

Wikipedia, Wikidata y Mix'n'match

Wikipedia, Wikidata and Mix'n'match

Xavier Agenjo-Bullón; Francisca Hernández-Carrascal

Agenjo-Bullón, Xavier; Hernández-Carrascal, Francisca (2020). "Wikipedia, Wikidata y Mix'n'match". *Anuario ThinkEPI*, v. 14, e14f01.

<https://doi.org/10.3145/thinkepi.2020.e14f01>

Publicado en *IweTel* el 22 de febrero de 2020

Xavier Agenjo-Bullón

<https://orcid.org/0000-0001-8338-8087>

Fundación Ignacio Larramendi
xavier.agenjo@larramendi.es

Francisca Hernández-Carrascal

<https://orcid.org/0000-0002-2389-0945>

Digibís
francisca.hernandez@digibis.com



Resumen: Se exponen las ventajas de que los catálogos de bibliotecas, archivos y museos estén presentes en *Wikipedia*, uno de los sitios web con mayor número de accesos, proporcionando a los usuarios datos que les permitan navegar desde los artículos de *Wikipedia* a los mencionados catálogos. Para ello, los catálogos deben disponer como mínimo de URLs permanentes, siendo muy recomendable que sean además URLs http según requiere *Linked Open Data*. Para que esto sea posible, es necesario vincular los items de *Wikidata*, base de conocimientos de *Wikipedia* y fuente para su control de autoridades, con las descripciones contenidas en los catálogos. Se describe el uso de *Mix'n'match*

que facilita la generación de estos vínculos, ejemplificándolo en la *Biblioteca Virtual de Polígrafos*. De este modo los catálogos de bibliotecas, archivos y museos pueden superar los obsoletos OPACs cerrados y convertirse en catálogos abiertos.

Palabras clave: Visibilidad; Catálogos abiertos; Bibliotecas; Archivos; Museos; OPACs; *Linked open data*; Datos abiertos enlazados; Control de autoridades; *Wikipedia*; *Wikidata*; *Mix'n'match*.

Abstract: This paper exposes the advantages for library's, archives' and museums' catalogs of being present in *Wikipedia*, one of the websites with the greatest number of visits, by providing users with data that allow them to navigate from *Wikipedia* articles to the aforementioned catalogs. For this, catalogs must have at least permanent URLs, being highly recommended http URIs as required by *Linked Open Data*. For that, it is necessary to link the descriptions contained in these catalogs with the *Wikidata* items, *Wikipedia* knowledge base and source for the control of *Wikipedia* authorities. The use of the *Mix'n'match* tool to facilitate links from *Wikidata* to LAM catalogs is described, exemplifying it in the *Virtual Polygraph Library*. In this way, LAM catalogs can overcome obsolete closed OPACs and become open catalogs.

Keywords: Visibility; LAM; Libraries; Archives; Museums; LOD; *Linked open data*; Open catalogs; Authority control; *Wikipedia*; *Wikidata*; *Mix'n'match*.

1. Introducción

Uno de los mayores problemas de las bibliotecas y de los bibliotecarios es conseguir la mayor visibilidad posible de la información bibliográfica. Es obvio que, si la información no es visible, tampoco será accesible. Pero parece que los bibliotecarios contemporáneos, sobre todo en España, han perdido los vientos sobre esta cuestión. En general, en las listas de correo de biblioteconomía, en los blogs y sobre todo en las revistas profesionales, son muy pocos, poquísimos, los artículos dedicados al proceso bibliográfico. Y lo mismo se puede decir de las ponencias o comunicaciones a congresos.

Sin embargo, a menudo oímos decir ¡pero si todo está en *Google*! A veces los bibliotecarios conciben a la biblioteca, sobre todo a aquella en la que trabajan, como una empresa pública donde hay que recortar los gastos y, por otro lado, piensan en la biblioteca como una especie de ONG. Es verdad que esos dos aspectos son muy importantes, pero eso no exime de ninguna manera de las tareas de proceso bibliográfico. Es decir, hablando en plata, la catalogación y la clasificación de la colección bibliográfica.

La colección bibliográfica es lo que confiere vida a una biblioteca. Una de las mejores definiciones de biblioteca es “un conjunto ordenado de libros” (o cualquier otro material). Lo que diferencia una biblioteca de un almacén es, sobre todo, la palabra *ordenado*, y el orden de la biblioteca está marcado por la catalogación y la clasificación. Ahora bien, además de *Google*, o cualquier otro buscador, hay grandes empresas que facilitan catalogaciones e indizaciones de libros o de cualquier material; todos tenemos muy patente el ejemplo de *Amazon*. Puede que nos venga a la memoria ese diputado británico (seguro que era partidario del Brexit) que después de ver los presupuestos de las bibliotecas públicas y su uso real, proponía pagar una tasa a *Amazon* para que los usuarios de bibliotecas pidieran libros en los que estaban interesados. Y decía que le salían las cuentas.

“Una de las mayores carencias de las bibliotecas actuales es la pobreza de las descripciones bibliográficas y su falta de visibilidad”

A nosotros, bibliotecarios, esto nos parece una aberración, incluso una aberración neoliberal, pero no deja de ser inquietante que alguien haya podido comparar el uso de libros en una biblioteca con la compra de libros en *Amazon*. Esto nos tendría que hacer reflexionar.

Para nosotros —lo que se refleja en la línea de los artículos que hemos publicado en el *Anuario ThinkEPI* (Agenjo-Bullón; Hernández-Carrascal, 2018)— una de las mayores carencias de las bibliotecas actuales es la pobreza de las descripciones bibliográficas y su falta de visibilidad. Cuando uno revisa los resultados que aparecen en primer lugar tras una búsqueda en *Google*, u otros buscadores, comprueba que los primeros lugares se refieren a una serie de librerías virtuales, cuyas descripciones son comparativamente más claras que las que ofrecen las bibliotecas en la web; a veces la comparación produce un cierto sonrojo.

Es evidente que esas librerías cuentan con muchos medios para que su SEO les haga aparecer en las primeras posiciones, algo que buscan y potencian los motores de búsqueda, cuyos fines comerciales son una de las razones de su existencia. Pero también se manifiestan otros hechos como:

- la opacidad de muchos catálogos de bibliotecas a la indexación en la Web;
- la pobreza de la inteligibilidad de las descripciones bibliográficas que proporcionan las bibliotecas;
- la descontextualización de los catálogos por el mero hecho de realizar una búsqueda en la web, que no se palia con los consabidos “busque la biblioteca más cercana” (OPAC opaco).

Por eso, hace dos años publicamos en *IweTel* y luego en el *Anuario ThinkEPI* el artículo “Registros de autoridades, enriquecimiento semántico y Wikidata” (Agenjo-Bullón; Hernández-Carrascal, 2018), mostrando que con diferentes herramientas gratuitas es posible enriquecer muchísimo los registros de autoridad y, a partir de ahí, proporcionar vínculos por un lado a la *Wikipedia*, y por otro lado a los grandes repositorios virtuales del mundo, desde la *Biblioteca Nacional de España* hasta el *WorldCat*, pasando por *Hispana*, *Europeana* y otras. Y, además, a través de *Wikidata* o de *DBpedia* poder definir vínculos semánticos especiales entre autoridades, mucho más específicos que los tradicionales véase y véase además. El año pasado, 2019, dedicamos nuestros esfuerzos a probar cómo a partir del uso de *Sparql* (Agenjo-Bullón; Hernández-Carrascal, 2019) se pueden agregar registros bibliográficos con sus correspondientes objetos digitales a nuestro catálogo.

Como siempre, estos artículos están basados en las experiencias de la *Fundación Ignacio Larramendi* y de *Digibís*, y muestran ejemplos de diferentes bibliotecas virtuales que han sido enriquecidas semánticamente, como por ejemplo la *Biblioteca Digital de Castilla y León*¹, que ofrece de esta manera los autores y lugares castellano-leoneses. El último caso que conocemos es el micrositio de *Diplomaticos*

escritores² de la *Biblioteca Digital de la Agencia Española de Cooperación Internacional para el Desarrollo (AECID)*. En cuanto a la agregación de metadatos bibliográficos mediante *Sparql*, lo aplicamos a la *Biblioteca Virtual de la Escuela de Salamanca 3.0* que se presentó en octubre de 2018³.
http://www.larramendi.es/esc_salles/micrositios/inicio.do

Esta biblioteca estaba enriquecida con registros descargados a través de los servicios *Sparql* de la *Biblioteca Virtual Miguel de Cervantes* y de la *Biblioteca Nacional de España*.
<http://data.cervantesvirtual.com/sparql>
<http://datos.bne.es/sparql>

En 2019, presentamos la *Biblioteca Virtual de Novatores*, que utilizó la misma metodología de agregación, lo que permitió reunir un total de 1.824 ediciones digitalizadas, de las que casi un 29% (527 ediciones) se habían incorporado a partir de los servicios *Sparql* mencionados.
<http://www.larramendi.es/novles/micrositios/inicio.do>

Mostramos también cómo era posible enriquecer un registro de autoridad del catálogo de cualquier biblioteca con una serie de vínculos semánticos, y dado que, y esto era muy importante, estas tareas conllevan mucho tiempo de dedicación, era conveniente utilizar las distintas herramientas que hay disponibles y que facilitan muchísimo el incremento de la riqueza y calidad de la información en un tiempo considerablemente menor. Hay que suponer que los bibliotecarios conocen esas herramientas y fuentes de información, y que a partir de ahí pueden conseguir establecer vínculos con VIAF, que a su vez permiten relacionar nuestros registros de autoridad con *WorldCat Identities* o con *datos.bne.es*, por poner ejemplos de algunos de los recursos más importantes.

Merece párrafo aparte mencionar que esos dos artículos han tenido su reflejo en la obtención de los correspondientes certificados acreditativos de proyectos de I+D+i que concede *Aenor*:

- *Diseño y usabilidad de datos abiertos. Exportación, transformación y carga* (proyecto realizado en 2017);
- *Linked Open Data en producción en archivos, bibliotecas y museos* (proyecto de 2018).

También merece párrafo aparte avanzar que esas tareas se hayan realizado por parte de *Digibís* para una nueva biblioteca virtual, muy ambiciosa, que se presentará seguramente en este otoño. Y hasta aquí podemos leer.

Pues bien, siguiendo con esta estrategia, el presente artículo incide en una serie de herramientas que dan más visibilidad a las *bibliotecas* digitales o virtuales en la Web, en concreto en *Wikipedia*; así como es posible beneficiarse del control de autoridades que esta proporciona. En nuestra opinión, el control de autoridades de *Wikipedia* ha sido uno de los mayores triunfos de la biblioteconomía desde hace 30 años, es decir, desde que se constituyó la Web.

Las posibilidades de navegación que proporciona el control de autoridades de las entradas de la *Wikipedia* son casi infinitas y, desde luego, cualquier biblioteca tendría que intentar estar presente allí. Para ello, hay que pasar primero por *Wikidata*, cuyo sistema de gestión de datos y de control de autoridades es ejemplar. A quienes trabajaron con *Ariadna* les recordará, o por lo menos les hará añorar, la ontología no explícita que incorporaba ese proyecto de la *Biblioteca Nacional*.

Cuando se compara el estado de *Wikidata* para datos bibliotecarios en 2017, cuando se publicó el primero de estos artículos, con el actual, se aprecia perfectamente el avance extraordinario en cantidad y, sobre todo, en calidad producido en solo dos años, especialmente en el número de fuentes y recursos vinculados.

Por eso, este artículo tiene como propósito recalcar la importancia de *Wikidata* y hacer hincapié en la herramienta *Mix'n'match*, que se ha utilizado de forma pionera, como siempre, en la *Biblioteca Virtual de Polígrafos* de la *Fundación Ignacio Larramendi*, como tareas complementarias a la edición de la *Biblioteca Virtual de Novatores*⁴.
<https://meta.wikimedia.org/wiki/Mix%27n%27match>

Tras esta introducción general, hemos dividido el artículo en dos partes. En la primera se hace un repaso del estado de *Wikipedia* y *Wikidata* desde el punto de vista de los bibliotecarios, archiveros y museólogos; y en la segunda se entra en el detalle de la aplicación de *Mix'n'match* en la *Biblioteca Virtual de Polígrafos*.

“Es vital que las instituciones de memoria estén presentes en la Web a través de *Wikipedia* y *Wikidata*”

2. Wikipedia, Wikidata y archivos, bibliotecas y museos

En general los proyectos *Wikimedia*, entre los que destacaremos *Wikipedia*, *Wikidata* y *Wikimedia Commons*, mantienen una estrecha relación con las bibliotecas, los archivos, y los museos, tanto en el uso de los propios proyectos *Wikimedia* en estas instituciones, en tanto que fuente de información, como a su presencia en estos proyectos aportando datos, referencias y contenidos digitales de libre acceso (Williams, 2019).

<https://www.wikimedia.org>

<https://www.wikipedia.org>

https://www.wikidata.org/wiki/Wikidata:Main_Page

https://commons.wikimedia.org/wiki/Main_Page

Para muchos bibliotecarios, archiveros y museólogos los proyectos *Wikipedia* son el nuevo “lugar” en el que tienen que estar presentes las bibliotecas, los archivos y los museos. Si la *Wikipedia* está entre los cinco primeros sitios web del mundo por el número de accesos, el objetivo de las instituciones de memoria, y de los profesionales que las sirven, de estar presentes en el “lugar donde están los usuarios” se amplía a otros ámbitos y funciones.

De los múltiples aspectos de los proyectos *Wikimedia*, nos ocuparemos aquí de la relación que tiene *Wikidata* con los ficheros de autoridad de bibliotecas, archivos y museos, y especialmente de la construcción semiautomática de vínculos entre descripciones de personas, entidades, lugares y temas de ambos recursos de información. Animamos a quienes lean este artículo a que observen con detenimiento los identificadores de bibliotecas, archivos y museos que están relacionados con la descripción de El Greco, así como el apartado ‘control de autoridades’ que aparece al pie de la correspondiente entrada de *Wikipedia*.

<https://www.wikidata.org/wiki/Q301>

Como se puede ver, este ítem de *Wikidata* está relacionado con identificadores de numerosas bibliotecas, desde la *Library of Congress* a la *Biblioteca Vaticana*, pasando por la *Biblioteca Nacional de España*; pero también con los ficheros de autoridades de diferentes museos, entre ellos *El Prado*, el *Museo Thyssen-Bornemisza* o el *British Museum*, y con el *Portal de Archivos Españoles*.

La principal ventaja de establecer relaciones semánticas de las descripciones y datos de un fichero de autoridades de una institución de memoria con *Wikidata* estriba en que *Wikidata* es la base de conocimientos común para todos los proyectos *Wikimedia*, por lo que una vez establecida una relación entre un ítem de *Wikidata* y los datos de bibliotecas, archivos y museos estará disponible para su uso en las *Wikipedias* en cualquier idioma —aunque no siempre estos datos están al mismo nivel—, lo que aumenta notablemente la presencia de bibliotecas, archivos y museos en la Web. Además, *Wikidata* es uno de los recursos más importantes de la web semántica por lo que la vinculación de los datos de bibliotecas, archivos y museos con esta fuente de información incrementa la productividad de los procesos masivos de vinculación semántica.

El principal problema al que han de hacer frente las instituciones de memoria en la relación entre *Wikidata* y, en general, entre todos los proyectos *Wikimedia*, y los datos de bibliotecas, archivos y museos en un entorno *linked open data* es la baja presencia de recursos de archivos, bibliotecas y museos españoles en los proyectos *Wikipedia*, en especial en la *Wikipedia* en inglés y en la *Wikipedia* en español. Esta baja presencia puede explicarse por diferentes motivos:

- para participar en los proyectos *Wikimedia* se requiere una dedicación que habitualmente los profesionales de estas instituciones no están en situación de proporcionar;
- es preciso tener un conocimiento muy amplio de los propios datos, de su estructura y codificación, así como un gran nivel de conocimiento de los datos y estructura de datos de *Wikipedia* y *Wikidata*;
- es necesario que los datos de bibliotecas, archivos y museos dispongan de unos requisitos técnicos específicos: las URLs que identifican esos datos deben ser permanentes y unívocas y dirigir a una descripción útil para los usuarios. En este aspecto, es muy recomendable que se utilice la tecnología *linked open data* para poder incrementar, al mismo tiempo, la presencia de los datos de archivos, bibliotecas y museos en la web semántica, lo que ha venido en llamarse LOD-LAM, acrónimo que combina *linked open data*, con *libraries, archives y museums* (Stinson et al., 2016).

3. Wikipedia, Wikidata e instituciones de memoria: proyectos conjuntos

El uso de los proyectos *Wikimedia* en las instituciones de memoria está descrito y analizado en numerosos documentos, artículos y casos de uso, entre los cuáles son de obligada referencia los elaborados por la *Federación Internacional de Asociaciones de Bibliotecas y Bibliotecarios (IFLA)* y *The Wikipedia Library: Opportunities for academic and research libraries and Wikipedia* (IFLA, 2016a) y *Opportunities*

for public libraries and Wikipedia (IFLA, 2016b).

https://en.wikipedia.org/wiki/Wikipedia:The_Wikipedia_Library

Igualmente, *Europeana*, que agrupa a más de 4000 instituciones de memoria, ha desarrollado diferentes iniciativas y vías de colaboración con *Wikimedia* y *Wikidata*, hasta tal punto que son uno de los pilares de la actividad de *EuropeanaPro*.

<https://pro.europeana.eu/tags/wikimedia>

<https://pro.europeana.eu/tags/wikidata>

Algunas de estas iniciativas conjuntas son:

- Working with Wikimedia
<https://pro.europeana.eu/services/discovery/working-with-wikimedia>
- Wikidata: Wiki Loves Open Data/Europeana
https://www.wikidata.org/wiki/Wikidata:Wiki_Loves_Open_Data/Europeana
- Wikimedia Commons: Category: Europeana
<https://commons.wikimedia.org/wiki/Category:Europeana>

En definitiva, la relación entre los proyectos *Wikimedia* y las instituciones de memoria puede resumirse en los siguientes grupos de proyectos:

a) Proporcionar a los usuarios de *Wikipedia* los medios para localizar una biblioteca:

- *Wikipedia:Find your library*
https://en.wikipedia.org/wiki/Wikipedia:Find_your_library

En este proyecto es de reseñar la colaboración entre *Wikipedia* y *OCLC* para ofrecer las referencias del catálogo colectivo mundial y poder localizar la biblioteca que dispone de un ejemplar.

b) Proporcionar a los usuarios de *Wikipedia* recursos de información, digitales o analógicos, de libre acceso como, por ejemplo:

- *Wikipedia:Find your source*
https://en.wikipedia.org/wiki/Wikipedia:Find_your_source
- *Wikipedia:The Wikipedia Library/Free resources*
https://en.wikipedia.org/wiki/Wikipedia:The_Wikipedia_Library/Free_resources
- *Wikipedia:List of free online resources*
https://en.wikipedia.org/wiki/Wikipedia:List_of_free_online_resources
- *Wikipedia:Imágenes de dominio público*
https://es.wikipedia.org/wiki/Wikipedia:Imágenes_de_dominio_público
- *Wikipedia Categoría:Bibliotecas digitales* (en español)
https://es.wikipedia.org/wiki/Categor%C3%ADa:Bibliotecas_digitales
- *Wikidata:WikiProject sum of all paintings/Collection*
https://www.wikidata.org/wiki/Wikidata:WikiProject_sum_of_all_paintings
- *Wikimedia Commons: Category: Digital Libraries*
https://commons.wikimedia.org/wiki/Category:Digital_libraries
- *Wikimedia Commons: Category: Europeana*
<https://commons.wikimedia.org/wiki/Category:Europeana>

c) Interrelacionar y vincular los datos de las instituciones de memoria, especialmente los ficheros de registros de autoridad, con *Wikidata* para facilitar el control de autoridades y la vinculación de recursos de uso público.

4. Archivos, bibliotecas y museos, control de autoridades, Wikidata y la web semántica

De los múltiples aspectos de *Wikidata*, resaltaremos algunos fundamentales para las instituciones de memoria. En primer lugar, el uso de los tradicionales ficheros de autoridad de las bibliotecas, archivos y museos, resultantes de los procesos de control de autoridades, como fuentes de referencias y vínculos para los ítems de *Wikidata* y *Wikipedia*.

Con el tiempo *Wikipedia* ha ido especializando estas fuentes, categorizándolas como Identificadores, Diccionarios y enciclopedias y Repositorios digitales, tal y como puede verse en el apartado Control de autoridades de la entrada "Diego Velázquez" de la *Wikipedia* en español.

https://es.wikipedia.org/wiki/Diego_Vel%C3%A1zquez



Ilustración 1. Control de autoridades del artículo Diego Velázquez de *Wikipedia* en español
https://es.wikipedia.org/wiki/Diego_Vel%C3%A1zquez

El procedimiento realizado por *Wikidata* para el establecimiento de relaciones con registros de ficheros de autoridades de bibliotecas, archivos y museos ha implicado distintas actuaciones. Obviamente, era necesario que la ontología de *Wikidata*⁵ y el modelo de datos de *Wikibase*⁶ —plataforma que almacena y da acceso a los datos estructurados—, dispusiera de las clases relacionadas con el control de autoridades y con la asignación de identificadores únicos de fuentes externas a *Wikidata*.

Ejemplos de ello son las clases y subclases de la “propiedad de *Wikidata* para control de autoría / *Wikidata property for authority control*”, cuya traducción al castellano no refleja los matices de autoridad tal y como estamos acostumbrados, y de “identificador único / *unique Identifier*”, y “propiedad de *Wikidata* que representa un identificador único / *Wikidata property for an identifier*” que en la práctica tienen diferentes grados de solapamiento, aunque no sean subclases de una clase común, entre ellas⁷:

- a) propiedad de *Wikidata* para control de autoría / *Wikidata property for authority control* (Q18614948).
 - propiedad de *Wikidata* para control de autoría de personas / *Wikidata property for authority control for people* (Q19595382).
 - propiedad de *Wikidata* para control de autoría sobre obras / *Wikidata property for authority control for works* (Q19833377).
 - propiedad de *Wikidata* para control de autoría sobre lugares / *Wikidata property for authority control for places* (Q19829908).
 - propiedad de *Wikidata* para control de autoría sobre organizaciones / *Wikidata property to identify organisations* (Q21745557).
 - propiedad de *Wikidata* para la identificación del patrimonio cultural (Q18618628).
- b) identificador único / *unique identifier*.
 - *geographic identifier* (Q36214810). Sin etiqueta en castellano.
 - *person identifier* (Q36218176). Sin etiqueta en castellano.
 - identificador de publicación / *publication identifier* (Q36646373).
 - *organization identifier* (Q11378663). Sin etiqueta en castellano.

Una vez definidas las clases en la ontología de *Wikidata* se incorporaron masivamente vínculos con distintos ficheros de autoridad por medio de acuerdos específicos de colaboración entre distintas entidades. Ejemplo de ello son la vinculación de *Wikidata* con *VIAF* por medio del programa *VIAFbot* desde 2012 (Klein, 2012), o entre *Cantic* (*Catàleg d’Autoritats de Noms i Títols de Catalunya*) y *Wikidata* llevada a cabo en 2014 (CSUC, 2014). En este proceso se incluyó la generación de un identificador específico para cada uno de los recursos; en nuestros ejemplos,

- “VIAF ID” (*Identifier for the Virtual International Authority File database*)⁸;
- “Cantic ID” (Identificador del catálogo de autoridades de nombres y títulos de Cataluña, gestionado por la Biblioteca de Cataluña)⁹.

Al mismo tiempo, *Wikidata* proporcionó los medios para que los wikipedistas pudieran proponer la creación de propiedades específicas para los distintos identificadores de ficheros de autoridades de bibliotecas, archivos y museos. Es decir, los wikipedistas pueden proponer, a través de la página “*Wikidata:Property proposal/Authority control*”¹⁰, la creación de propiedades, que son instancias de las clases mencionadas anteriormente y sus subclases, que deben ser aprobadas por un “creador de propiedades” o un administrador. En el enlace referenciado se pueden consultar las condiciones y requisitos de aprobación o rechazo, así como los plazos para ello. Como en cualquier otro aspecto de los proyectos *Wikimedia*, esta propuesta la puede efectuar cualquier miembro de estos proyectos o cualquier usuario identificado. Es muy importante señalar que este es un paso previo obligatorio para la vinculación masiva de registros de autoridad de cualquier fuente externa y *Wikidata*, y que para que los registros

de autoridad de bibliotecas, archivos y museos se puedan vincular con *Wikidata*, estos deben disponer obligatoriamente de un identificador permanente. Por ejemplo, la estructura del identificador permanente de la *Biblioteca Virtual de Polígrafos* es:

[http://www.larramendi.es/aut/\\$1](http://www.larramendi.es/aut/$1)

como el correspondiente a Menéndez Pelayo:

<http://www.larramendi.es/aut/POLI2009001509>

O en *Pares* que es del tipo:

[http://pares.mcu.es/ParesBusquedas20/catalogo/autoridad/\\$1](http://pares.mcu.es/ParesBusquedas20/catalogo/autoridad/$1)

como el correspondiente a Felipe II:

<http://pares.mcu.es/ParesBusquedas20/catalogo/autoridad/46876>.

Desde el momento de su puesta en funcionamiento en 2012, *Wikidata* se ha ido consolidado como una de las fuentes de datos estructurados más importantes de la web semántica. Según la última *The Linked Open Data Cloud*¹¹ de marzo de 2019, el conjunto de datos de *Wikidata* está formado por 5800 millones de tripletas¹², y *DBpedia* por 9500 millones de tripletas¹³. Estas cifras dan una idea de la importancia que tienen estos dos proyectos en la web semántica como recursos de información en sí mismos y como fuente de datos estructurados para otros conjuntos de datos *linked open data*. Aunque *Wikidata* y *DBpedia* son proyectos con una estructura, gestión y dirección diferentes, en la práctica es difícil separarlos por las estrechas relaciones entre los datos de ambas.

Según el estado de la “nube” en marzo de 2019, los ítems de *Wikidata* están relacionados con diferentes identificadores de bibliotecas, archivos y museos, entre los que destacamos los recursos y número de entradas vinculadas de algunas de las instituciones bibliotecarias más importantes del mundo:

- *Virtual International Authority File (OCLC)*: 13.441.955.
- *Gemeinsame Normdatei (GND) - Deutschen Nationalbibliothek*: 802.336
- *Library of Congress*: 608.446.
- *datos.bne.es (Biblioteca Nacional de España)*: 97.716. Sin embargo, en el momento de redactar este texto, enero de 2020, el número de identificadores de la *BNE* en *Wikidata* es de 159.064, lo que da idea de la progresión de estos vínculos.

5. *Mix'n'match*: vinculación masiva de ficheros de autoridad de archivos, bibliotecas y museos con *Wikidata*

Llegados a este punto estaríamos ya en situación de continuar con la adición de nuevas relaciones con ficheros de autoridad de bibliotecas, archivos y museos de forma masiva en nuestro proyecto.

Como procedimiento para añadir relaciones entre *Wikidata* y los ficheros de autoridad, los wikipedistas tienen a su disposición el conjunto de herramientas denominado *Mix'n'match*, que permite subir catálogos de datos para encontrar automáticamente la entrada correspondiente de *Wikidata*.

<https://tools.wmflabs.org/mix-n-match/#>

Lógicamente, en algunos casos la correspondencia es directa, mientras que en otros es necesario que un usuario confirme una correspondencia determinada entre las posibles que la aplicación detecta. Para ello dispone de dos modos, el modo manual y el modo semiautomático.

Mix'n'match es un desarrollo de Magnus Manksee que ha recibido en 2019 el *WikidataCon Award*, celebrado por primera vez en este año. Uno de los grandes difusores de esta herramienta ha sido *Europeana*, que ha utilizado *Mix'n'match* para sus procesos de vinculación semántica y ha elaborado una recomendación para que los proveedores de datos de *Europeana* la utilicen, de modo que estos vínculos estén disponibles para su reutilización tanto por *Europeana* como por cualquier proveedor de datos de *Europeana* o de otras instituciones. El uso de *Mix'n'match* por *Europeana* se ha descrito en el documento de *EuropeanaTech: Get your vocabulaires in Wikidata... so Europeana and others can get them* (*Europeana*, 2017). También *Mix'n'match* se ha utilizado en talleres celebrados en el marco de las reuniones anuales de *EuropeanaTech*, como la de 2018¹⁴.

“Se dispone de herramientas avanzadas como *Mix'n'match* para optimizar la presencia de los objetos digitales en *Wikipedia* y *Wikidata*”

6. La *Biblioteca Virtual de Polígrafos* y *Mix'n'match*

La *Fundación Ignacio Larramendi* ha utilizado esta herramienta justamente para facilitar la vinculación masiva de los registros de autoridad de la *Biblioteca Virtual de Polígrafos* con *Wikidata*. El proceso

completo se ha realizado sobre los 858 registros de autoridad de polígrafos de la *Biblioteca Virtual de Polígrafos*, que previamente habían sido vinculados y enriquecidos semánticamente con distintas fuentes, entre ellas *Wikidata* (Agenjo-Bullón; Hernández-Carrascal, 2018).

El principal resultado ha sido la concreción de un procedimiento de trabajo que puede realizarse para distintas instituciones culturales interesadas. En definitiva, ha dado lugar a aquilatar los subprocesos que se deben seguir, su orden y el esfuerzo en horas necesario para su realización.

Como paso previo se creó una entrada en la *Wikipedia* para la *Biblioteca Virtual de Polígrafos* de la *Fundación Ignacio Larramendi*. Esta entrada está en la dirección: https://es.wikipedia.org/wiki/Biblioteca_Virtual_de_Polígrafos

y está identificada en *Wikidata* con el identificador Q53114817. El objetivo era lograr que al final del proceso de cotejo se pudiera obtener para las personas incluidas en la *Biblioteca Virtual de Polígrafos* un identificador en *Wikidata*

<https://www.wikidata.org/wiki/Property:P6907>

que fuera una instancia de la clase "*Wikidata property for authority control for people*"

<https://www.wikidata.org/wiki/Q19595382>

correctamente interrelacionadas, siendo la *Biblioteca Virtual de Polígrafos* el origen de los datos.

“La herramienta *Mix'n'match* se ha probado en la *Biblioteca Virtual de Polígrafos* de la *Fundación Ignacio Larramendi* a plena satisfacción”

Wikidata:Property proposal/BVLarramendi

< Wikidata:Property proposal

BVLarramendi ID [edit]

Originally proposed at Wikidata:Property proposal/Authority control

Done: BVLarramendi ID (P6907) (Talk and documentation)

Description	Identifier for a person on the Polymath Virtual Library
Data type	External identifier
Domain	human (Q5)
Allowed values	[A-Z0-9]{15}
Example 1	Marcelino Menéndez y Pelayo (Q716799) → POLI20090015098 ↗
Example 2	Santiago Ramón y Cajal (Q150526) → POLI20090015074 ↗
Example 3	Juan de Mariana (Q939889) → POLI20090012677 ↗
Source	http://www.larramendi.es/es/inicio/inicio.do ↗
External links	Use in sister projects: [ar] ↗ • [de] ↗ • [en] ↗ • [es] ↗ • [fr] ↗ • [he] ↗ • [it] ↗ • [ja] ↗ • [ko] ↗ • [nl] ↗ • [pl] ↗ • [pt] ↗ • [ru] ↗ • [sv] ↗ • [vi] ↗ • [zh] ↗ • [commons] ↗ • [species] ↗ • [wd] ↗
Planned use	Displaying links in Wikipedia through the authority control module and facilitate links to other sources through Wikidata
Number of IDs in source	874
Expected completeness	eventually complete (Q21873974)
Formatter URL	http://www.larramendi.es/aut/\$1
Robot and gadget jobs	https://es.wikipedia.org/wiki/Plantilla:Larramendi ↗

Motivation [edit]

Polymath Virtual Library (Q53114817) is a virtual library managed by the Fundación Ignacio Larramendi (Q24943800). Through its identifier gives access to open access digitized works of Spanish, Portuguese, and Iberoamerican thinkers of any era whose works have had a great influence on the history of thought and philosophy. Gnjo55 (talk) 15:34, 16 June 2019 (UTC)

Discussion [edit]

- Support David (talk) 06:54, 17 June 2019 (UTC)
- Support Dhx1 (talk) 15:11, 18 June 2019 (UTC)

@2 ديديد عادل وحة خليل Gnjo55, Dhx1: Done: BVLarramendi ID (P6907). – Pinto (talk) 22:41, 28 June 2019 (UTC)

Category: Property proposal authority control

Ilustración 2. Plantilla de propuesta de propiedad para un identificador de personas en la *Biblioteca Virtual de Polígrafos*

Además, según se ha dicho, ya se había efectuado también la reconciliación de todos los registros de polígrafos (858) con *Wikidata* y se había obtenido el nombre (*label*) en castellano de las entradas correspondientes en *Wikipedia* y *Wikidata*. Esta reconciliación facilitó el cotejo automático y minimizó los efectos de la comparación de los nombres de las personas en orden inverso (en la *Biblioteca Virtual de Polígrafos*) y en orden directo como se usa en *Wikipedia* y *Wikidata*.

Como ya hemos recomendado, el primer paso fue crear en *Wikidata* una propuesta para la propiedad "*BVLarramendi ID*", para lo que utilizamos los mecanismos y plantillas disponibles en "*Wikidata:Property proposal/Authority control*".

Tras la aceptación de esta propuesta de propiedad para identificar a una persona en la *Biblioteca*

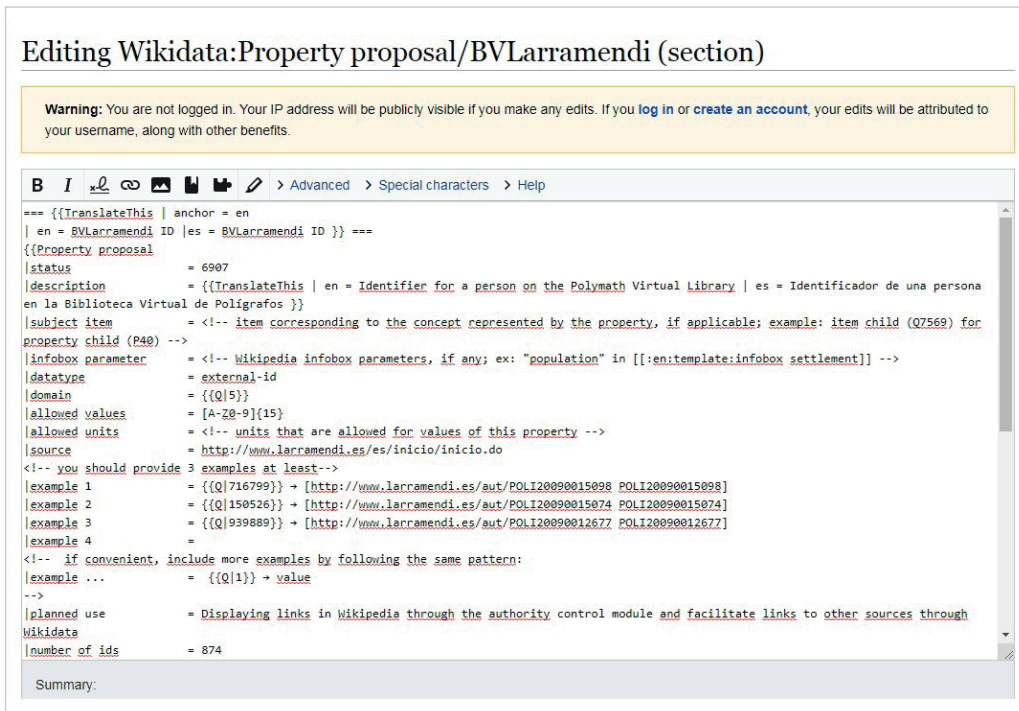


Ilustración 3. Plantilla de discusión sobre el identificador *BVLarramendi ID*

Virtual de Polígrafos, iniciamos la vinculación automática a través de *Mix'n'match*, que supuso diferentes tareas. En primer lugar, se extrajeron los registros de autoridad de polígrafos que conformaron un fichero TSV (valores separados por tabulaciones, por sus siglas en inglés) con todos aquellos campos que facilitarían la identificación de las personas descritas y, por tanto, proporcionarían los suficientes datos para permitir la vinculación semiautomática o manual.

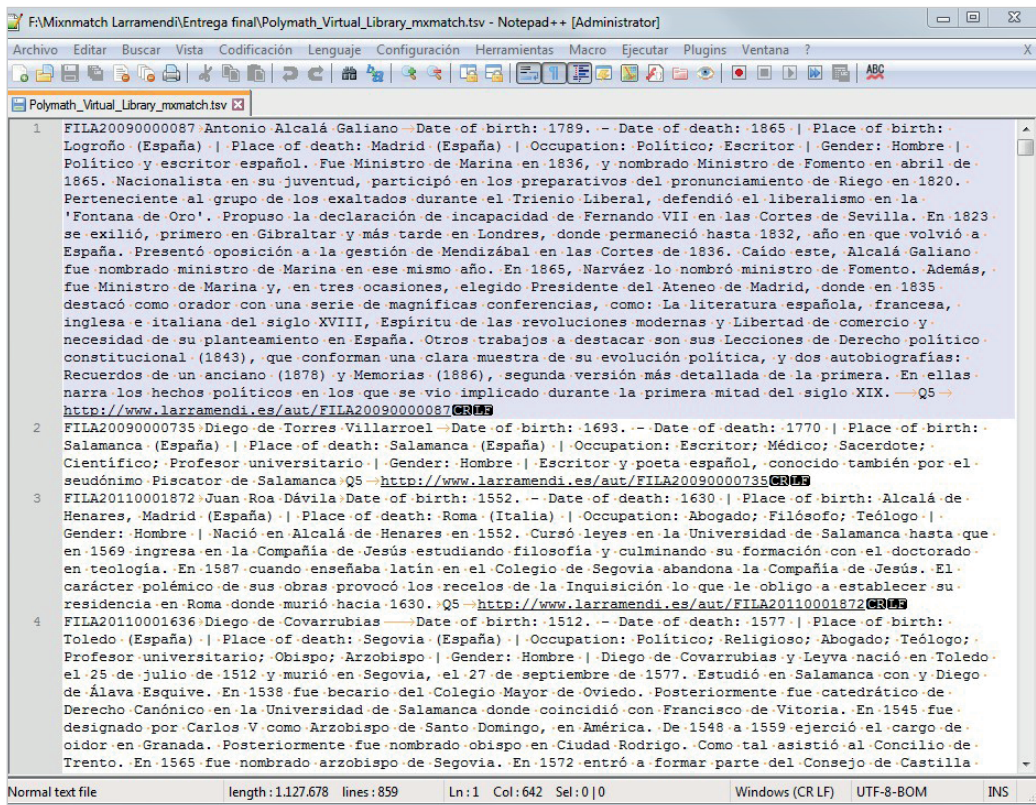


Ilustración 4. Fichero TSV con las descripciones de polígrafos

Los campos que se extrajeron incluían:

- el 001 identificativo del registro en la *Biblioteca Virtual de Polígrafos*, elemento básico para generar el identificador de esta biblioteca;
- el campo 100 con el nombre preferido de cada polígrafo, que siguiendo la práctica habitual se proporciona en orden inverso (p.e.: Menéndez Pelayo, Marcelino);
- el campo 678 con una breve descripción biográfica elaborada por la *Fundación Ignacio Larramendi*; y otros datos como las fechas de nacimiento y defunción, el lugar de nacimiento y/o defunción, el campo de actividad, o la ocupación.

Y, finalmente, se añadió el nombre de la entrada correspondiente a cada polígrafo en *Wikidata*, en orden directo. Además de las diferencias en el orden de nombres y apellidos, puede haber muchas otras como la ausencia o presencia de algún apellido, sobrenombres, u otras. Como ya disponíamos de los nombres de las entradas correspondientes a *Wikipedia* y *Wikidata*, las utilizamos como “nombre” para facilitar que la vinculación automática fuera lo más amplia posible.

De todos los datos extraídos que hemos mencionado, solo algunos son obligatorios, pero en conjunto tienen como finalidad, como se ha dicho, proporcionar a los usuarios de *Mix'n'match* información complementaria para cotejar manualmente los casos en los que no se haya producido una vinculación automática con la suficiente información¹⁵.

Uno de los componentes de *Mix'n'match* es la herramienta de importación, *Import Tool* o *Mix'n'match Importer*¹⁶, que para el usuario es un formulario en el que se describe el catálogo al que se refiere la importación y cotejo (nombre, descripción, URL, tipo de catálogo, propiedad de *Wikidata*, idioma, patrón de URL y tipo de entradas). En nuestro caso todos estos datos estaban referidos a la *Biblioteca Virtual de Polígrafos*. En esta misma página se puede subir el fichero TSV a cotejar, que estará compuesto por las siguientes columnas obligatorias: identificador, nombre y descripción. De forma opcional se pueden añadir otras columnas como el tipo de entrada (en nuestro caso personas, es decir Q5) y URL de cada entrada.

Mix'n'match importer
Tools Git Talk

Here, you can import a new catalog into Mix'n'match.
Just paste your tab-delimited text (copy from Excel should be fine) into the textbox, and fill out the form!
This import function is experimental! Please be careful, and contact Magnus if something goes wrong!

Catalog name		A short name for your catalog (like "VIAF")
Catalog description		A brief description of your catalog
Catalog URL		The URL (main page) of your catalog
Catalog type	Unknown ⌵	The group under which your catalog will appear on the Mix'n'match main page
Wikidata property		The Wikidata property for this catalog; <i>optional</i>
Main language	en	Which Wikipedia to search for this catalog (language code)
URL pattern		A URL pattern to get to entries from their \$id (e.g "http://dummy.org/\$id"); <i>optional</i>
Default type		The default type of the entries (e.g "Q5"); <i>optional</i> If you use a Q-ID (e.g. Q11424) here, automatches will be limited to items linking to this item and its subclasses, resulting in fewer matches of higher quality

The following (tab-delimited, un-quoted) columns are required, *in this order*:

1. **Entry ID** (your alphanumeric identifier; must be unique within the catalog)
2. **Entry name** (will also be used for the search in mix'n'match later)
3. **Entry description**

Optionally, the following columns can be added *in this order*:

4. **Entry type** (item identifier, e.g. "Q5"; recommended)
5. **Entry URL**; if omitted, it will be constructed from the URL pattern and the entry ID. *Either a URL pattern or a URL column are required!*

Text needs to be UTF-8. You can either paste the tab-delimited text into the box (for shorter catalogs), or upload it as as tab-delimited file (max. 20MB).

Ilustración 5. Ficha de catálogo de la herramienta *Mix'n'match importer*

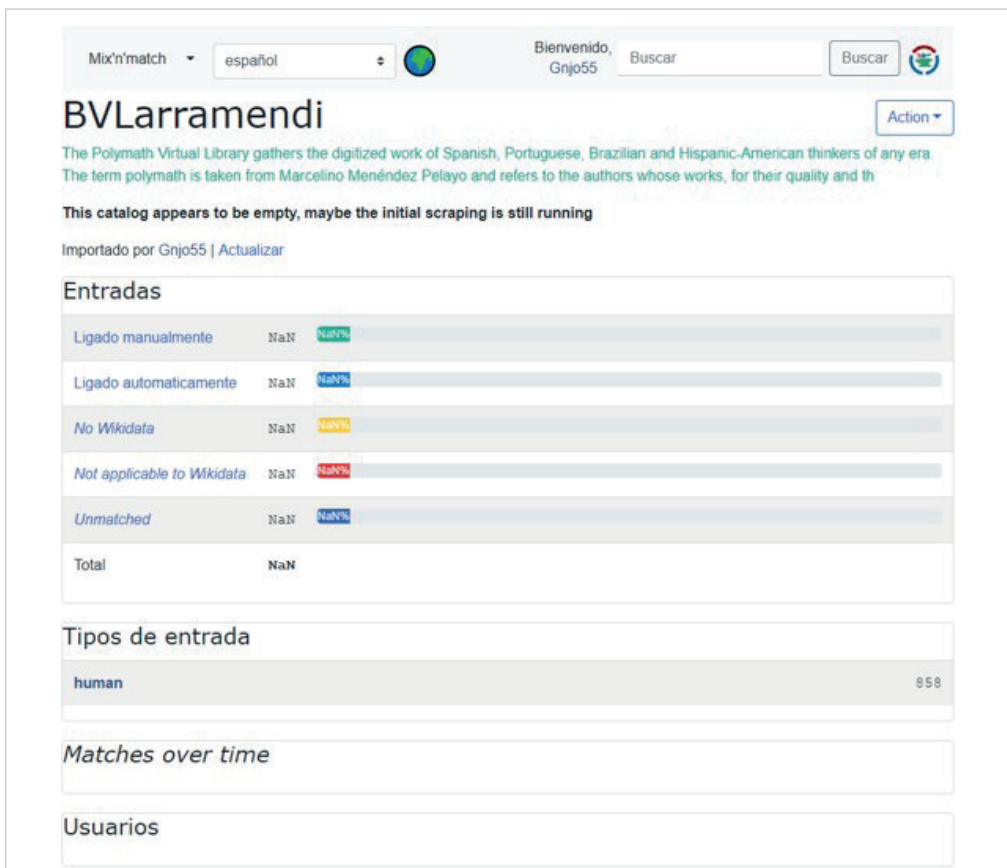


Ilustración 6. Inicio del proceso de vinculación automática con *Mix'n'Match importer*

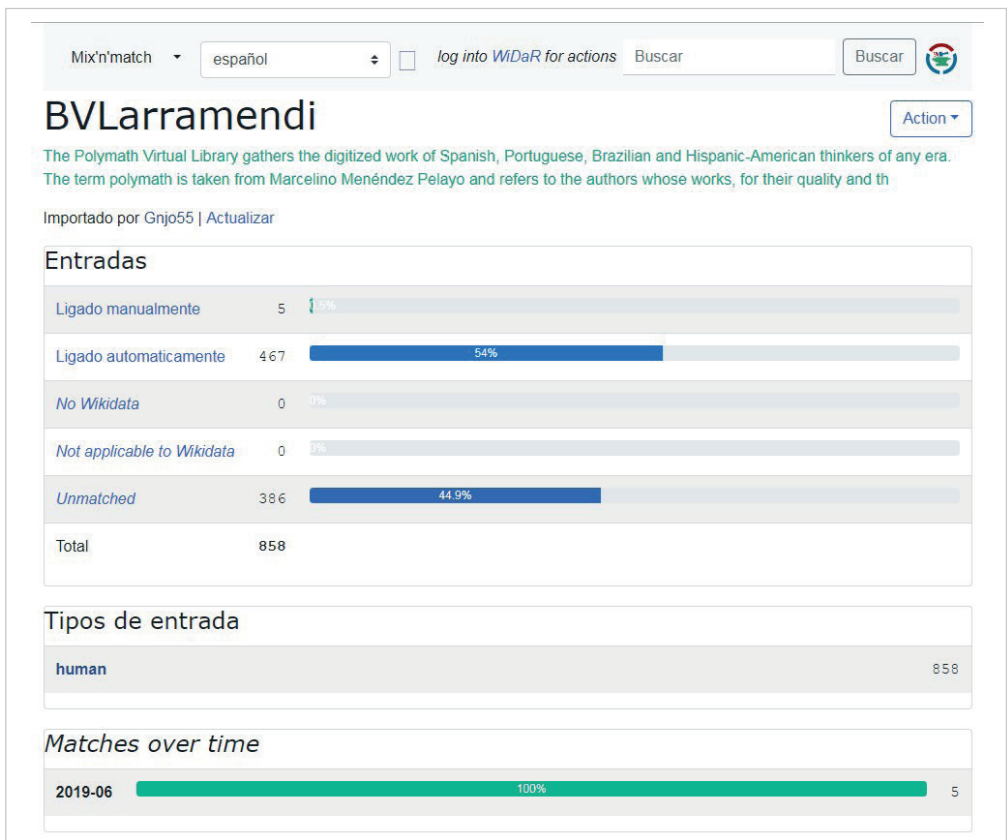


Ilustración 7. Resultado de la vinculación automática con *Mix'n'Match importer*

The screenshot shows the BVLarramendi interface with a search bar set to 'español'. Below the header, there are navigation tabs numbered 1 to 10. The main content area displays a list of entries, each with a name, a brief biographical summary, and a status indicating they are 'Ligado automáticamente' (automatically linked). The entries include:

- Juan María Gutiérrez**: Date of birth: 1809. - Date of death: 1878 | Place of birth: Buenos Aires (Argentina) | Place of death: Buenos Aires (Argentina) | Occupation: Historiador; Político; Ministro; Profesor universitario | Gender: Hombre | Escritor, historiador, educador y po...
- Gumersindo de Azcárate**: "Date of birth: 1840. - Date of death: 1917 | Place of birth: León (España) | Place of death: Madrid (España) | Occupation: Abogado; Político; Profesor universitario; Historiador; Filósofo | Gender: Hombre | Abogado, político y reformador social. Era
- Juan de Mariana**: "Date of birth: 1536. - Date of death: 1624 | Place of birth: Talavera de la Reina, Toledo (España) | Place of death: Toledo (España) | Occupation: Teólogo; Historiador; Profesor universitario; Economista; Sacerdote; Abogado | Gender: Hombre | Teólogo España historiador, profesor universitario, economista (1536–1624) ♂
- Andrés Bello**: Date of birth: 1781. - Date of death: 1865 | Place of birth: Caracas (Venezuela) | Place of death: Santiago de Chile (Chile) | Occupation: Filólogo; Profesor; Abogado; Escritor | Gender: Hombre | Educador, humanista, poeta, ensayista, crítico, político, Chile-venezuela poeta, legislador, filósofo, educador, filólogo, diplomático, político, traductor, senador de Chile (1781–1865) ♂
- Juan Bautista Alberdi**: "Date of birth: 1810. - Date of death: 1884 | Place of birth: San Miguel de Tucumán (Argentina) | Place of death: París (Francia) | Occupation: Escritor; Historiador; Filósofo; Abogado; Periodista | Gender: Hombre | Escritor, historiador, filósofo, jur
- Carlos de Sigüenza y Góngora**: "Date of birth: 1645. - Date of death: 1700 | Place of birth: Ciudad de México (México) | Place of death: Ciudad de México (México) | Occupation: Historiador; Humanista; Geógrafo; Matemático; Cartógrafo; Sacerdote; Científico | Gender: Hombre | Estudi

Ilustración 8. Entradas vinculadas automáticamente con *Mix'n'Match importer* (467 entradas, 54,4 %)

The screenshot shows the BVLarramendi interface with a search bar set to 'español'. Below the header, there are navigation tabs numbered 1 to 8. The main content area displays a list of entries, each with a name, a brief biographical summary, and a status indicating they are 'Not matched'. The entries include:

- Antonio Alcalá Galiano**: Date of birth: 1789. - Date of death: 1865 | Place of birth: Logroño (España) | Place of death: Madrid (España) | Occupation: Político; Escritor | Gender: Hombre | Político y escritor español. Fue Ministro de Marina en 1836, y nombrado Ministro de Fom
- Diego de Torres Villarroel**: Date of birth: 1693. - Date of death: 1770 | Place of birth: Salamanca (España) | Place of death: Salamanca (España) | Occupation: Escritor; Médico; Sacerdote; Científico; Profesor universitario | Gender: Hombre | Escritor y poeta español, conocido tam
- Juan Roa Dávila**: Date of birth: 1552. - Date of death: 1630 | Place of birth: Alcalá de Henares, Madrid (España) | Place of death: Roma (Italia) | Occupation: Abogado; Filósofo; Teólogo | Gender: Hombre | Nació en Alcalá de Henares en 1552. Cursó leyes en la Universi
- Domingo de Soto**: "Date of birth: 1494. - Date of death: 1560 | Place of birth: Segovia (España) | Place of death: Salamanca (España) | Occupation: Abogado; Filósofo; Teólogo; Físico; Profesor universitario | Gender: Hombre | Figura esencial del pensamiento jurídico, fi
- Diego Pérez de Mesa**: Not matched
- Rafael Altamira**: "Date of birth: 1866. - Date of death: 1951 | Place of birth: Alicante (España) | Place of death: Ciudad de México (México) | Occupation: Abogado; Historiador; Juez | Gender: Hombre | Jurista e historiador. Investigador de temas americanos y teórico y

Ilustración 9. Entradas no vinculadas con *Mix'n'Match importer* (389 entradas, 45,3%)

Una vez cargado el fichero a cotejar, *Mix'n'match* realiza un cotejo automático y presenta una lista de resultados en la que se muestran, en porcentajes, las entradas enlazadas automáticamente, las que deben ser ligadas manualmente, las que no encuentran ninguna relación con algún ítem de *Wikidata*, etc.

La pobre *Wikidata* se las vio y deseó para cotejar los escritores visigóticos. Bien es verdad que la *Biblioteca Virtual de Escritores Visigóticos y Mozárabes*¹⁷ de la *Biblioteca Virtual de Polígrafos* constituye el mayor recurso del mundo sobre este período y los autores que lo componen.

El proceso concluyó con un manual de procedimiento y una estimación de esfuerzos, para la realización del mismo proceso para otras instituciones. En la tabla que se incluye como anexo se puede apreciar el estado de la vinculación de los ficheros de autoridad de bibliotecas, archivos y museos en España, en el momento de la redacción de este texto, a finales de enero de 2020, tal y como se obtiene del servidor *Sparql* de *Wikidata* tras la correspondiente pregunta.

<https://w.wiki/GsC>

Para aquellos que estén interesados pueden consultar este mismo estado, actualizado al momento en el que hagan clic en el enlace referenciado.

Por último, solo nos queda mencionar el lanzamiento del proyecto *Soweego*¹⁸, aprobado por *Wikimedia Foundation* en agosto de 2019 con la finalidad de reunir diferentes algoritmos de aprendizaje automático para obtener enlaces de la mayor calidad entre *Wikidata* y grandes catálogos. Su objetivo es complementar las tareas que ejecuta *Mix'n'match* para volúmenes de información muy grandes.

7. Conclusiones

Como se ha dicho en la introducción, las bibliotecas, archivos y museos tienen que cuidar muchísimo la información bibliográfica, archivística o museológica que se pone a disposición de los usuarios para dar la máxima visibilidad posible, lo que redundará en la accesibilidad de los documentos. De esta forma, los usuarios que, casi nunca, van a consultar las páginas web de esas instituciones puedan buscar, por ejemplo, a través de la *Wikipedia* esa información y a partir de ahí navegar a las instituciones de memoria. Si no se hace así se corre el riesgo de que los catálogos se conviertan en casi irrelevantes. Para hacer una analogía, puede pasar algo similar a lo que ocurre con esas notaciones detalladas, basadas en la CDU o cualquier otro sistema de clasificación, que ningún usuario utiliza, y que en el 90% de los casos no sirven ni para facilitar la colocación de los libros en las estanterías de libre acceso.

Para conseguir que las tareas catalográficas, sobre todo las de control de autoridades, lleguen a tener una visibilidad suficiente, una herramienta como *Mix'n'match* ha probado en la *Biblioteca Virtual de Polígrafos* una utilidad admirable. Se invita al improbable lector de este artículo a efectuar una búsqueda en *Google* de la cabecera de prensa "El liberal de Reus". Y a que vean cuáles son los resultados. Si hacemos una apuesta, cuando se teclaa "El liberal de Reus" en *Google* o en cualquier otro navegador, y el lector accede a la *Biblioteca Virtual de Prensa Histórica*, los autores de este artículo ganan; en el caso contrario, el lector no gana nada, incluso ha perdido el tiempo leyendo este artículo.

Por otro lado, la web semántica, o incluso la Web normal, proporciona ingentes recursos de información vinculada. ¿Por qué vamos a prescindir de esta información cuando puede ser vital para el lector? Cuando se llega a un registro bibliográfico, ese registro tiene que tener una serie de funcionalidades que permitan que el lector conozca la bibliografía del autor, las otras obras de este autor en todas las bibliotecas disponibles; y en el caso de que haya digitalizaciones de la obra, ofrecérselas, etc. Ahora, la mayoría de los catálogos no proporcionan eso, y en el caso de que la biblioteca forme parte de un catálogo colectivo, se trata de un catálogo cerrado, cuando con las herramientas que hemos descrito en este artículo, se puede facilitar (y además con un control de calidad muy importante) un catálogo abierto, para que el bibliotecario cumpla la cuarta ley de Ranganathan que es tan fácil como "save the time of the user".

"El porcentaje de correspondencia automática del catálogo de autoridades en Wikidata es del 54,4%"

"Con herramientas como *Mix'n'match* se puede conseguir que los catálogos sean totalmente abiertos y vinculados con otros recursos de información en la Red"

8. Notas

1. Castilla y León: Datos Abiertos Vinculados. *Biblioteca Digital de Castilla y León*.
<http://bibliotecadigital.jcy.es/autcl/es/micrositios/inicio.cmd>
2. *Diplomáticos escritores*. *Biblioteca Digital DECID*.
http://bibliotecadigital.aecid.es/bibliodig/diplomaticos/es/consulta_aut/busqueda.cmd
3. Presentada la *Biblioteca Virtual de la Escuela de Salamanca 3.0* en la *USAL*.
<http://www.larramendi.es/fundacion/presentada-la-biblioteca-virtual-de-la-escuela-de-salamanca-3-0-en-la-usal-2/>
4. *Biblioteca Virtual de Novatores*. Nota a esta edición digital. 2019.
http://www.larramendi.es/nov/es/cms/elemento.do?id=ms%2Fnov%2Fpaginas%2FNota_a_esta_edicion_digital.html
5. *Wikidata:WikiProject Ontology*.
https://www.wikidata.org/wiki/Wikidata:WikiProject_Ontology
6. *Wikibase Data Model*.
<https://www.mediawiki.org/wiki/Wikibase/DataModel>
7. Todas las propiedades subclase de la propiedad *Authority Control*, 59 en el momento de redactar este texto, se pueden ver en la correspondiente consulta *Sparql*.
<https://w.wiki/GsH>
8. *VIAF ID*.
<https://www.wikidata.org/wiki/Property:P214>
9. *Cantic ID*.
<https://www.wikidata.org/wiki/Property:P1273>
10. *Wikidata:Property proposal/Authority control*.
https://www.wikidata.org/wiki/Wikidata:Property_proposal/Authority_control
11. *The Linked Open Data Cloud*. March 2019.
<https://lod-cloud.net>
12. *The Linked Open Data Cloud*. *Wikidata*.
<https://lod-cloud.net/dataset/wikidata>
13. *The Linked Open Data Cloud*. *DBpedia*.
<https://lod-cloud.net/dataset/dbpedia>
14. *EuropeanaTech 2018 Wikidata workshop day*.
https://www.wikidata.org/wiki/Wikidata:Events/EuropeanaTech_2018_Wikidata_workshop_day
15. *Mix'n'match/Manual*.
[https://meta.wikimedia.org/wiki/Mix%27n%27match/Manual#Semi-automatic_mode_\(Game_mode\)](https://meta.wikimedia.org/wiki/Mix%27n%27match/Manual#Semi-automatic_mode_(Game_mode))
16. *Mix'n'match importer*.
<https://tools.wmflabs.org/mix-n-match/import.php>
17. *Biblioteca Virtual de Escritores Visigóticos y Mozárabes*.
<http://www.larramendi.es/visigles/micrositios/inicio.do>
18. *Grants:Project/Rapid/Hjfofcs/soweego 1.1*.
<https://meta.wikimedia.org/wiki/Grants:Project/Rapid>

9. Referencias

- Agénjo-Bullón, Xavier** (2019). *Biblioteca Virtual de Novatores*. Presentación realizada el 30 de octubre de 2019 en la Real Academia de Doctores de España.
http://www.larramendi.es/fundacion/wp-content/uploads/2019/10/novatores_presentacion.pdf
- Agénjo-Bullón, Xavier; Hernández-Carrascal, Francisca** (2018). "Registros de autoridades, enriquecimiento semántico y Wikidata". *Anuario ThinkEPI*, v. 12, pp. 361-372.
<https://doi.org/10.3145/thinkepi.2018.61>
- Agénjo-Bullón, Xavier; Hernández-Carrascal, Francisca** (2019). "Agregación de datos bibliográficos por medio de servicios Sparql". *Anuario ThinkEPI*, v. 13, e13f02.
<https://doi.org/10.3145/thinkepi.2019.e13f02>
- CSUC (2014). "Cantic a Viquipèdia i Wikidata". *El Blog del CSUC: Actualitat del Consorci de Serveis Universitaris de Catalunya*, 23 septiembre..
<http://blog.csuc.cat/?p=1987>

Europeana (2017). "Get your vocabularies in Wikidata... so Europeana and others can get them". *EuropeanaTech*, 1 September.

<https://pro.europeana.eu/page/get-your-vocabularies-in-wikidata>

IFLA (2016a). *Opportunities for academic and research libraries and Wikipedia*. Federación Internacional de Asociaciones de Bibliotecas y Bibliotecarios.

<https://www.ifla.org/files/assets/hq/topics/info-society/iflawikipediaopportunitiesforacademicandresearchlibraries.pdf>

IFLA (2016b). *Opportunities for public libraries and Wikipedia*. Federación Internacional de Bibliotecas y Bibliotecarios.

<https://www.ifla.org/files/assets/hq/topics/info-society/iflawikipediaandpubliclibraries.pdf>.

Klein, Max (2012). "VIAFbot debriefing. *Hanging together, the OCLC research blog*, 28 November.

<http://hangingtogether.org/?p=2306>

Stinson, Alex, Fauconnier, Sandra; Anäs, Susanna; Wyatt, Liam; Darnell, Jane (2016). "Why you should be paying attention to Wikidata and GLAM". *Wikimedia blog*, 23 August.

<https://blog.wikimedia.org/2016/08/23/wikidata-glam>

Williams, Kathleen (2019). *DPLA cultural artifacts coming to Wikipedia through new collaboration with Wikimedia Foundation*. DPLA. Digital Public Library of America, 10 December.

<https://dp.la/news/dpla-cultural-artifacts-coming-to-wikipedia-through-new-collaboration-with-wikimedia-foundation>

Anexo: Propiedades de Wikidata relacionadas con el control de autoridades o identificadores de bibliotecas, archivos, museos y otras instituciones en España

La tabla que sigue es resultado literal de una pregunta realizada sobre *Wikidata Query Service* (<https://w.wiki/GsC>).

Propiedad	Nombre de la propiedad	Descripción	Ítems
P950	identificador BNE https://www.wikidata.org/wiki/Property:P950	Identificador de la <i>Biblioteca Nacional de España</i>	159.064
P1273	identificador Cantic https://www.wikidata.org/wiki/Property:P1273	Catálogo de autoridades de nombres y títulos de Cataluña, gestionado por la <i>Biblioteca Nacional de Cataluña</i>	90.142
P480	identificador FilmAffinity https://www.wikidata.org/wiki/Property:P480	Número de identificación <i>FilmAffinity</i>	78.508
P4459	identificador del <i>Diccionario Biográfico Español</i> https://www.wikidata.org/wiki/Property:P4459	Identificador de una persona con artículo en el <i>Diccionario biográfico español</i> (números y letras)	40.572
P1600	Código del <i>Inventario del Patrimonio Arquitectónico de Cataluña</i> https://www.wikidata.org/wiki/Property:P1600		27.945
P808	Código de Bien de Interés Cultural https://www.wikidata.org/wiki/Property:P808	Identificador de un elemento en la base de datos de bienes inmuebles del <i>Registro de Bienes de Interés Cultural del Ministerio de Cultura de España</i>	27.484
P1580	Identificador <i>Universidad de Barcelona</i> https://www.wikidata.org/wiki/Property:P1580	Identificador externo	18.908
P2473	Código IGPCV https://www.wikidata.org/wiki/Property:P2473	código de identificación de un bien cultural en el <i>Inventario General del Patrimonio Cultural Valenciano</i>	13.014
P2799	Identificador BVMC de autor https://www.wikidata.org/wiki/Property:P2799	Identificador de un autor en la <i>Biblioteca Virtual Miguel de Cervantes</i>	11.814
P1807	Identificador <i>Gran Enciclopedia Aragonesa</i> https://www.wikidata.org/wiki/Property:P1807		10.100
P772	Código INE https://www.wikidata.org/wiki/Property:P772	código asignado a las entidades de población por el <i>Instituto Nacional de Estadística</i> de España	9.332
P3218	Identificador <i>Auñamendi</i> https://www.wikidata.org/wiki/Property:P3218	Identificador de un elemento en la <i>Enciclopedia Auñamendi</i> (solo números)	7.201
P3427	Identificador As https://www.wikidata.org/wiki/Property:P3427	Identificador de un deportista en <i>as.com</i>	7.146
P5321	Identificador <i>Museo del Prado</i> de artista https://www.wikidata.org/wiki/Property:P5321	Identificador de un artista en el sitio web del <i>Museo del Prado</i>	4.382
P6222	Identificador DIR3 https://www.wikidata.org/wiki/Property:P6222	Identificador oficial de organismo público español	3.526

P7388	Identificador de la <i>Gran Enciclopedia de Navarra</i> https://www.wikidata.org/wiki/Property:P7388	Identificador de un elemento en la <i>Gran Enciclopedia de Navarra</i>	2.726
P7827	Artículo en <i>Vikidia</i> en español https://www.wikidata.org/wiki/Property:P7827	nombre del artículo en <i>Vikidia</i> en español	2.310
P1586	Código de <i>Bien Cultural de Interés Nacional</i> https://www.wikidata.org/wiki/Property:P1586		2.155
P5513	P5513 https://www.wikidata.org/wiki/Property:P5513		2.076
P6175	P6175 https://www.wikidata.org/wiki/Property:P6175		2.002
P3525	Identificador <i>acb.com</i> https://www.wikidata.org/wiki/Property:P3525	Identificador de un jugador de baloncesto en <i>acb.com</i>	1.973
P5496	Identificador de <i>Médicos históricos</i> https://www.wikidata.org/wiki/Property:P5496	Identificador de un médico en la base de datos <i>Médicos Históricos de la Universidad Complutense de Madrid</i>	1.904
P3177	Identificador <i>Patrimonio Web JCyL</i> https://www.wikidata.org/wiki/Property:P3177	Identificador de un ítem en <i>Patrimonio Web de la Junta de Castilla y León</i>	1.782
P5128	Identificador <i>ICAA</i> de catálogo de películas https://www.wikidata.org/wiki/Property:P5128	Identificador de una película en el catálogo de películas del <i>Instituto de Cinematografía y de las Artes Audiovisuales</i>	1.496
P2498	<i>Diccionario Biográfico de Mujeres</i> https://www.wikidata.org/wiki/Property:P2498		1.310
P4813	Identificador <i>Pares</i> https://www.wikidata.org/wiki/Property:P4813	Autoridad en el <i>Portal de Archivos Españoles</i>	1.201
P3580	Código <i>SIPCA</i> https://www.wikidata.org/wiki/Property:P3580	Identificador de un ítem en el <i>Sistema de Información del Patrimonio Cultural Aragonés (SIPCA)</i>	1.165
P4439	Identificador <i>MNCARS</i> de artista https://www.wikidata.org/wiki/Property:P4439	Identificador de artista en el <i>Museo Nacional Centro de Arte Reina Sofía (MNCARS)</i>	1.151
P3318	Identificador <i>Patrimonio Inmueble de Andalucía</i> https://www.wikidata.org/wiki/Property:P3318	Identificador de un bien cultural en la base de datos <i>Patrimonio Inmueble de Andalucía (BDI)</i>	1.080
P6539	Identificador <i>Catálogo de Patrimonio Cultural de Castilla-La Mancha</i> https://www.wikidata.org/wiki/Property:P6539	Identificador de un bien cultural en el <i>Catálogo de Patrimonio Cultural de Castilla-La Mancha</i>	1.044
P2917	Identificador <i>COAM</i> inmueble https://www.wikidata.org/wiki/Property:P2917	Identificador de un edificio o estructura destacados en la ciudad de Madrid en la base de datos del <i>Colegio Oficial de Arquitectos de Madrid (COAM)</i>	1.010
P4335	código territorial <i>Idescat</i> https://www.wikidata.org/wiki/Property:P4335	Identificador del <i>Instituto Estadístico de Cataluña (Idescat)</i> de una unidad territorial catalana	992
P6173	P6173 https://www.wikidata.org/wiki/Property:P6173		982
P5957	Identificador del <i>Inventario del Patrimonio Arqueológico y Paleontológico de Cataluña</i> https://www.wikidata.org/wiki/Property:P5957		918
P4868	Código <i>Lista Roja Hispania Nostra</i> https://www.wikidata.org/wiki/Property:P4868	Identificador para un bien del patrimonio cultural amenazado (a veces un bien natural) tal como aparece en la <i>Lista roja</i> de patrimonio en peligro de <i>Hispania Nostra</i>	905
P6907	<i>BVLarramendi ID</i> https://www.wikidata.org/wiki/Property:P6907	Identificador de una persona en la <i>Biblioteca Virtual de Polígrafos</i>	855
P1610	Identificador <i>Dialnet</i> de artículo https://www.wikidata.org/wiki/Property:P1610	Identificador de artículo en <i>Dialnet</i>	787
P3976	Identificador <i>BVMC</i> de obra https://www.wikidata.org/wiki/Property:P3976	Identificador de una obra en la <i>Biblioteca Virtual Miguel de Cervantes</i>	745
P2985	Identificador <i>DBSE</i> https://www.wikidata.org/wiki/Property:P2985	Identificador de una persona en el <i>Diccionario Biográfico del Socialismo Español</i>	692
P4488	Identificador <i>COAM</i> persona https://www.wikidata.org/wiki/Property:P4488	Identificador de un individuo que ha construido o proyectado edificios o estructuras significativas en la ciudad de Madrid disponibles en la base de datos del <i>Colegio Oficial de Arquitectos de Madrid (COAM)</i>	627
P5498	P5498 https://www.wikidata.org/wiki/Property:P5498		619
P2431	Identificador <i>Thyssen-Bornemisza</i> de artista https://www.wikidata.org/wiki/Property:P2431		577

P3964	Identificador <i>BDCYL</i> de autoridad https://www.wikidata.org/wiki/Property:P3964	Identificador de un autor, tema o lugar en la <i>Biblioteca Digital de Castilla y León</i>	551
P3436	Identificador <i>AELG</i> https://www.wikidata.org/wiki/Property:P3436	Identificador de un autor en la página web de la <i>Asociación de Escritores en Lengua Gallega</i>	523
P4956	Identificador <i>BDH</i> de edición https://www.wikidata.org/wiki/Property:P4956	Identificador de una edición (no obra) en la <i>Biblioteca Digital Hispánica</i> , incluyendo el prefijo «bdh»	496
P3178	Identificador de monumento de Zaragoza https://www.wikidata.org/wiki/Property:P3178	Identificador de un monumento catalogado en <i>zaragoza.es</i>	478
P3945	Identificador <i>RANM</i> https://www.wikidata.org/wiki/Property:P3945	Identificador de un miembro de la <i>Real Academia Nacional de Medicina (RANM)</i> de España	459
P4802	Identificador <i>BVPB</i> autoridad https://www.wikidata.org/wiki/Property:P4802	Identificador de un autor, tema o lugar en la <i>Biblioteca Virtual del Patrimonio Bibliográfico</i>	316
P3758	Identificador <i>Docomomo Ibérico</i> https://www.wikidata.org/wiki/Property:P3758	Identificador de un edificio, estructura o conjunto de edificios en la base de datos <i>Registros del Movimiento Moderno</i> , de <i>DOCOMOMO Ibérico</i>	302
P7006	Identificador <i>BIC</i> de Canarias https://www.wikidata.org/wiki/Property:P7006	Identificador de un <i>Bien de Interés Cultural</i> en el catálogo del Gobierno de Canarias	290
P4797	ID diputado en la Asamblea de Madrid https://www.wikidata.org/wiki/Property:P4797		274
P2768	Identificador <i>BNE</i> de publicación periódica https://www.wikidata.org/wiki/Property:P2768	Identificador de un periódico o revista en la Hemeroteca Digital de la Biblioteca Nacional de España	270
P4568	Identificador <i>SNCZI-IPE</i> de embalse https://www.wikidata.org/wiki/Property:P4568	Identificador de un embalse en España, en <i>SNCZI-Inventario de Presas y Embalses</i>	256
P5468	P5468 https://www.wikidata.org/wiki/Property:P5468		254
P4256	Identificador del <i>Boletín Oficial del Estado</i> https://www.wikidata.org/wiki/Property:P4256	Identificador de un decreto, ley, orden o en general cualquier anuncio oficial en el <i>Boletín Oficial del Estado</i> , publicación oficial del Gobierno de España	248
P5763	Identificador del <i>Directorio de Museos y Colecciones de España</i> https://www.wikidata.org/wiki/Property:P5763	Identificador de un museo en el sitio web del <i>Directorio de Museos y Colecciones de España</i>	226
P2961	Identificador <i>BVPH</i> de publicación https://www.wikidata.org/wiki/Property:P2961	Identificador de publicaciones en la <i>Biblioteca Virtual de Prensa Histórica</i>	223
P5946	Código <i>SUCA</i> https://www.wikidata.org/wiki/Property:P5946	Identificador de una {{Q183620}} urbana específica en la Comunidad de Madrid ({{Q15756}}) suministrado por el conjunto de catálogos {{Q152637792}} de direcciones postales del {{Q152775017}}	222
P3307	Identificador <i>Galiciana</i> de autor https://www.wikidata.org/wiki/Property:P3307	Identificador de un autor en <i>Galiciana</i>	211
P3004	Identificador <i>Galiciana</i> de obra https://www.wikidata.org/wiki/Property:P3004	Identificador de una obra en <i>Galiciana</i>	207
P4593	<i>CPE</i> atleta ID https://www.wikidata.org/wiki/Property:P4593	Perfil de una persona en la web del <i>Comité Paralímpico Español</i>	188
P6164	Identificador <i>Depósito legal</i> https://www.wikidata.org/wiki/Property:P6164	Identificador para una publicación en España	154
P6489	<i>Joan Miró Online Image Bank</i> ID https://www.wikidata.org/wiki/Property:P6489	Identificador de una obra de Joan Miró en <i>Joan Miró Online Image Bank</i>	131
P5562	Identificador <i>Tebeosfera</i> https://www.wikidata.org/wiki/Property:P5562	Identificador de un autor en la página web <i>Tebeosfera</i>	112
P4558	Identificador <i>SNCZI-IPE</i> de presa https://www.wikidata.org/wiki/Property:P4558	Identificador de una presa en España, en <i>SNCZI-Inventario de Presas y Embalses</i>	106
P6297	P6297 https://www.wikidata.org/wiki/Property:P6297		71
P6595	P6595 https://www.wikidata.org/wiki/Property:P6595		65
P6490	Identificador <i>Bidicam</i> de autoridad https://www.wikidata.org/wiki/Property:P6490	Identificador de autoridades en la <i>Biblioteca Digital de Castilla-La Mancha</i>	52
P6493	Identificador <i>CCBAE</i> de publicación https://www.wikidata.org/wiki/Property:P6493	Identificador de autoridades en el <i>Catálogo Colectivo de la Red de Bibliotecas de los Archivos Estatales</i> y del <i>CIDA</i>	51

P6496	Identificador <i>Biblioteca Virtual Andalucía</i> de autoridad https://www.wikidata.org/wiki/Property:P6496	Identificador de autoridades en la <i>Biblioteca Virtual de Andalucía</i>	49
P3932	Identificador <i>Biblioteca Valenciana Digital</i> de autor https://www.wikidata.org/wiki/Property:P3932	Identificador de un autor en la <i>Biblioteca Valenciana Digital (BiValDi)</i>	40
P6130	P6130 https://www.wikidata.org/wiki/Property:P6130		38
P6239	Identificador en la base de datos del <i>IEC</i> de monumentos conmemorativos https://www.wikidata.org/wiki/Property:P6239	Identificador de un monumento conmemorativo o estructura destacados en Cataluña en la base de datos del {{Q1756890}} (IEC)	38
P2897	Identificador <i>Eldoblaje</i> de película https://www.wikidata.org/wiki/Property:P2897	Identificador de doblaje en la web <i>eldoblaje.com</i> , una base de datos de doblajes españoles	28
P7229	P7229 https://www.wikidata.org/wiki/Property:P7229		26
P920	<i>LEmb</i> https://www.wikidata.org/wiki/Property:P920	Encabezamientos de materia mantenido por el <i>Ministerio de Educación, Cultura y Deporte de España</i>	23
P5295	Identificador <i>Lafonoteca</i> de un artista https://www.wikidata.org/wiki/Property:P5295		21
P1022	<i>Código Nacional de Ocupaciones 2011</i> https://www.wikidata.org/wiki/Property:P1022	Clasificación española de ocupaciones mantenida por el <i>Instituto Nacional de Estadística de España</i>	13
P6514	Identificador <i>DOGC</i> https://www.wikidata.org/wiki/Property:P6514	Identificador de un decreto, ley, orden o en general cualquier anuncio oficial en el <i>Diari Oficial de la Generalitat de Catalunya</i>	12
P5569	Identificador de personaje de <i>Tebeosfera</i> https://www.wikidata.org/wiki/Property:P5569	Identificador de un personaje en la página web <i>Tebeosfera</i>	4

Te esperamos en



www.sedic.es

c/Rodríguez San Pedro 2,
oficina 606. 28015 Madrid

Tfno: +34 915 934 059

secretaria@sedic.es

Sociedad
Española de
Documentación e
Información
Científica



 <https://twitter.com/SEDIC20>

 <https://www.facebook.com/AsociacionSEDIC>

 <https://www.linkedin.com/groups?home=&gid=5060038>