

Datos.bne.es: a LOD service and a FRBR-modelled access into the library collections

Ricardo Santos

National Library of Spain
ricardo.santos@bne.es

Ana Manchado

National Library of Spain

Daniel Vila-Suero

Ontology Engineering Group – Technical University of Madrid
dvila@fi.upm.es



Copyright © 2015 by Ricardo Santos, Ana Manchado and Daniel Vila This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License:
<http://creativecommons.org/licenses/by/3.0/>

Abstract:

Datos.bne.es was first launched in 2011 as the data service from the National Library of Spain. In November 2014 a new beta version was built broadening the scope and data wealth included and building, on top of the entity-based data model, an innovative search and display tool targeted to the common user.

FRBR has been the model of choice for transforming, sorting, arrange and finally display the data extracted from the bibliographic and authority MARC21 records. This paper explains the goals and fundamentals behind this ongoing project, the workflows and the decisions made in the conceptualization, the data services developed and the underlying technology supporting all these processes.

Finally, taking into account the feedback received from users, some assessment are made about benefits and downsides of this FRBR-inspired visualization and LOD exposure experiment, drawing some conclusions and possible steps for the future.

Keywords: Linked Open Data, FRBR, FRBRization, User experience

Introduction

Datos.bne.es 2.0 was released in November 2014, as a beta version and experimental test-bed for proof of concept and feedback gathering purposes. This flagship project from the National Library of Spain with the support and development from the Ontology Engineering Group (from the Technical University of Madrid) is a new step into the Linked Open Data domain, which already commenced in 2011, with the first installment of datos.bne.es, upon which experience this new version is built. Two are the main features to highlight:

- The use of Linked Open Data to build a huge set of data, described according best practices of LOD publication, transforming library data into models, structures and vocabularies appropriate for the Semantic Web environment, making it more interoperable, reusable and more visible to the Web, and effectively connecting and exchanging our data with other sources.
- Use, as a reference model, of FRBR basic conceptions and entities to arrange and express bibliographic data, and exploring the benefits and the downsides of this choice to common and scientific users, compared with the traditional wisdom.

Background

BNE and UPM (Universidad Politécnica de Madrid - Technical University of Madrid) first contacted in September 2010 to study the possibility of working together in the field of exposure of cultural and library data as Linked Data.

OEG (Ontology Engineering Group¹), is a research group within the university, pioneering in Spain in Linked Data related projects. OEG is widely recognized in Europe in the areas of Ontology Engineering, Semantic Infrastructure, Linked Data, and Data Integration. This group has developed applications in a wide range of domains and, specifically, in Linked Open Data publication (in fields like tourism or geographical data); in addition, OEG maintains the Spanish chapter of DBpedia, and participates in many standardization activities at the W3C such as the Linked Data Platform specification or PROV-O provenance ontology. Cultural heritage was a very much desired field of research for this group.

BNE's goal was trying to publish data according to both IFLA conceptual model and vocabularies and semantic web principles. To begin with the iterative and incremental life cycle process of the project, a first set of data was transformed, as a proof of concept (Manchado, 2011). That set was composed around Cervantes, affecting all the bibliographic records and all the authorities related to them (around 8.000 bibliographic records and 7.300 authority records).

With the experience gained, by the end of that year the first set of bibliographic records (2.4 million, comprising modern and old monographs and printed and recorded music) and authority data (persons, corporate body and Works) were transformed and published (Vila Suero, 2013), summing up to around 58 million of triplets.

Data was accessible through files datadumps and through a Sparql end-point under a Public Domain license CC0.

¹ <http://www.oeg-upm.net/>

Datos.bne.es 2.0: main features

In November 2014 the new portal datos.bne.es was launched. The new version contains many new features and improvements over the previous version.

It has now been equipped with a visual interface with a search engine, designed to allow users to experiment with new ways of navigating and discovering data.

The range and number of bibliographic and authority records has been increased and upgraded, covering now nearly all the library's materials, adding manuscripts, video recordings, photographs, drawings and maps. Only serials have been left out in this new version. Information about items and holdings is not either available.

The subject headings file has been transformed into SKOS and integrated into datos.bne.es. More than 350.000 subject headings have been transformed and around 30.000 have been linked to LCSH.

Links have also been included to digitalized works in the Hispanic Digital Library, providing access to nearly 100.000 library's digital holdings.

For this iteration, a BNE proper ontology was designed², in order not to incur in inconsistencies with pre-existing vocabularies. However, a work of alignment is still to be done.

The main principles guiding the work done in this updating, and in the future directions can be summed up in these three guidelines:

- Offer the user an integrating platform of all the resources available at the BNE, and a gateway to selected resources from the outside, in a new and innovative navigation experience. The proposed architecture must meet the needs of different kind of users, from academic to common user.
- Offer a solid and scalable underlying data model, upon which new data can be aggregated, either imported from external sources or added by cataloguers to enrich or complete the already existent.
- Offer all this wealth of data in an open access environment, with an open license to make them reusable by other parties.

Front-end and navigation

Datos.bne.es is still a beta version, and search and information display are now under revision, so maybe some of the features explained hereinafter might be changed at the time of reading this paper.

² <http://datos.bne.es/def/ontology.html>

Datos.bne.es is presented in a multi-device, highly responsive front-end, and can thus be used by computers, tablets, phones and similar devices.

The front-end are built around the three main entities underlying the data model, that is, Authors, Works and Subjects. From these entities, only the preferred name label is indexed from each one. Also proper titles from manifestations are indexed for retrieval.

Queries are conducted through the general search box. Suggestions are offered at the time of writing, if the typing matches some of the aforementioned labels indexed. For this to occur, also the order of element must match (signs and capitals are ignored), that is to say, “James, He” would return “James, Henry”, but “Henry Ja” would not.



Automatic suggestion

Along with the suggestion is shown the type of entity to which the result belongs. This can be:

- *Authors* (Autor), encompassing personal and corporate authors.
- *Work* (Obra), in the FRBR sense.
- *Subject* (Tema).
- *Resource* (i.e., Manifestations), this is to say, any given document. This kind of “entity” is not named as such throughout the portal; instead is given the type of media to which the resource belongs:

doñana en la cultura contemporánea	
Doñana en la cultura contemporánea	Registro sonoro musical
Doñana, Parque Nacional	Video
Doñana nombre a nombre	Obra
Doñana, diversidad y ciencia	Libro
Doñana, las otras huellas	Música impresa
Doñana parque nacional :	Libro
Doñana :	Video
Doñana	Música manuscrita

Example of a search suggestion (altered for demonstrative proposals) showing type of document and a Work Entity [translated top-bottom: sound recording, video, Work, book, printed music, book, video, music manuscript]

The suggestions can be skipped clicking the Intro key, prompting then a list of results containing the words searched, divided by the three entities and, in a second line, the documents grouped by type of media.

Buscar un autor, obra o tema

Búsqueda: zaragoza (10771 resultados)

Filtrar por tipo

3861 Obras	1632 Autores	1304 Temas			
2642 Libros	414 Dibujos	224 Materiales cartográficos	142 Manuscritos	124 Registros sonoros musicales	116 Fotografías

Example of a direct search. First line shows the results from the three entities [from left to right, Works, Authors and Subjects]. Second line shows the results among the documents, arranged by type of material [in the screenshot, from left to right:, books, drawings, maps, manuscripts, sound recordings and photographs]

Clicking on any of the boxes eventually applies a filter, narrowing the search to the respective entity. The search engine uses the FRBR graph to retrieve and rank entities, according to the respective entities' relationships.

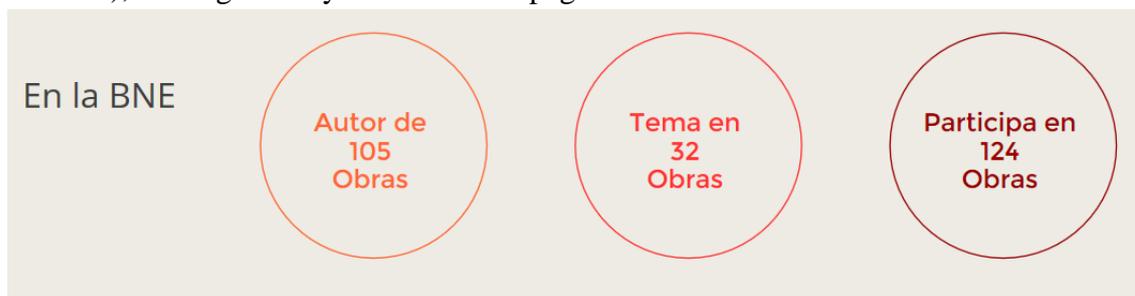


Filter by Works.

An entity-driven search and web architecture allows the user to dive more precisely into the subset of data of interest. That forces the user to a preliminary task of selecting the desired entity, and adds a new click before exploring real documents, but we think the task is worthwhile.

The Author page is composed of contextual information extracted from Dbpedia (abstract and thumbnail), whenever available, and the resources connected to the authors, split into three categories:

- Works by and about the author: these nodes may lead to a Work page or directly to a Resource page.
- Other resources connected to the author (performing a secondary function other than author), leading directly to a Resource page.



From left to right: Author of ; Subject in ; Contributes to

Below the resources from inside the library, outside equivalent links to selected datasets are displayed:

Enlaces	
Ver más	http://dialnet.unirioja.es/servlet/autorExterno/BNE/871548
VIAF	http://viaf.org/viaf/98038975
Sudoc	http://www.idref.fr/02666304X
Biblioteca Nacional Francesa	http://data.bnf.fr/11886882
ISNI	http://isni-url.oclc.nl/isni/0000000121447774
DBpedia	http://dbpedia.org/resource/Julián_Marías
Biblioteca del Congreso de los EEUU	http://id.loc.gov/authorities/names/n79093050
Biblioteca Nacional Alemana	http://d-nb.info/gnd/118781898

These links are sameAs relationships with other LOD datasets (VIAF, LC, BnF, Sudoc, DNB, Libris, ISNI, Dbpedia), and seeAlso relationships to others non-LOD sources, mostly extracted from MARC21 authority records (fields 856 and 670\$u)

Finally are displayed sources consulted for authority work, extracted from MARC21 authority records.

The Work page also is one of the cornerstones of datos.bne.es; it is pivotal in gathering and sorting data about the various editions of the same intellectual creation. It also provides the foundation for eventually connecting these important building blocks of the bibliographic universe to other related entities.

The Work page groups the existent editions in *versions*, that is to say, the languages in which the Work is presented; also shows Works **about** the Work (descriptive relationship).

Tomás de Aquino, / Summa theologica

Summa theologica

En la BNE

Disponibles
7
Versiones

Tema en
20
Obras

Available versions (7) and works about (20) St. Thomas Aquinas' Summa theologica

For the sake of saving navigation levels the most populated version is shown by default. Manifestations are shown abbreviated; full data is shown clicking in *Ver más* (see more). Editions related to digital content are shown first.

Finally, the edition (i.e., manifestation) page contains the attributes usually found in bibliographic records:

Darwin, Charles / On the origin of species / Español / El origen de las especies

El origen de las especies

[Acceder a la versión digitalizada](#) [Acceder a la entrada en el catálogo](#)

Título	El origen de las especies;Charles Darwin ; traducción de Aníbal Froufe ; prólogo de Faustino Cerdón;
Lugar de publicación	Madrid
Editorial	Edaf
Fecha de publicación	2010
Tipo de recurso	Libro
Tema	Evolución
Participante	Froufe, Aníbal
Pertenece a versión	Español
Descripción física o extensión	695 p.
Forma de contenido	[Texto impreso]
Dimensiones	21 cm
Depósito Legal	M 5761-2010
ISBN	978-84-414-2501-9

From this page the user can jump to the traditional OPAC to find item information and circulation services. Also, there is a link to the digital library if the edition has digital content related (usually, a digitized version).

Finally, the Subject page contains links to topics related in the thesaurus, the equivalent term in LSCH and links to the Works:

Evolución

Término aceptado	Evolución
Usado por	Evolución (Biología) Origen de las especies Evolución biológica Especies (Biología)--Evolución
Término específico	Cerebro--Evolución Evolución humana Plantas--Evolución
Término relacionado	Emergencia (Filosofía) Filogenia Selección natural Adaptación (Biología) Genética evolutiva Creación Evolucionismo Biogénesis
Nota de alcance	Úsase para obras sobre la descripción y explicación de los hechos que ilustran las teorías de la evolución. véaseEvolucionismo
Término equivalente en otros vocabularios	http://id.loc.gov/authorities/subjects/sh90004042

Subject page

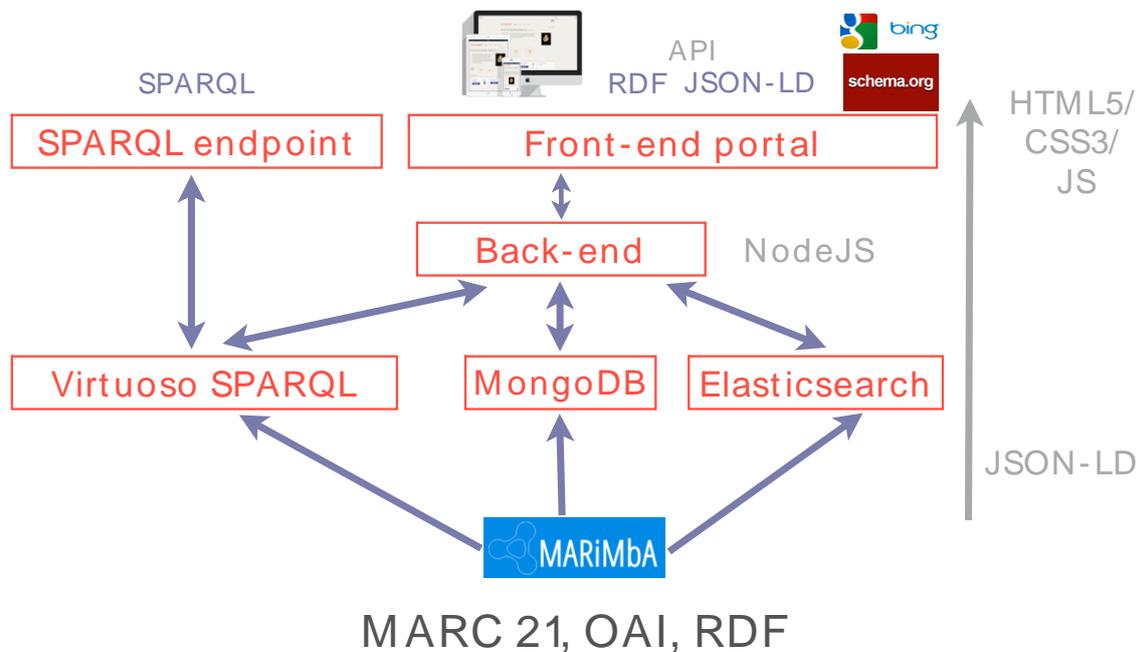
Possible enhancements for the future in visualization and navigation:

- Include more values to be searchable.
- Allows search only into one kind of entity.
- Foresee and manage searches involving two entities (for instance, Shakespeare Hamlet)
 - Improve information display and sorting. Explain better the difference between works and resources.
 - Explain or refine the concept of Authors (not all Persons are authors, not all authors are persons).

Technological background

A wide range of modern web technologies composes the technological platform underlying the datos.bne.es service. In order to process, analyze, map and transform catalogue data into Linked Data, a scalable tool named Marimba has been developed. Marimba allows library experts to participate in the mapping and generation of Linked Data, hiding the complexities of the semantic models and algorithms driving the generation process (Vila-Suero, 2013). Once the data is transformed and interlinked, a publication pipeline has been established. First, Linked Data is indexed into a search engine especially tailored to deal with and to benefit from the semantic structure of the data. The graph nature of the data allows for a straightforward ranking mechanism directly based on the underlying FRBR model. Second, the data is also stored into two different systems: a NoSQL database to support the service

front-end, and a triple-store to facilitate querying and reuse by semantic applications. At this point, it is important to mention that the JSON-LD linked data serialization is used throughout the whole platform. This serialization brings a clear benefit as no data transformation is made after the generation process, and thus the search engine, the data repositories and ultimately the front-end consistently deal with the same data objects. Also, as the service is fully built in Javascript, both server and client codebases, JSON is a natural choice. This is an overview of the technical architecture:



(Linked) Data services

The platform described above has been designed to provide a solid, interoperable and scalable data-driven service. Based on these principles, a suite of data services is being developed. In particular the system offers:

- A public SPARQL endpoint is provided to make the data query-able using the SPARQL language. (<http://datos.bne.es/sparql>)
- A search API is provided to web developers for developing applications on top of the BNE data. This search API is backed by the previously mentioned ranking mechanism and can serve as a reconciliation service as well, to be used with tools like Open Refine. For instance, a search for the keyword “Cervantes”³ would provide as a first result the URI of Miguel de Cervantes Saavedra.
- A content-negotiation mechanism for every resource with the following available formats: HTML, RDF/XML, Turtle and JSON-LD.
- Schema.org annotations in the JSON-LD format for each web resource in order to increase visibility and semantic indexing by search engines of datos.bne.es resources. For example, authors are annotated as schema:Person and works are annotated as schema:CreativeWork.

³ <http://datos.bne.es/find?s=cervantes?format=json>

Besides the services listed above, the plan is to extend the APIs and its documentation in order to promote reuse by third parties, as well as organizing developer gatherings in the context of Open Data events in Spain.

Extracting and linking FRBR entities from MARC21 structures

Designing a strategy for transforming MARC21 data into FRBR involves a thoughtful look at the legacy data and codification procedures from every institution (Boulet, 2013). In datos.bne.es mapping, MARC21 fields and subfields were considered, and creation, delimitation and relationship were extracted based in the presence, or absence, of the values of the coding.

From 1991 to 2007 the ad-hoc cataloguing software in use at the BNE forced cataloguers to explicitly assign every bibliographic record to a previously existing Work authority record, or in a translation, to an expression authority record. That wealth of authority records were quite often done at a minimal level, and did not even sometimes properly identify Works, but, summing up, we found ourselves with a remarkable amount of authority records for Works and Expressions to deal with and take advantage of.

FRBR was the logical answer to how to best profit from these data. Firstly these basic assumptions were made as assigning data to entities:

Person/Corporate body authority record → Person / Corporate body entity
Work (Person/CB+title) authority record → Work entity
Expression (Person/CB+title+language) authority record → Expression entity
Bibliographic record → Manifestation entity
Subject authority record → Subject entity

In 2007 new software came to the library, and links from bibliographic records to work/expression authority records were only used in the usual uniform title cases.

When no explicit relationship were recorded (fields, 240, 130), manifestation records were assigned to the corresponding work/expression through the sum of field 1XX + 245 \$a, matching the resultant string to the corresponding authority record. Language of the original expression was inferred from fixed field 008 from bibliographic records, when no explicit language in 130/240 \$l was given. Obviously many codification errors surfaced in these assignments, and many manifestations were work “orphans”, but it was worth the risk, as many links could be traced.

Having established the entities from the existence of self-contained records, mapping of properties of elements for those entities were traced from their respective subfields. For instance:

Bibliographic record / Field-subfield / 260 \$b = bne:P1117

Mapping of properties were established from subfields. Various levels of granularity were thus reached, according to the proper characteristics of subfield: ambiguous subfields (245 \$b, for instance) had to be represented as general properties, and local practices were sometimes of help to extract semantics, and other nor. That took us to the inevitable conclusion that brick-

and-mortar cataloguing practices have to be purposefully reconsidered in order to help Linked Data applications manage the data.

No aggregated statement or blank nodes were used to stablish groups of elements (such as could be *publication area*, for instance).

FRBR as a reference model

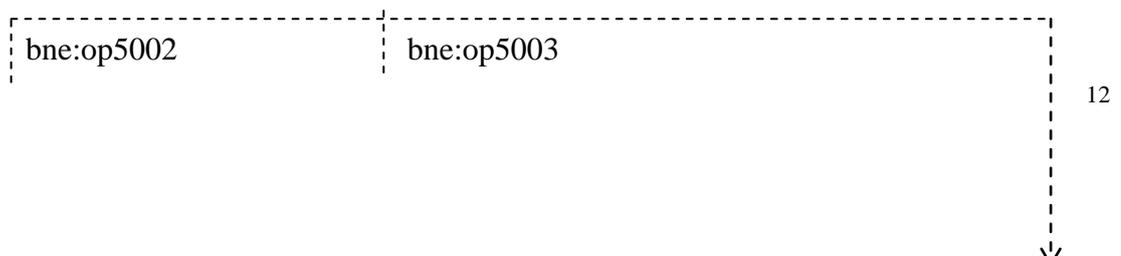
In our view, FRBR is the most suitable model for representing relationships, especially in large quantities of bibliographic data, as there is usually in a national library the size of that of Spain. It allows collocating related resources, guiding the users into the subset of data more precisely suitable to their needs. It also helps in giving the user what he is actually asking for: a user searching for authors should have authors as an answer, for instance. Usual search capabilities in library interfaces provide immediately resources, no matter what the query is about. Data.bne tries to understand what the user is asking to, and acts accordingly.

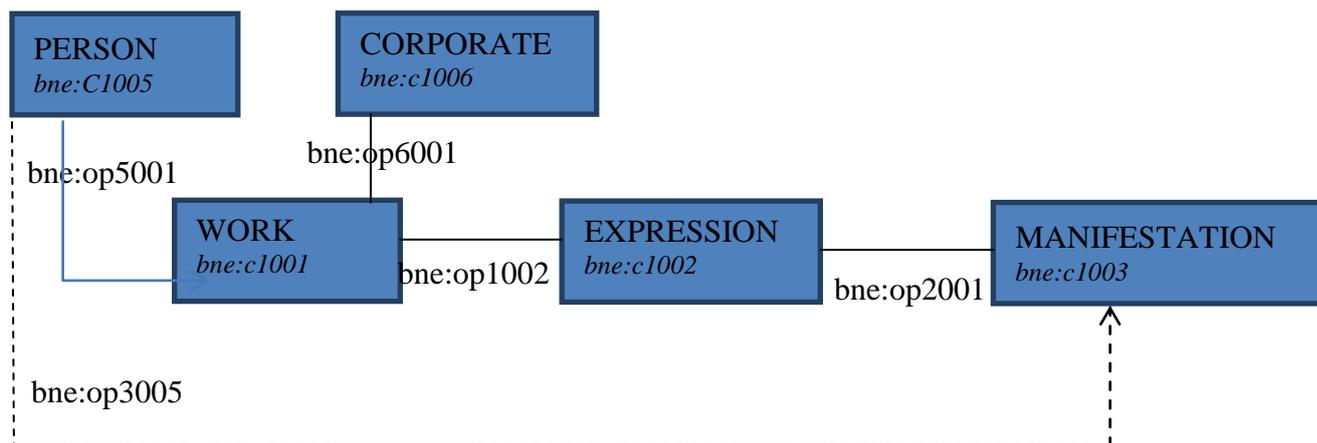
However, embrace FRBR as a whole proved no easy task. MARC legacy data and structure is far older than FRBR model, thus what we can dig or infer from existing data does not always fit properly into FRBR structure. As said before, although some complex mechanism of data analysis were carried upon records, for the purpose of identifying and relating FRBR entities, some data in the BNE catalogue fell through the creaks of this building. Datos.bne.es version 1 skipped data not fitting into the model, so we had a solid model, but a dramatic loss of data:



In this model, data that could not meet the criteria (that is, manifestations without a related work), did not make it into the final pool of data.

For datos 2.0, a more pragmatic vision was adopted, introducing more flexibility and some loopholes on the model, so that we could express a wider range of relationships without losing the spirit and benefits or FRBR features:





Two significant new relationships have been implemented in the model between authors and manifestations:

- OP5002 and OP5003 relate a person or corporate body directly to a manifestation as *creators*; these relationships were created to build links between authors and manifestations when a work couldn't be extracted or inferred; the properties are mostly used to circumvent FRBR logical chain of events when 1:1:1, this is, works with only one expression and one manifestation. For this highly common relationship, no work/expression is considered.
- OP3005 relates a person to a manifestation as a *contributor*. Authors recorded in bibliographic records as additional entries are related with this property as nothing more precise can be ascertained. No relationship between authors and expressions are considered.

More complex relationships between Works or between Agents were not fully considered at this stage. Relationships from manifestation to W/E/M, expressed through additional access point in bibliographic records has been expressed through a more general *has relation* (bne:OP3007), with domain *Manifestation*.

For Subjects we tried to experiment with some data wrangling from MARC21 to FRBR structure. In manifestations, Subjects relationships has been modeled as *has [entity] as subject* relationship. But Subjects has been uplifted and assigned to Works, whenever they exist:

Software libre

Término aceptado	Software libre
Usado por	Software de código abierto
Término genérico	Software
Término relacionado	Propiedad intelectual--Programas de ordenador
Nota de alcance	Úsase para obras sobre aquellos programas de ordenador que permiten al usuario ejecutarlos, copiarlos, distribuirlos, estudiarlos libremente
Término equivalente en otros vocabularios	http://id.loc.gov/authorities/subjects/sh99003437
Esquemas	http://datos.bne.es/scheme/conceptScheme http://datos.bne.es/scheme/topicalTerms

Obras

<p>Implantació de sistemes de programari lliure</p> <p>Albós Raya, Amadeu</p> <p>(2 versiones)</p> <p>Obra</p>	<p>Software libre: técnicamente viable, económicamente sostenible y socialmente justo</p> <p>Mas, Jordi</p> <p>(2 versiones)</p> <p>Obra</p>	<p>E-business colaborativo</p> <p>García Valcárcel, Ignacio</p> <p>(1 versión)</p> <p>Obra</p>
---	---	---

Subject page for Free software; below connected Works

Possible enhancements for the future in FRBRization, data elements or workflows:

- Consider inferring and creating works through some data mining, maybe giving them later back to catalogue, as has been proposed elsewhere (Ilien, 2015).
- Better representation of related and complex works, aggregations and whole/part relationships.
 - Use of medium of performance for constructing musical expressions.
 - Reconsider and improve Subjects relationships not expressed as related concepts in the thesaurus, but as aspects (for instance, relating *Libraries* to *Libraries-History*, or *Libraries –Hungary..*), or between entities (relating *Poe, Edgar Allan* to *Poe, Edgar Allan-Musical adaptations*)

Reception and feedback from users:

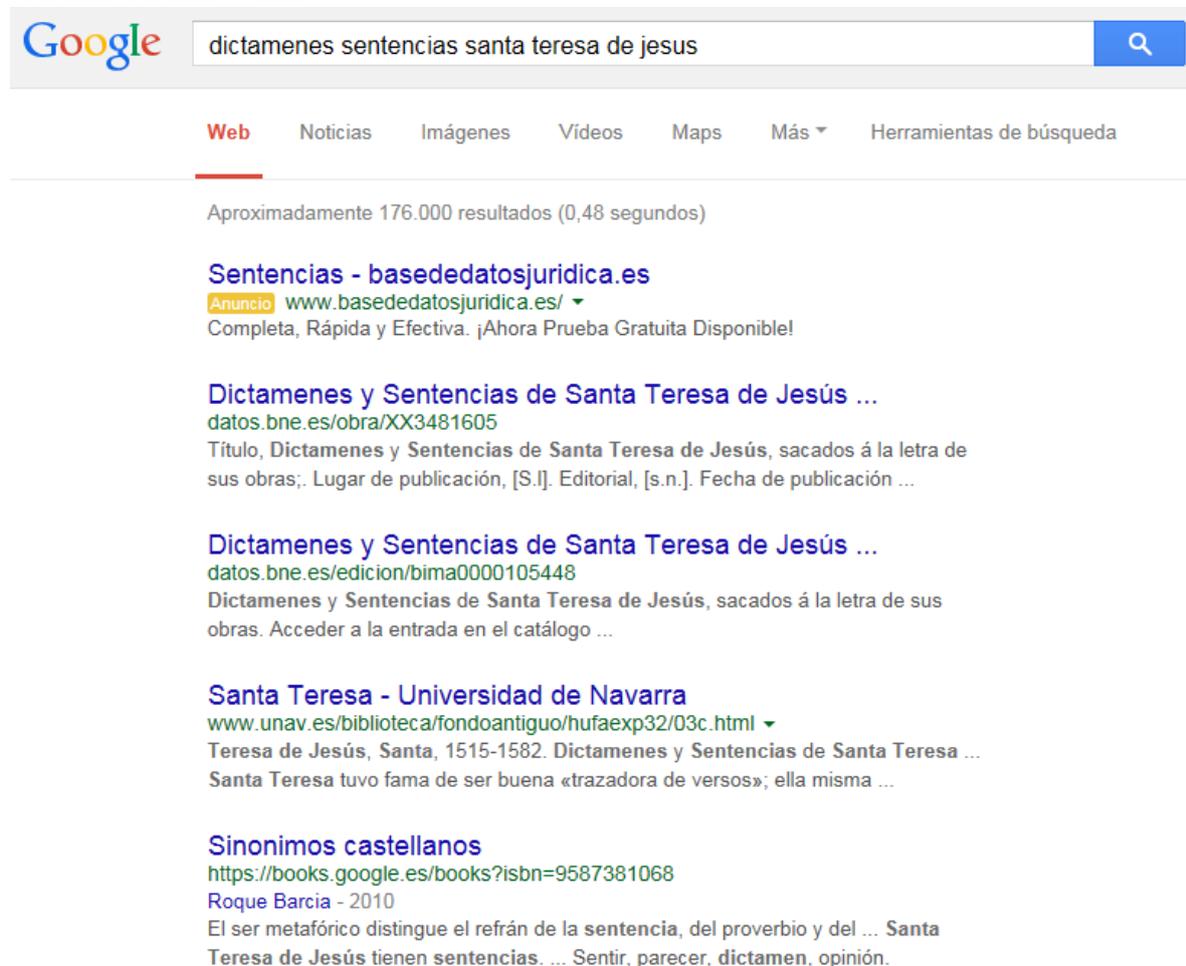
Whereas in datos.bne version 1.0 the target audience were primarily computer and specialized users, focus in 2.0 have been broadened to include academic researchers and common users. We were very much intrigued about how people used to OPAC structure and information display would search and navigate through entities, and whether they would be able to select preferred entities and go down from there to resources. Particularly, we were most interested in the reception, understanding and use of the abstract “Work” concept.

Additionally, we wanted to measure the impact of Web Semantic and Linked Open Data best practices followed in this project in the outreach beyond library-bound users.

Datos.bne.es lives separately from the library catalogue main entrance (<http://catalogo.bne.es>), so a direct link were traced from this OPAC and that of authority file (<http://catalogo.bne.es/uhb/authoritybrowse.cgi>), enticing users to try an alternative experience of exploring the library resources. Main traffic was channeled through this way.

Of course, expert use of datos.bne data from Web Semantic and Linked Open Data communities, interest groups and specialists carried on in new experiments, projects and tools using cultural data (García, 2014).

It has been remarkable the uptick in traffic coming straight from searching engines, made possible because every entity and resource having a permanent URL. Search positioning was improved dramatically, especially in scarce resources, lesser known authors and precise title searches for rare works:



Screenshot for a Google search results for a rare edition of Saint Teresa of Avila

In the one hundred odd e-mails received so far, there were technical questions about ontology and vocabulary, Sparql access, modelization, but more abundant were requests from non-technical users. These have meant a valuable wealth of information about strengths and

weaknesses, and gave us some hints for future developments. Here is a summary of the most relevant conclusions:

- Exposure and awareness of authors. Authors are at last given the central role they deserve in library data. They have their own page, their own set of data about them, one that can be potentially filled with new data, relationships to other authors, or topics, or whatever. Their contents are well delimited, meaning for authors that their body of work can be easily accessible. So, living writers (or sometimes descendants or pundits) have been eager to contribute personal data (birth date or place, personal websites) or to flag authorship mistakes. This has been also valid for highlighting missing works, or even prompting the donation of items.
- Exposure of resources. As has been said, proper titles of given resources have improved their SEO positioning, taking users searching for particular resources straight into the library services. Thus, one the main task of answering the Datos.bne.es mailbox has been effectively redirecting users' requests to proper services (music, very often, but also rare books, modern monographs and videorecordings). This requests were about how to get a copy of an item, or how to access them, or asking about reproduction rights.
- Overexposure. On the other hand, in the very short existence of datos.bne 2.0, some authors have complained of finding themselves up there "on the cloud". Authors already existing in our traditional catalog for years suddenly have asked their data to be removed from the service, because their data showed up after typing their name in a search engine.
- Misconceptions about the services provided by datos.bne.es. Maybe because of the new display and data arrangement, quite distant from the traditional library environment, some users coming straight from Google or other search engines were confused about where they just had come into, asking for information about how to purchase a copy of a book, or contact data about authors. Also was relatively common the misunderstanding of certain concepts used elsewhere in the *about* section or in online help, like Open Data, Data Reuse, Data Downloads, confusing metadata for content, and therefore asking for how to download a book or a song, or even a movie.
- FRBR-like sorting and display. From our experience, FRBR concepts must be embraced, but FRBR terms must be avoided. The feedback received has taken us to think again how to clarifying the use of the concept and the word Work. This is a concept we deem undoubtedly useful, but must be presented to users in a transparent way. Its benefits are twofold: we save time to users by putting together similar editions, especially in large bibliographical institutions where scholars use to examine and compare items to build the bibliographical history of a Work, and we provide a solid ground to establish relationships to other related Works. For instance, one user was extremely thankful to check quickly Spanish-speaking translators of Saint Exupery's *Le Petit Prince*, for some research about it (he was looking into Expressions without knowing). But tangling with Works and Manifestations proved to be trickier for some users, incapable of telling the difference between one and another, especially in 1-1 relationships, as stated above. We hope to address this pitfalls for understanding FRBR basic relationships in future improvements, either at the display level, at the data level, or both.

Conclusion and more possible developments for the future

Besides all improvements sketched before, some other strategic decisions have to be taken in the future. Having established the underlying architecture supporting datos.bne overall design, more and more lines of research, experimentation and development can be approached in new versions of the service. Here are some:

- More content and more entities, please. Data model and visualization based on entities has proved to be essential in building flexible and robust data architecture. More entities are thus desirable to be designed, and with them, more relationships to be established. Some are inherent to library data, and some don't. Obviously, the first step would be the data missing in datos.bne, (serials, harder to be fitted into this FRBR world, and items, with relevant information, like owners and provenance, to be carefully woven into the rest on entities). But also are of interest Years, Places, Publishers, harder to extract and manipulate.
- Dig deeper. External sources of data are not yet fully exploited. A close examination of available sources (from library and non-library worlds) is needed, picking up pieces of data valuable, designing and establishing gateways for exchanging and mining data that could be used.
- Change minds. Cataloguers need to think in the way data are modelled when they proceed to the daily cataloging; the semantic web begins with the human data creator facilitating linking when recording data. Training and awareness are absolutely necessary tools to make things work properly.

Entity-based architecture and data display is full of benefits for users, expert, academic or common, or for reference librarians. Although datos.bne.es is at this moment at an early stage of development, we feel that we have dared into a no-return path.

References

GARCÍA, A. (2014). *Integración de contenidos semánticos en un portal web de científicos y humanistas valencianos: Vestigium*. Ph. D. Thesis Universidad Católica de Valencia San Vicente Mártir [Spanish]. Retrieved on February 20th in: <http://eprints.rclis.org/24649/>

BOULET, V. (2013). S'appuyer sur la structure des données et les liens - le format MARC comme tremplin pour le Web de données : l'exemple de data.bnf.fr. *79th IFLA World Library and Information Congress*

ILLIEN, G.; PEYRARD, S. (2015) Trust me if you dare: data quality in the age of Linked data. *Europeana Tech Conference* [Power point slides]. Retrieved on 14th April in: <http://www.slideshare.net/Europeana/illien-peyrardeurtech2015?qid=52c1f6fb-70d3-4782-bf62-dac3620d107a&v=qf1&b>

MANCHADO, A.; VILA SUERO, D. (2011) Linked data: un caso de uso en la BNE [Power point slides, in Spanish]. *XII Jornadas Españolas de Documentación*, Málaga, 25-27 mayo de 2011. Retrieved on March 12th in: <http://www.slideshare.net/bne/un-caso-de-uso-en-la-bne>

VILA SUERO, D.; VILLALÓN-TERRAZAS, B.; GÓMEZ-PÉREZ, A. (2013) Datos.bne.es: A library linked dataset. *Semantic Web*, vol. 4, n° 3 (2013), p. 307–313