DIGIBÍS® Newsletter. No. 17. January-June, 2017

DIGICLIC

Information about enriched digitization, software for Libraries, Archives and Museums and international standards.

Semantic reconciliation

Linked Open Data (LOD) enrichment technology that is applied almost automatically

Linked Open Data (LOD) resulting after applying the semantic reconciliation technology. To the right, the MARC 21 record on the same authority record allows us to see the corresponding URI.

SUMARIO

Editorial 3

INTERNATIONAL

Europeana

Update of the Europeana Business Plan 2020 4

GLAM APPLICATIONS

Virtual archives

Norman Foster chooses DIGIARCH to manage his digital archive 5

LOD enrichment

Semantic reconciliation 6

Virtual libraries

Hispana reaches 7 million digital objects 8

Microsites in the Virtual Library of La Rioja 9

Upgrades to DIGIBIB 9.1

Upgrade of the Digital Library of the Royal National Academy of Medicine 10

The Digital Library of Galicia adds SPARQL searches 11

E V E N T S

Presentation

Launching of the Virtual Library of Enlightened Scientific Travels 12

Fairs and congresses

Americanae at the 15th FESABID Congress 14

News

The digitization of Defence collections in the "Success stories" of Red.es 16

DIGIBIS

DIGICLIC
DIGIBÍS® Newsletter

Managing Director
Tachi Hernando de Larramendi

Project Director
Xavier Agenjo Bullón

Director of Finance and Administration
Nuria Ruano Penas

IT Dept. Director
Jesús L. Domínguez Muriel

Art Director
Antonio Otiñano Martínez

Sales Director
Javier Mas García

Technology Coordinator
Francisca Hernández Carrascal

Administration
María Luz Ruiz Rodríguez (coord.)
José María Alcega Barroeta

IT Department
Feli Matarranz de Antonio (coord.)
Alejandra Arri Pacheco
Andrés Felipe Botero Zapata
Julio Diago García
Carlos Henche Hernández
Luis Panadero Guardeño
Rafael Roldán Tejedor
Fernando Román Ortega

Innovation Department
Paulo César Juanes Hernández (coord.)
Noemí Barbero Urbano
María Isabel Campillejo Suárez
Susana Hernández Rubio
Montserrat Martínez Guerra

Digitization Department
Francisco Viso Parra (coord.)
María José Escuté Serrano
Amando Martínez Catalán
Javier Ramos Altamira
Ricardo Vela Alegría

Documentation Department
Isabel Martín Tirado
Beatriz Pascual Fernández
Alicia Varela Villafranca
Luis Fernando Méndez Prado

Sales Department
José María Tijerín Gómez

Coordination with the Ignacio Larramendi Foundation
Patricia Juez García

DIGIBÍS®
C/ Alenza, 4, 5ª planta
28003 Madrid
Tel.: 91 432 08 88
Fax: 91 432 11 13

digibis@digibis.com | www.digibis.com

Legal Deposit:: M-18050-2009
ISSN (print): 1889-4968
ISSN (online): 1989-5194

"Semantic reconciliation" LOD technology

In DIGIBÍS we have always paid great attention to semantic enrichment, as can be seen in different issues of DIGICLIC, from the editorial in issue seven, where we reported on the enrichment of the records of the Virtual Library of Bibliographical Heritage with the List of Subject Headings for Public Libraries in SKOS, to the article on the description of the enriched digitization in issue 16.

There have been many resources that have benefited from that Linked Open Data (LOD) technology and, among them, particularly the Polymath Virtual Library of the Ignacio Larramendi Foundation. Therefore, in order to carry out and present of one of the thematic libraries that make it up, that of the Enlightened Scientific Travels, we have taken on a technology called "semantic reconciliation" to link almost automatically the authority records with existing value vocabularies, such as VIAF, datos.bne.es, DBpedia, Wikipedia and Wikidata, among others.

Almost automatic enrichment of the authority records

In this process all the authority records of the Polymath Virtual Library, not only those of the Virtual Library of Enlightened Scientific Travels, have been linked to different sources and have been enriched with many new data. Semantic reconciliation between the value vocabularies that the Polymath Virtual Library has established has affected both those records that were not linked as well as those that were partially linked.

Authority records are linked to various sources and enriched with new data

As always, the Polymath Virtual Library has been the testing ground for our DIGIBIB application, in this case in its version 10, as well as for semantic reconciliation technologies (see pgs. 8 and 9).

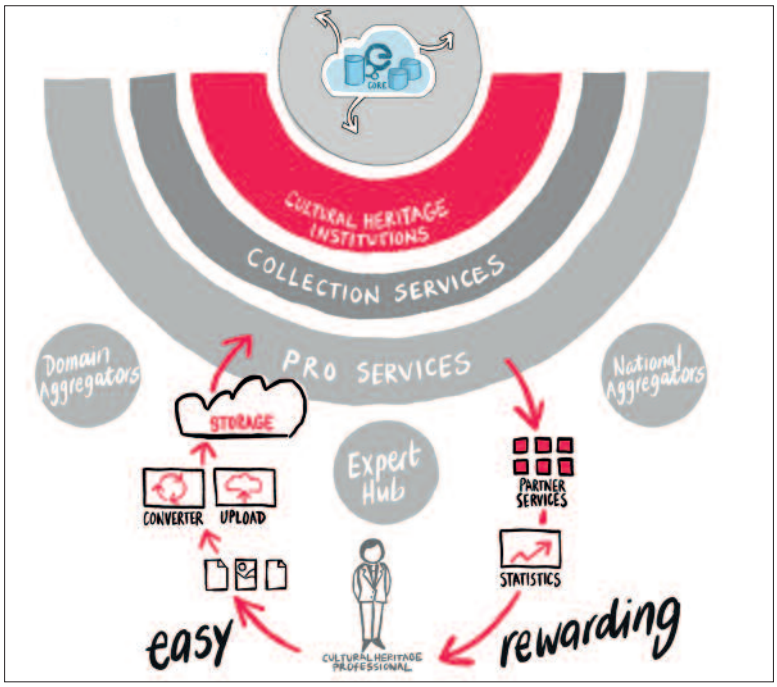
Once tested, this technology is being applied to different Virtual Libraries implemented with DIGIBIB. We anticipate that there will be spectacular results. We have also begun to use Wikidata as one of the Linked Open Data resources suitable for linking RDF descriptions and for the enrichment of the data, but this will be the subject of the next issue of our newsletter. ■

Upgrading of the Europeana Business Plan 2020

Three years after launching the 2015-2020 strategy "We transform the world with culture", Europeana presents an upgrade of its work plan under the title "A call to culture"; this is a review of the work carried out, up to the moment, with the purpose of focusing on the objectives, directing resources to accelerate the progress of the tasks and achieve the objectives on time.

The three priorities that Europeana has establshed are:

1. To facilitate and reward Cultural Heritage Institutions for sharing high-quality content.
2. To get closer to the participants to achieve the market and audience targets.
3. To involve people in Europeana websites and through participation campaigns.



Computer graphics collected from "A call to culture. Europeana 2020 Strategic Update".

It focuses on the quality of the data, standards, enrichment and interoperability

The first objective is the one that most directly affects the cultural institutions. To achieve this, Europeana has proposed:

- Organisational changes: this will mean reviewing and reforming Europeana's aggregation infrastructure, so that cultural institutions control their data, national platforms direct the standards and contribute to the infrastructure, experts set out their experience as well as the management of their knowledge and relationships, and Europeana is the platform that brings everything together.
- Changes in the platform and the product: this leads to tools being developed that directly connect Europeana to commercial data management systems and as a direct upload point for digital collections of cultural heritage institutions using the latest Web technologies.
- Content changes: closing the loop between what users want and the acquisition of the metadata and content of the partner institutions; focus on the quality of the data rather than on the quantity; working with Member States to set data quality targets that encourage and support online accessibility to high quality cultural heritage.
- To clearly establish the actions of each party: the cultural heritage institutions and Europeana Network Association, data quality and control of its use; member states, the establishment of the data and their quality; aggregators, the development of standards and the enrichment and improvement of data, as well as the assistance of experts; and Europeana, publication, data interoperability, frameworks, distribution and usage report . ■

Norman Foster chooses Digiarch to manage his digital archive

The Norman Foster Foundation has chosen DIGIARCH as the system to manage digital archives for the dissemination of its digitized, standardized and described materials, on Internet. In this way, users from all over the world can access the digital copies of drawings, models and projects of the renowned architect Norman Foster.

The contents of the Norman Foster Foundation Archive chronologically runs from the 1950's until today and encompasses a wide variety of materials, such as drawings, plans, sketches, models and personal items.

The digitized materials which will increase with the successive phases of digitizing, including the work of the architect during his student years at the Universities of Manchester and Yale, or utopian proposals, created with partners such as Buckminster Fuller, up to the projects of the studies founded by Norman Foster, among them: Team 4, Foster Associates, Sir Norman Foster and Partners and Foster+Partners.

The Norman Foster Foundation has chosen DIGIARCH software that, thanks to the use of Linked Open Data (LOD), will allow the use of existing vocabularies to enrich its contents, such as the Getty Vocabularies: The Art & Architecture Thesaurus, for subject descriptors; The Getty Thesaurus of Geograpic Names, for geographic descriptors, and The Union List of Artist Names, for the names of architects and artists.. It works in such a way that all of the descriptors used in the cataloguing are standardised and linked through these international vocabularies.

The DIGIARCH application automatically generates them at the same time in a OAI-PMH repository, which implies that the administrator only needs enter the data once in the catalogue. The repository meets the requirements and technology so that the digital objects of the Digital Archive of the Norman Foster Foundation can be harvested and disseminated by Hispana, Europeana or OAIster, thus very sensitively broadening their dissemination in the network.

DIGIARCH features a high-definition viewfinder that can display objects with 360º and 3D views. In the case of the Norman Foster Archive the models have been digitized in order to be able to be viewed on 360º. ■

By the end of the first half of 2017, the user can access 1129 digital objects:

- 1065 drawings
- 37 textual documents
- 20 drawing notebooks
- 5 slides
- 1 model

Semantic reconciliation

Since 2014, virtual libraries that use our software to manage digital records, DIGIARCH, and physical and virtual libraries, DIGIBIB, have the opportunity to enrich their records with Linked Open Data (LOD) links.

This improvement was applied from the bigening to the Polymath Virtual Library of the Ignacio Larramendi Foundation, our test lab, which meant the semantic enrichment of its authority records (personal names, entities, subjects, places...) and, in addition, as reported in the 2014 double issue of our newsletter, it was the only Spanish institution to participate in the study of the OCLC on implementation of Linked Open Data.

Manual *versus* "automatic"

In these three years, LOD sources and their content have grown vastly and quickly. In the last ten years, published LOD data have increased by more than 4000 %.

This work is increasingly arduous: standardisation of data, link with LOD, removal of complementary data... In addition, these tasks do not only affect the incorporation of new records, but the maintenance of those that already exist and need to be updated.

If we add to this the fact that as discussed above, the amount of data that has to be enriched and linked is increasingly becoming greater, it will be difficult for the cultural institutions to be able to carry out this work exclusively by using manual procedures, which require a large investment in time and in human resources.

In this sense, one of the semantic enrichment tasks that can considerably facilitate the work is the so-called "reconciliation of data", which could be defined as the computer process of searching for an equivalent entity in the LOD sources almost automatically, rate the level of correspondence between the values obtained and remove the URI of the matching resource.

Exportation, Transformation and Load (ETL)

To carry out the process we have just described, in addition to the specific developments that can be performed, there are different ETL tools (Exportation, Transformation and Load) and other more comprehensive programmes that allow all the tasks involved in the LOD enrichment process to be partially or fully performed massively

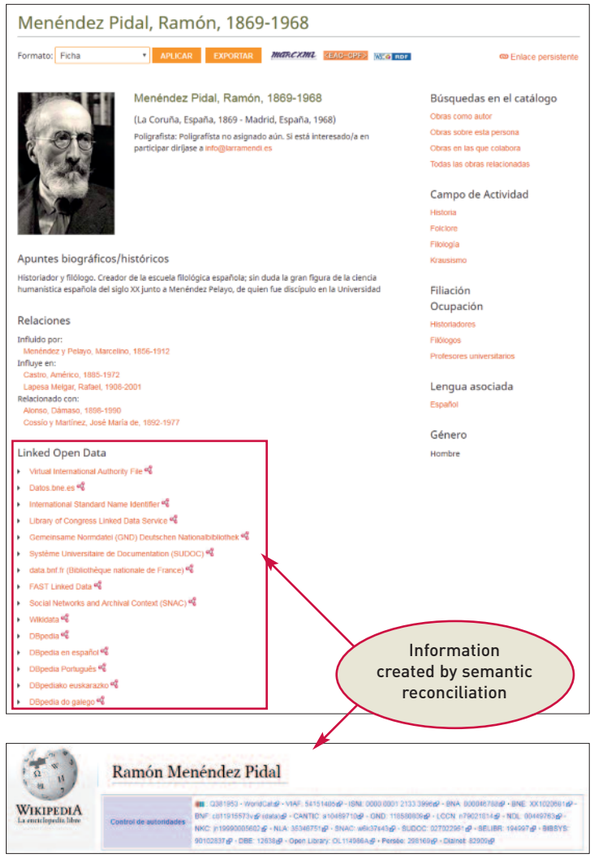
> Semantic enrichment through "data reconciliation" occurs nearly automatically

and "almost automatically"; although it is necessary to clarify that "almost automatically" does not mean that this can be done in an unattended way and without the mediation of specialized professionals, but that the productivity of this technology makes viable what manually would be almost impossible.

Semantic reconciliation in the Polymath Virtual Library

Once again the Polymath Virtual Library is located at the head of the Spanish virtual libraries that apply ETL tools and has been the subject of reconciliation semantics by DIGIBÍS, although it is not the only virtual library that has begun to apply it in our country.

Among the LOD sources used for semantic reconciliation of authority records of polygraphs, in addition to VIAF, Data.BNE and Library of Congress, among others, DBpedia and the DBpedias of cultural areas of origin of the polymaths (Portugal, Brazil...) have been used, as well as the DBpedias in all the official languages of Spain (DBpedia in Spanish, DBpedia do galego and DBpediako euskarazko kapitulua; the DBpedia in Catalan, unfortunately, has not been able to be located).



Above, the Menéndez Pidal file in the Polymath Virtual Library shows the LOD paragraph, created by "semantic reconciliation". Under, the section that the Wikipedia dedicates to the control of authorities, carried out by means of the same system; locates it at the end of the files.

The linkage and enrichment of the authority records with DBpedia has proved to be very useful, since it automatically provides a small biography, in addition to other possible data.

Wikidata: the control of authorities in non library sources

Another source of enrichment that has emerged as a fundamental Wikidata for different reasons has also been used in this enrichment process.

Its large amount of structured data, in constant improvement; its close relationship with Wikipedia and other Wikimedia projects; and, in turn, its relationship with many other external value vocabularies to Wikipedia, belonging to archives, libraries and museums, provide it with a great value as an LOD source.

Since "Wikipedia: Authority control integration proposal" was launched on 2012, to use VIAF identifiers, a programme has been developed that detects the appropriate VIAF identifiers and the "Template: Control of authorities" of Wikipedia has been debugged, whose purpose is to automatically add the identifiers of Wikidata taken from several databases at the bottom of the articles.

Currently, the Polymath Virtual Library already appears as one of those sources used by Wikidata. Therefore, the authorities control has been extended to the Wikipedia area, which has resulted in linking large amounts of data between the authority files of cultural institutions and Wikipedia, using the large database that is Wikidata.

The importance of Wikidata is due to the possibility that any cultural institution (library, archive, museum or other) with LOD enrichment –as all those that use DIGIBÍS software: DIGIARCH, DIGIMÚS, DIGIBIB– and carry out a control over all of its records, may be included in the Wikidata vocabularies base and set out in Wikipedia in the "control of authorities" box of their articles.

As a medium and long term added value, the aggregation of authority data from different reliable LOD sources will allow the identification of authorities to be debugged; for example, two people with the same name which in one or several sources have been able to be confused as the same person, although it will never be one hundred per cent debugged. ■

Hispana reaches 7 million digital objects

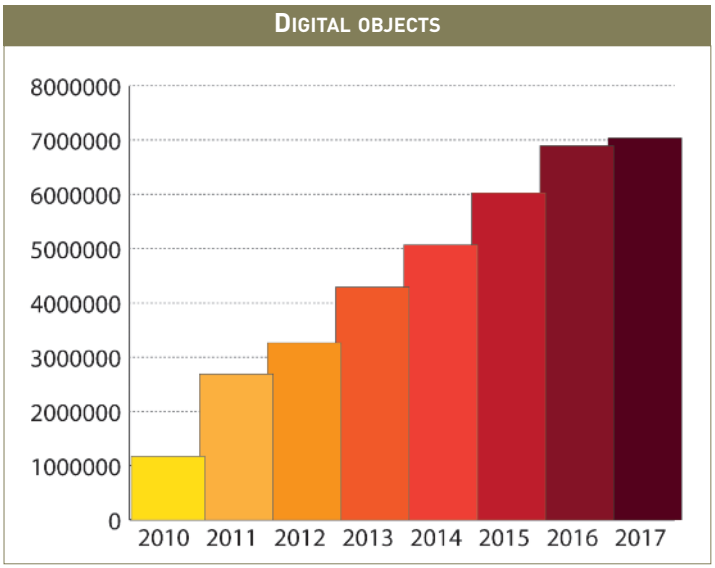
In different issues of our newsletter we provide information about the continuous growth of Hispana, the access portal to our digital culture and national aggregator of digital objects run by the Subdirectorate General for Library Coordination under the Secretary of State for Culture of the Ministry of Education, Culture and Sport (MECD).

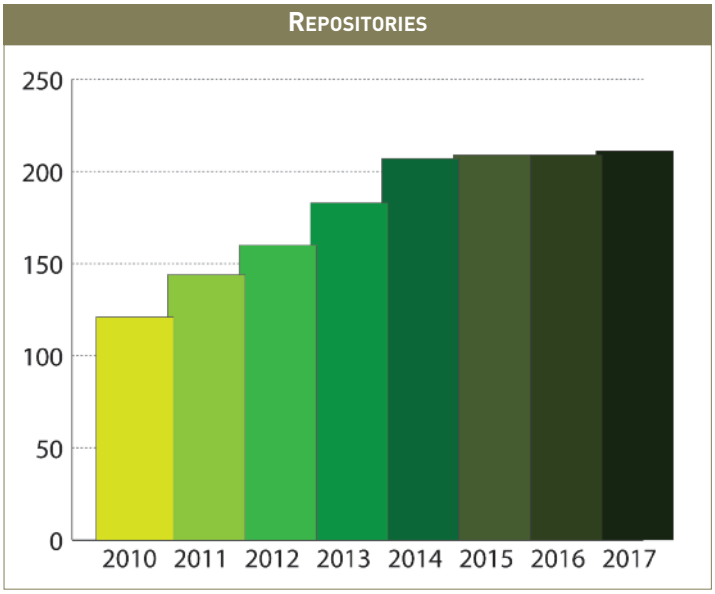
Hispana works with DIGIHUB, a powerful tool with a dual function: an OAI-PMH harvester and an OAI-PMH repository

The graphics that go along with this text show that, in spite of the complicated situation that has been affecting the country, both the MECD as well as public and private institutions contributing to Hispana, have committed to culture through the digitisation of the rich Spanish cultural heritage.

Hispana uses DIGIHUB, an application created and developed by DIGIBÍS

Since its launching in 2005, Hispana uses the DIGIHUB management system, a powerful tool with a dual function, since it allows Hispana to work as a OAI-PMH harvester, aggregating all the Spanish OAI-PMH repositories interested in participating, and as another OAI-PMH repository, which can be harvested by other meta-aggregators as occurs with the European Commission initiative Europeana. ■





Microsites in the Virtual Library of La Rioja

The Virtual Library of La Rioja has now a microsites module, thus joining the virtual libraries that have implemented microsites with DIGIBIB, software that it has used since its creation in 2007. Two microsites have been created with this module: one of them devoted to the Library of La Rioja, and another one devoted to the Institute of La Rioja Studies (IER).

With the publication of the latter, the Virtual Library of La Rioja highlights the presence of the Institute of La Rioja Studies. The microsite provides access to the digitized collections of the rich library of the Institute. The digitization of its collections, of a purely la Rioja subject, began in 2009 with a selection of works that has increased in successive digitization stages.

The microsites can be accesed from the main page of the Virtual Library of La Rioja. Both microsites, that of the Institute of La Rioja studies and the Library of La Rioja, have their own header and menu, which can be customized by the administrator. Both the searches, as well as the lists of titles, authors, places and printers, and the statistics will only show results for each microsite.

The two microsites maintain the full text search of the main website and a carousel, also customised, which shows the most outstanding works of each one of them.

The creation of new microsites has been envisaged throughout 2018, corresponding to new cultural institutions of La Rioja. ■

Upgrade of the Digital Library of the Royal Academy of Medicine

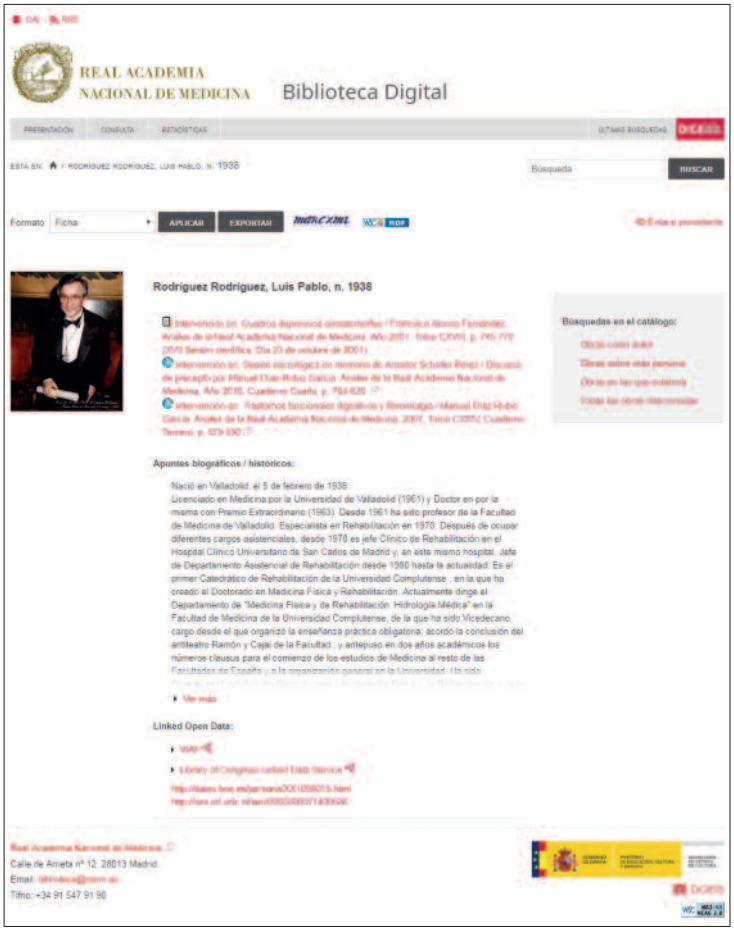
The National Royal Academy of Medicine, which was already using DIGIBIB to manage its Digital Library, has just adopted the same software for the management of its physical library, for which it has migrated the records that were managed by the Absys application.

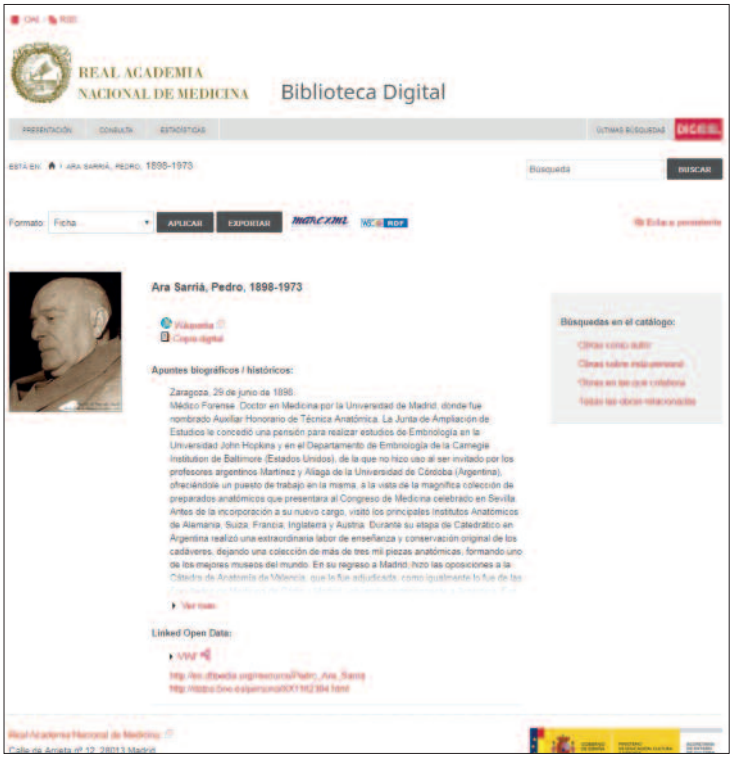
In this way, it combines all copies, both digitised and physical, in a single point of inquiry, offering users a single catalogue in which to discover and consult all its collections.

This change has coincided with the upgrade to version 9.1 of our library management system DIGIBIB. The tool is now adapted to the RDA (Resource Description and Access) and allows the semantic enrichment of authorities through Linked Open Data (LOD). In addition, the OAI-PMH repository of this Digital Library has the metadata required to be harvested by Hispana and Europeana.

> The National Royal Academy of Medicine manages with DIGIBIB both its digital library and its physical library

The Digital Library of the Royal National Academy of Medicine, thanks to a cooperation agreement with the Tatiana Pérez de Guzmán Foundation, increases the number of its scanned collections every year. It currently offers 1667 digital resources in digital format, of the almost 75000 that make up its collections. ■





Two examples of enriched authority records from the Digital Library of the Royal National Academy of Medicine.

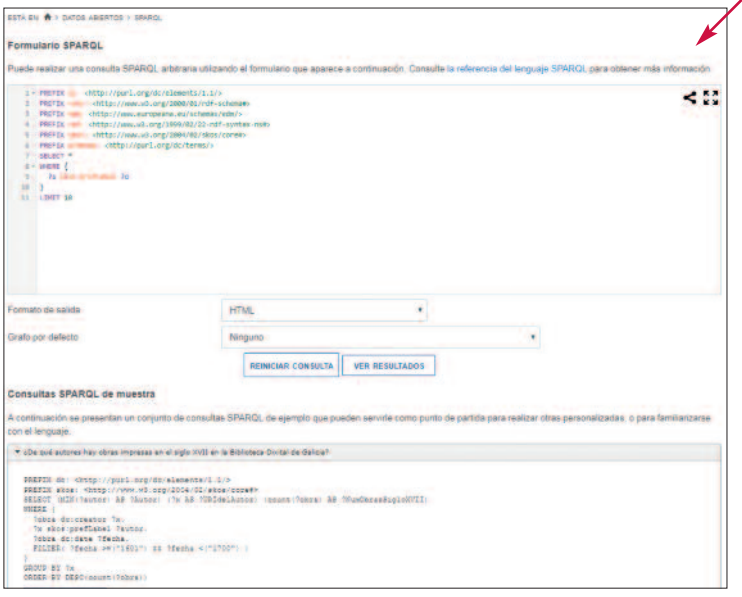
The Digital Library of Galicia adds SPARQL searches

The Digital Library of Galicia, which uses the DIGIBIB software, has been upgraded to the new version 9.1. With this upgrade, we have added a SPARQL search, requested by the Xunta de Galicia and developed in collaboration with SERDOC.

SPARQL is a query language for data in RDF format, which allows parameters that cannot be combined in other types of searches to be combined. In this way, mandatory and optional patterns together with conjunctions and disjunctions can be set in the same search, and even indicate how the results should be sorted.

The Digital Library of Galicia is part of the web portal Galiciana, Galicia Digital Heritage, which works with DIGIHUB and which also brings together the resources of the Digital Archive of Galicia.

Thus, Galicia, which is at the forefront in the digitization of cultural heritage, offers a more complex search tool that other agencies and institutions, such as the Ministry of Education, Culture and Sport in the List of Subject Headings also provide. ■



Below, SPARQL query page in the Digital Library of Galicia, which is part of, along with the Digital Archive of Galicia, of the web portal Galiciana, Digital Heritage of Galicia.

Launching of the Virtual Library of Enlightened Scientific Travels

On 27 March the Virtual Library of Enlightened Scientific Travels was presented in the auditorium of the Spanish National Research Council (CSIC). The improvement of the wind sledge, carried out by the team of Ramón Hernando de Larramendi was also presented with this. The organisation of the event was undertaken by MAPFRE Foundation and Ignacio Larramendi Foundation.

In an auditorium filled to overflowing and broadcast in streaming, José Ramón Urquijo Goitia, vice chairman of Organisation and Institutional Relations of the CSIC, was responsible for opening the event. After him spoke: Antonio Huertas Mejías, chairman of MAPFRE; Luis H. De Larramendi, chairman of the Ignacio Larramendi Foundation; Juan Velarde Fuertes, Chairman of the Royal Geographical Society; Pío Cabanillas Alonso, vice-chairman of the Spanish Geographical Society; Andrés Barbosa Alcón, a researcher at the Museum of Natural Sciences, and Ramon H. De Larramendi, polar explorer, inventor of the wind sleigh and director of Tierras Polares.

Tachi Hernando de Larramendi, director of DIGIBÍS, with participants at the event; from left to right: Andrés Barbosa, Juan Velarde, Antonio Huertas, Luis H. de Larramendi, José Ramón Urquijo, Ramón H. de Larramendi and Pío Cabanillas. On the opposite page, general view of the auditorium. Posters of the organising entities and home page of the Virtual Library.

Spanish scientific research in the18th and 19th centuries

The Virtual Library of Enlightened Scientific Travels is in charge of the Spanish historical contribution to the development of scientific research during the 18th and 19th centuries. The great enlightened navigation projects, such as those of Felix de Azara (direct precursor of Darwin), Celestino Mutis (great American flora classifier), Lorenzo Hervás and Panduro (creator of the comparative philology), Jorge Juan or Alejandro Malaspina are made known.

This digital library, which contains more than 1000 works of 23 authors, is part of the Polymath Virtual Library, created by the Ignacio Larramendi Foundation, and is the result of collaboration between this Foundation and MAPFRE Foundation.

This Virtual Library is in charge of the Spanish historical contribution to the development of scientific research during the 18th and 19th centuries

Spain at the forefront of digital preservation in the 21st Century

The Virtual Library of enlightened scientific travels is an example of how Spain is at the forefront of Europe in digital preservation of cultural heritage and its dissemination; in particular, in the development of information and communication technologies promoted by the European Digital Agenda 2020.

This virtual library is implemented, as all the Polymath Virtual Library, with the DIGIBIB application, a next generation library management system that allows free access, and from a single point of the Internet, to digital works scattered over different libraries from all over Europe and America, (through the most important repositories: Europeana and Digital Public Library of America (DPLA).

Thematic and author libraries, beneficial for scholars and non-specialist readers

Thus, all author and thematic libraries that make up the Polymath Virtual Library not only greatly facilitate the work of scholars but also put at the disposal of non-specialist readers works of great beauty and interest that they otherwise could never enjoy. Everything, from a single point of access. In addition, they also offer a very broad and encyclopaedic information about the authors, their literary work, the different names by which they are known in different languages, their relations with other authors, for which information is also offered, and opens doors to what is said about them on other websites, rigorously selected, so that the browsing from an author or a work can be extended *ad libitum.* ■

Spain at the forefront of Europe in digital preservation and dissemination of cultural heritage

AMERICANAE in the 15th FESABID Congress

The 15th Spanish Conference for Information and Documentation of the Spanish Federation of Associations for Librarians and Documentalists (Fesabid), was held in Pamplona on 25 and 26 May 2017 under the slogan "Synergies among professionals for digital transformation" Opened by Asun Maestro, director of the Library Service of the Government of Navarre, and closed by Glòria Pérez Salmerón and Alicia Selles, the conferences were divided in four areas: Synergies between the public sector and the private sector; Libraries and archives in the digital transformation; Digital narrative: transmedia and social media; and Starting up in digital transformation.

Participation of DIGIBÍS

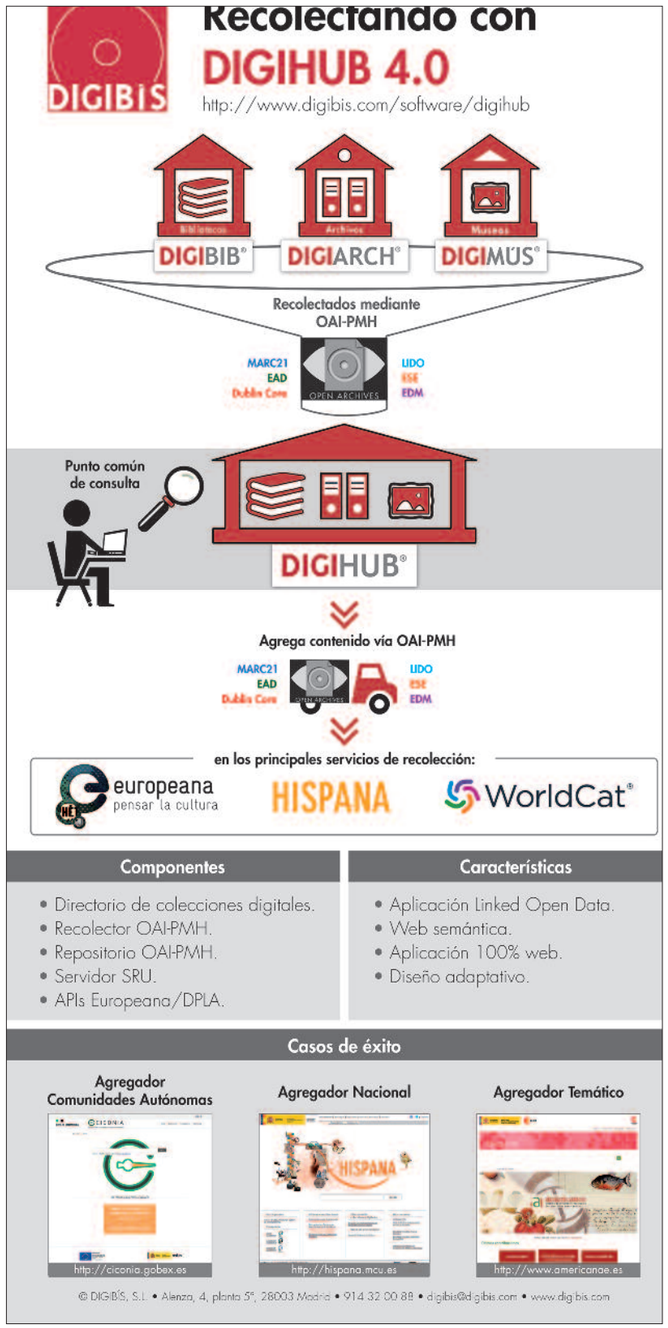
DIGIBÍS was particularly involved in the first of them, in particular in a public meeting devoted to the success cases between the public administration and private companies, with an exhibition on AMERICANAE, the result of the collaboration between the Spanish Agency for International Cooperation for Development (AECID), the Ignacio Larramendi Foundation and DIGIBÍS, which was entitled "Americanae case of success between the AECID and the Ignacio Larramendi Foundation and DIGIBÍS".

A la izquierda logo de las XV Jornadas Españolas de Información y Documentación. Arriba, algunos DIGICLIC y otras publicaciones técnicas de DIGIBÍS.

Case of success between the public administration and companies: Americanae, result of the collaboration between AECID and DIGIBÍS

It was expected that Araceli García Martín, director of the Library of the AECID, would joint Francisca Hernández and Xavier Agenjo, DIGIBÍS consultant and project director of the Ignacio Larramendi Foundation respectively, in the presentation, but, in any case, the latter two prepared a presentation giving an account of the background of that collaboration between the public administration and the private sector, the process of the presentations that took place, in this case in the application DIGIHUB, of the current reality of the project and the prospects for the future.

DIGIBÍS presented and explained its DIGIHUB application, which due to its functions was very much in line with the title of these Conferences. This application takes advantage of the synergies of the other three of our applications, those for library management (DIGIBIB), file management (DIGIARCH) and museum and gallery management (DIGIMÚS). The data of one, two or all



DIGIHUB: when synergy is very useful for professionals and users

three can be harvested, managed and disseminated in a way that is transparent for administrator users through DIGIHUB.

As we have explained in other occasions, DIGIHUB is a powerful tool that can work both as a harvester of OAI-PMH repositories, as the repository itself which, in turn, can also be harvested. Thanks to Americanae DIGIHUB, supports ESE and EDM formats.

DIGIBÍS also presented a poster on harvesting with DIGIHUB 4.0 prepared and presented by Susana Hernández, from the Department of Innovation. ■

Poster presented at Fesabid on DIGIHUB, the harvester and repository developed by DIGIBÍS that in turn allows it stored data to be harvested by other repositories.

The digitization of Defence collections in the "Success stories" of Red.es

To celebrate its 15 years of life, on 30 March 2017 Red.es has published, as part of its "Success stories," a story entitled "Red.es contributes to the recovery of archives from the Indies, photographic collections of great wars and unusual military pieces."

The news deals with the collaboration between the Ministry of Defence and the public company Red.es, in the digitization and dissemination of this large and important heritage of the Ministry of Defence, tasks in which DIGIBÍS has had the honour of participating in very directly.

The result of the collaboration between this Ministry and Red.es has been the digitalization of pieces of three collections. Two of them with the enriched digitalization that DIGIBÍS carries out, and of which we have already spoken in issue 14 of DIGICLIC: have been pieces belonging to the General Navy Archive of Álvaro de Bazán and the Army Museum, in Toledo.

The third digitized collection is formed by pieces from the Royal Institute and Observatory of the Navy in San Fernando, Cadiz, whose pieces has been scanned in a panoramic view (360º) and can be seen on the viewfinder that DIGIBÍS has prepared for its applications.

Pieces and documents have become a part of the collections offered by the Virtual Library of the Ministry of Defence, managed with the DIGIBIB application, which incorporates an OAI-PMH repository, so all the digital heritage, now and in the future, becomes part not only of the Virtual Library of the Ministry of Defence but also of Hispana and Europeana. ■

Sequence of images harvested in the 360º viewfinder of the Virtual Library of the Ministry of Defence.

Edición en línea
http://www.digibis.com/publicaciones/digiclic-boletin.html

Online Edition
http://www.digibis.com/en/publications/digiclic-newsletter.html