European Linked Data Contest 2016 submission:

DATOS·BNE·es *beta

The data service of the National Library of Spain (BNE), datos.bne.es, makes available a wide collection of Linked Open Data under a Public Domain License (CC0). This knowledge graph with more than 10 million connected entities using open web standards serves two main goals: (1) enhanced end-user experiences when searching and browsing the catalogue, and (2) enable reuse and repurpose of the data via open APIs and SPARQL . The service, in continuous development, has been identified by BNE library managers as the future main entry point to the vast cultural collection of the library for both end-users and applications.

Stairs and main entrance to the National library of Spain with monuments to San Isidoro, Alonso Berruguete, Alfonso X el Sabio by José Alcoverro (1835-1908). - Madrid, Spain

THE PROJECT

The data service is the result of a long-term collaboration, started in 2010, between the National Library of Spain and the Ontology Engineering Group from Universidad Politécnica de Madrid.
Since 2011, the project has reached three milestones:

[EXPLORATION]
The first milestone was the transformation of a subset of the catalogue into RDF (Resource Description Framework), and modelled with the ontologies developed by the IFLA (International Federation of Library Associations). During this phase, we explored different techniques for mapping and extracting entities and relationships out of records in the MARC 21 format. This preliminary results was published as a note in the Cataloguing news magazine from IFLA:

"D. Vila Suero, E. Escolano Rodríguez, Linked Data at the Spanish National Library and the application of IFLA RDFS models, IFLA Scat-News (35) (2011) 5–6."

[CONSOLIDATION]
The second milestone of the project was the generation and publication of a significant part of the catalogue following the Linked Data principles. The main result of this phase was the release of a large and highly interlinked dataset under a Public domain license. The dataset is made available using the SPARQL language through a public endpoint, and through a standard Linked Data front-end, Pubby, that provided access to the data in different formats using content-negotiation. During this phase, we improved and extended the methods and techniques for mapping and extraction and built a tool, Marimba, that leverages the knowledge of cataloguing experts during the mapping process. The results of this phase have been published in two journal articles:

"D. Vila-Suero, B. Villazón-Terrazas, A. Gómez-Pérez, datos. bne. es: A library linked dataset, Semantic Web Journal 4 (3) (2013) 307–313."

"D. Vila-Suero, A. Gómez-Pérez, datos. bne. es and marimba: an insight into library linked data, Library Hi Tech 31 (4) (2013) 575–601."

[INNOVATION]
The current version of the data service presents several major improvements and additions. The complete catalogue has been transformed and interlinked covering 4,784,303 bibliographic records, 3,083,671 authority records, and generating 143,153,218 unique RDF triples. Moreover, the number of owl:sameAs links to external datasets has been significantly increased up to three times to a total of 1,395,108 links. Additionally, 108,834 links to digitized materials were added.

The data is modelled using an integrated ontology, based on more than ten different bibliographic ontologies. This rich ontology is published following the Linked Open Data principles and publicly available: http://datos.bne.es/def/

A end-user service has been developed to give access to the vast amounts of interconnected entities. This user interface is built exclusively using the Linked Data knowledge graph and leverages the data connectivity and the underlying ontology to index, present, and arrange information.

INNOVATION

datos.bne.es started as a research collaboration between BNE and OEG and has since then transitioned into an innovative and key product for the National Library of Spain. The service has been identified both inside and outside the library as the "future online catalogue". Although in constant development the goal since early stages of the project has been to create a final product instead of a research product without leaving aside cutting-edge methods and technologies.

Data integration
Marimba is the tool for generating Linked Open Data from several sources (authority catalogue, bibliographic catalogue, digitized materials metadata, DBpedia, and VIAF). Marimba scales up to millions of records and works with several formats including MARC21, RDF, and OAI-PMH. The most innovative aspect of the tool is that it facilitates the ontology engineering process by allowing domain experts to directly map source metadata into a selection of vocabularies. The mapping is performed and validated through easy-to-use mapping templates directly generated from the structure of the data sources. Moreover, Marimba enables connecting the data sources to other sources such as VIAF, DBpedia and a wide network of National Libraries such as data.bnf.fr or the German National Library of Spain. In particular, besides from including links to several datasets, RDF resources in datos.bne.es get enriched with information coming from DBpedia and will include data from other sources in the next versions of the service. Finally, the technology has been recently licensed to the National Library of Spain ensuring the sustainability of datos.bne.es.

Data curation
The data service is changing the ways the library collections are modelled, exposed and consumed in the Web, which is leading to improvements in the data curation of the collection itself. As a by-product of the project thousand of records have been fixed, enhanced, and/or transformed increasing the quality of the catalogue.

Semantically modelling and exposing the full catalogue on the Web is impacting cataloguing practices and the internal structure of the catalogue to answer questions such as:

Q: Could we add links to other datasets directly in the cataloguing process?
A: Yes, addition of links to VIAF and DBpedia are now a current practice.

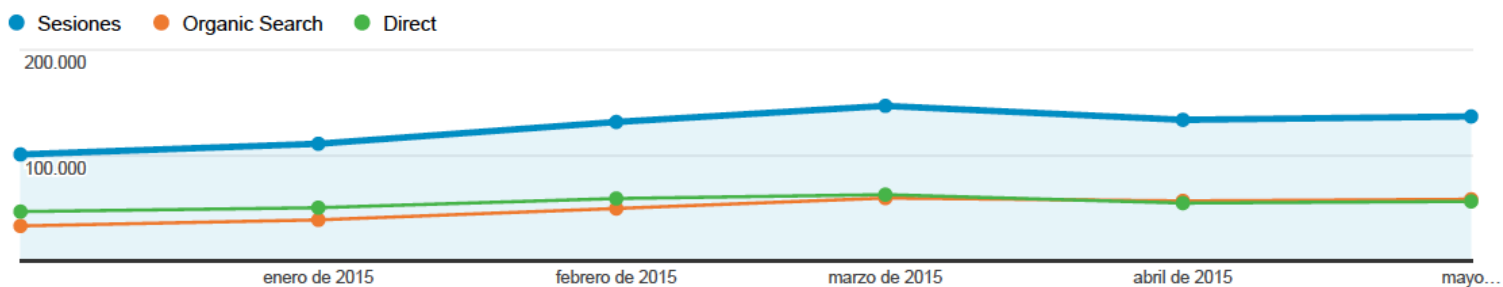*Q: Could we improve precision and recall of search results by adding titles in different languages?*
A: Yes, titles in different languages are now being added to work records.
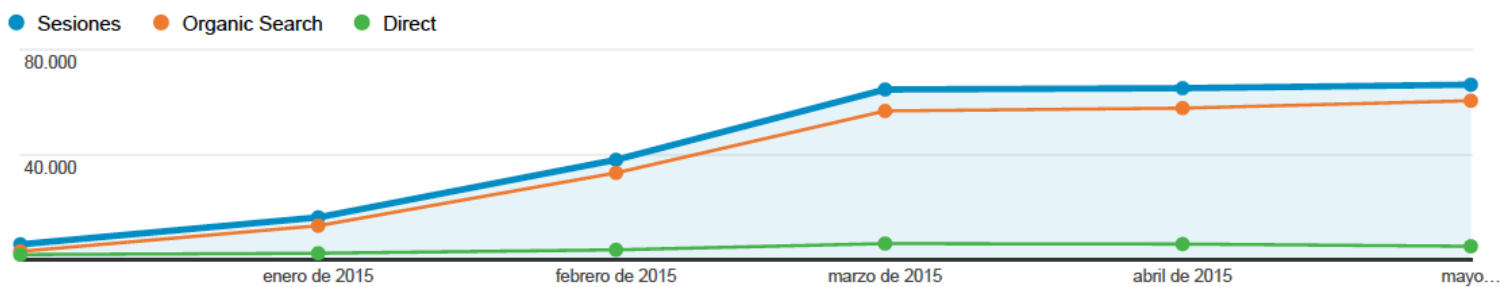
User experience
A major goal of the service since the latest release is to focus on exploiting LOD to create better experiences for end-users. This is a key aspect that differentiates datos.bne.es from other LOD projects. datos.bne.es provides a cross-browser, cross-device user interface to search, browse and discover the library collections. A modern and scalable web architecture enables fast entity-search and minimalistic faceting.

NETWORK EFFECTS

**SEO: Users coming to the library collection
directly from search engines have doubled.**

- Sesiones
- Organic Search
- Direct

200.000

100.000

enero de 2015   febrero de 2015   marzo de 2015   abril de 2015   mayo...

Session in catalogo.bne.es (traditional online catalogue) between January and May 2015

- Sesiones
- Organic Search
- Direct

80.000

40.000

enero de 2015   febrero de 2015   marzo de 2015   abril de 2015   mayo...

Session in datos.bne.es between January and May 2015

Social media

Biblioteca Nacional @BNE_biblioteca · 13 abr.
Federico García Lorca en datos.bne: datos.bne.es/autor/XX841898…
#LODenBNE #MdTFinal

♺ 69   ★ 57   • • •

Interacting with national TV historic series Ministerio del Tiempo (Time Ministry)

Directora @BNE_directora · 26 may.
la obra de Vicente Aranda permanece en la @**BNE**_biblioteca . Una gran aportación al cine español.   datos.bne.es/autor/XX105665…

↩   ♺ 3   ★ 3   • • •

Interacting with the news

Biblioteca Nacional @BNE_biblioteca · 30 mar.
La novela picaresca en datos.bne #MdT6 #LazarillodeTormes
datos.bne.es/tema/XX530631….

↩   ♺ 54   ★ 51   • • •

Interacting with national TV historic series Ministerio del Tiempo (Time Ministry)

EL IMPARCIAL @elimparciales · 10 mar.
#TalDíaComoHoy Nació hace 255 años Fernández de Moratín. Aquí podéis leer 'El sí de las niñas' datos.bne.es/autor/XX894897…

↩   ♺ 1   ★ 1   • • •

Promoting information during historical dates

OPEN STANDARDS AND BEST PRACTICES

Content negotiation: serving Turtle, RDF/XML, JSON-LD, and HTML

JSON-LD objects from storage (MongoDB) to search engine (Elasticsearch) to client (D3.js)

MARC 21, OAI, RDF
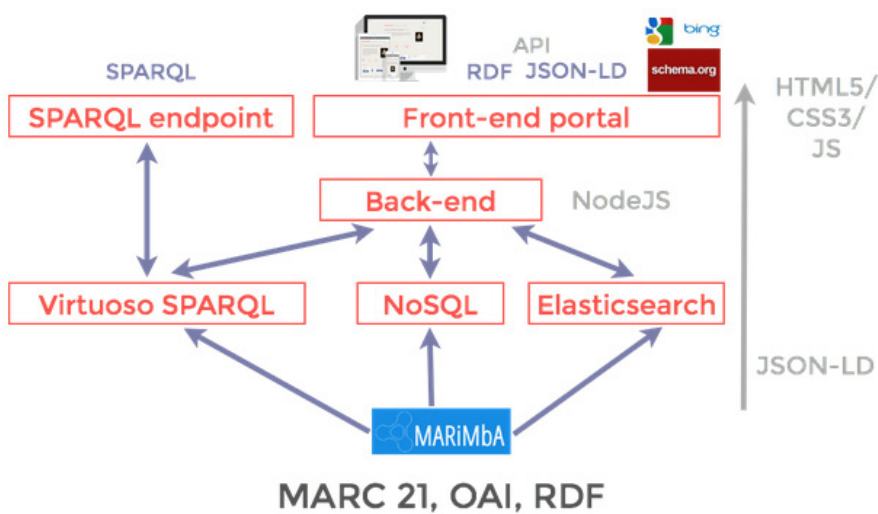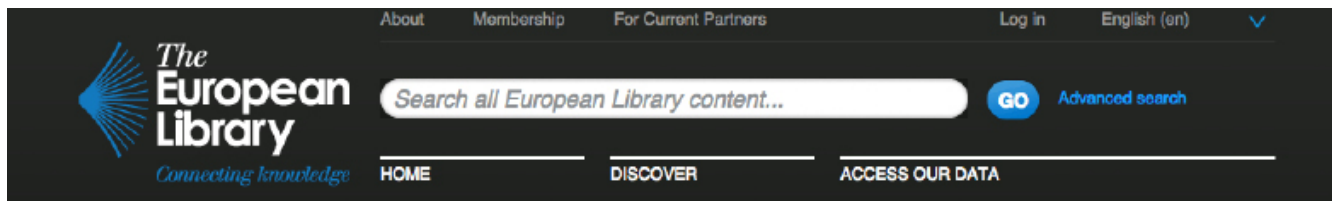
schema.org annotations
using JSON-LD

```
<script type=»application/ld+json»>
{    «@context»:»http://schema.org»,
     «@type»:»Person»,
     «name»:»Cervantes Saavedra, Miguel de
(1547-1616)»,
«image»:»http://commons.wikimedia.org/
wiki/Special:FilePath/Cervates_jauregui.
```

BNE Ontology:

Reuses standards (FRBR, ISBD)
Well documented
published as LD
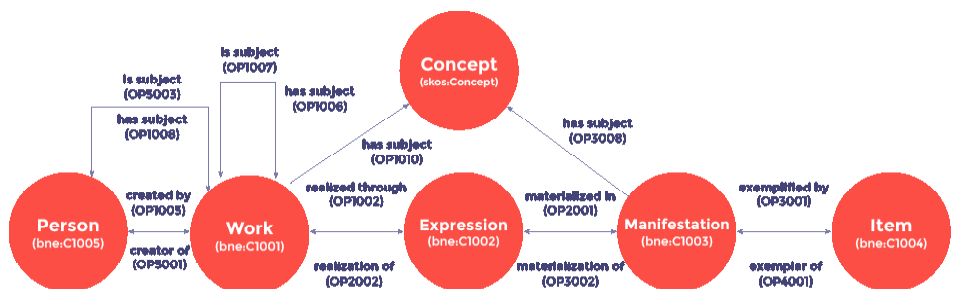
USER EXPERIENCE

CASE STUDY: NATIONAL LIBRARY OF SPAIN TRANSFORMS THE USER EXPERIENCE OF LINKED OPEN DATA

In the lead-up to the RLUK Linked Open Data (LOD) Hackathon on Wednesday 14th May in London, The European Library is running an awareness campaign to promote LOD innovations in libraries across Europe. This case study is one of a series that highlights achievements of pioneering libraries to date in this exciting area.

The launch of the datos.bne dataset in January 2012 marked a crucial milestone in the LOD journey of the National Library of Spain (BNE). As the Library continues to refine the data and deliver innovative user-facing solutions, it envisages a range of benefits, including increased public access, wider reuse and improvements in quality.

Work on LOD started in early 2011 as a collaboration between the National Library of Spain and the Ontology Engineering Group of the Universidad Politécnica de Madrid. The Ontology Engineering Group undertook the technical development of datos.bne, while Library staff provided support on areas such as data analysis and mapping. We set about exploring how LOD could connect library catalogues across the country. Spain has several languages, though, and that makes the records highly heterogeneous. says Daniel Vila-Suero, Researcher and Technical Lead of datos.bne, Ontology Engineering Group.

http://www.theeuropeanlibrary.org/tel4/newsitem/5800
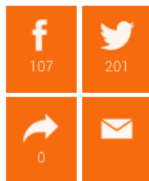
La Biblioteca Nacional estrena un catálogo online más intuitivo

datos.bne.es featured at Spanish Radio and TV Corporation, hub for news in Spanish
http://www.rtve.es/noticias/20150310/biblioteca-nacional-estrena-catalogo-online-mas-intuitivo/

**Ranking and arrranging information using LD
and the BNE ontology
Enriching data with DBpedia**

Powerful search engine
Providing rich entity-search results

DATOS·BNE·es *beta*   INICIO   PERSONAS   ENTIDADES   OBRAS   TEMAS   Ayuda

BNE

Buscar una persona, grupo, entidad, el título de una obra o un tema. Ej... Leonardo da Vinci, ONU, Grapes of Wrath, Aztecas

Búsqueda: *pedro almodovar* (97 resultados)

Filtrar por tipo

36 Obras   5 Personas   2 Entidades

107 Ediciones

21 de 97

Almodóvar, Pedro (1949-)

Pedro Almodóvar Caballero es un director de cine, guionista y productor español, el que mayor reconocimiento ha logrado fuera de España en las últimas décadas

Año de nacimiento: 1949
Año de muerte:
Obras catalogadas: 12

Obras destacadas

Patty, Diphusa y otros textos
Hable con ella
Todo sobre mi madre
"Pepi, Luci, Bom y otras chicas del montón"

Sugerencias

Abril, Victoria (1959-)   Bosé, Miguel (1956-)

Mistral, Jorge (1920-1972)   Portabella, Pere (1929-)

Poetika Pedro Almodovara
Markus, Sara
(2 versiones)
Obra

Escuela de Arte de Ciudad Real "Pedro Almodóvar"
Entidad

Pedro Almodovar
Acevedo-Muñoz, Ernesto R.
(1 versión)
Obra

Todo sobre mi madre
Almodóvar, Pedro
(2 versiones)
Obra

Spando Pedro Almodóvar
Forgione, Anna Pasqualina
(1 versión)
Obra

Almodóvar, Pedro_Jiménez de Góngora, Duque de (1727-1794)
(5 obras)
Persona

El anuncio
Almodóvar, Pedro
(1 versión)
Obra

Patty Diphusa y otros textos
Almodóvar, Pedro
(5 versiones)
Obra

Pedro Almodóvar
Méjean, Jean-Max
(1 versión)
Obra

Hable con ella
Almodóvar, Pedro
(2 versiones)
Obra

Todo sobre mi madre, Pedro Almodóvar
Colmenero Salgado, Silvia
(1 versión)
Obra

Pedro Almodóvar y el "kitsch" español
Polimeni, Carlos
(1 versión)
Obra

Década epistolar sobre el estado de las artes...
Almodóvar, Pedro_Jiménez de Góngora...
(1 versión)
Obra   Disponible en digital

Laberinto de pasiones
Almodóvar, Pedro
(1 versión)
Obra

Almodovar, Pedro
(1 obra)
Persona

"Pepi, Luci, Bom y otras chicas del montón"
Almodóvar, Pedro
(1 versión)
Obra

CROSS-DOMAIN POTENTIAL

Applications to education and tourism
Cross-institutional collaboration
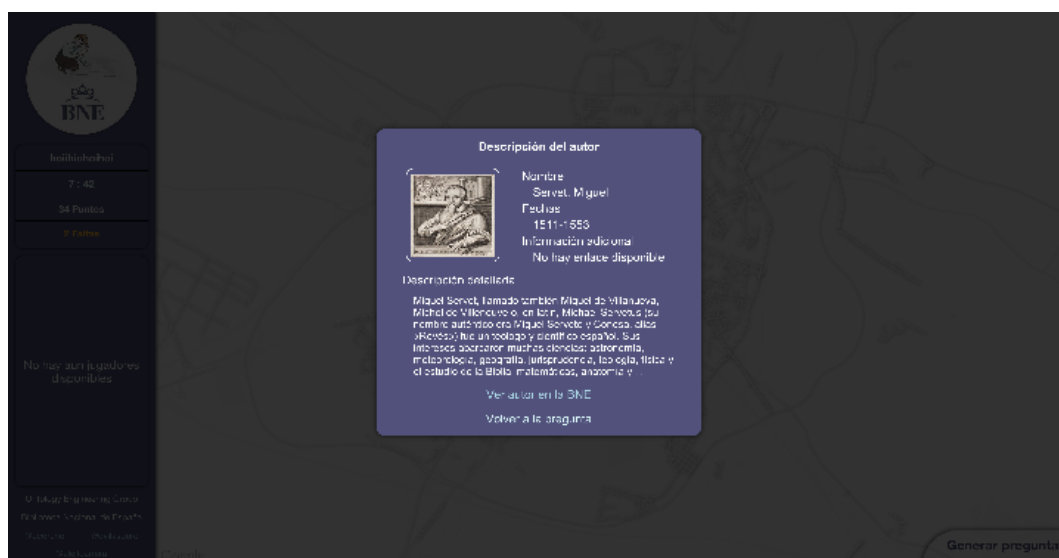
http://www.fujitsu.com/es/linked-open-data-2015/

Winner of Fujitsu Laboratories of Europe Innovation Award
"Citizen Science and Games with a Purpose for LOD data curation"
Daniel Vila, Alejandro Fernández Carrera, Oscar Corcho
Universidad Politécnica de Madrid

Screenshot of the app (1)

Screenshot of the app (2)