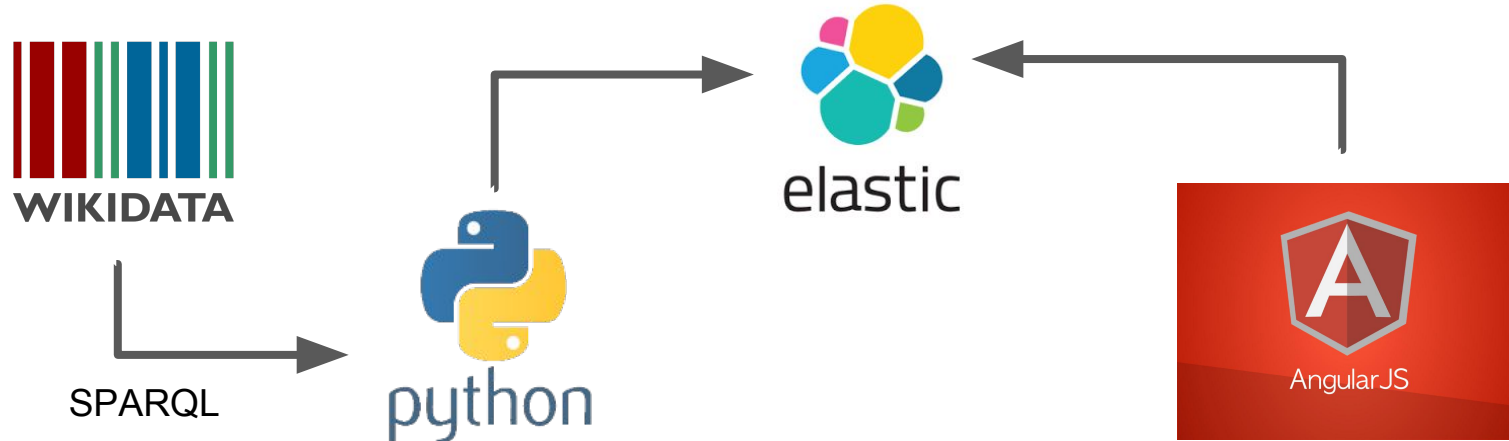
```
$result = $sparql->query(
    'SELECT * WHERE {'.
        '?writer wdt:P106 wd:Q36180 .' .
        ' OPTIONAL {?writer wdt:P18 ?image } .' .
        ' OPTIONAL {?writer wdt:P569 ?_date_of_birth .' .
            '      BIND( year(?_date_of_birth) as ?birthYear )} .' .
        ' OPTIONAL {?writer wdt:P570 ?_date_of_death .' .
            '      BIND( year(?_date_of_death) as ?deathYear )} .' .
        '?writer wdt:P135 wd:'. $movement .' .' .
        '?writer rdfs:label ?label .' .
        ' FILTER ( lang(?label) = "es" )' .
    '} ORDER BY ?label ' .
    'LIMIT '.$numresults
);
```



3. Wikidata and BVMC

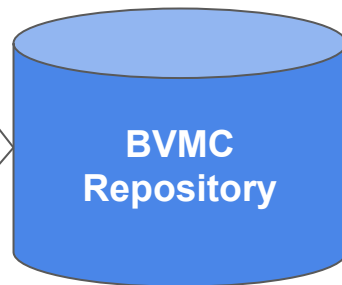
Reusing Wikidata

- Elasticsearch application example for BVMC linked authors.
- <http://data.cervantesvirtual.com/autor/buscador.html>
- <https://github.com/fbvmc/elasticsearch-tutorial>



4. Future work

Keeping on exploring and innovating



I Still Haven't Found What I'm Looking For...

4. References

- <http://data.cervantesvirtual.com>
- <http://data.cervantesvirtual.com/geosearch>
- galatea.cervantesvirtual.com
- [SPARQL endpoint](#)
- Migration of a library catalogue into RDA linked open data, Semantic Web Journal, 2017 [online](#)
- Transformation of a Library Catalogue into RDA Linked Open Data. [TPDL 2015](#)
- Semantic Enrichment on Cultural Heritage collections: A case study using geographic information. doi>[10.1145/3078081.3078090](https://doi.org/10.1145/3078081.3078090)
- <http://www.rdaregistry.info/>
- <https://www.ifla.org/publications/functional-requirements-for-bibliographic-records>

