

RESEÑAS DE INFORMES: Tecnologías de información, normativa y gestión de la información

Objetos culturales y metadatos: hacia la liberación de datos en Wikidata

Jesús Tramullas

Universidad de Zaragoza
Departamento de Ciencias de la Documentación
tramullas@unizar.es

Tramullas, Jesús (2017). "Objetos culturales y metadatos: hacia la liberación de datos en Wikidata". *Anuario ThinkEPI*, v. 11, pp. 319-321.

<https://doi.org/10.3145/thinkepi.2017.59>

Publicado en *Blok de BiD* el 30 de marzo de 2016



Generalmente, los proyectos de desarrollo GLAM (*galleries, libraries, archives & museums*) se asocian a la liberación de materiales textuales, gráficos o sonoros, que se ponen a disposición de todos los posibles usuarios, bajo licencias libres, sin que existan cortapisas para su reaprovechamiento y

uso, sea comercial o no. Nota para incautos: poner en la web reproducciones de alta calidad que sólo se pueden usar para disfrute privado e individual no es un proyecto GLAM de liberación y uso de contenidos culturales; puede ser estéticamente agradable, pero resulta baladí (Tramullas, 2015).

En los próximos años se va a apreciar un creciente número de proyectos que van a acompañar a la liberación de contenidos con la liberación de los datos semánticos correspondientes. Y es que las obras culturales, en el entorno digital, no son nada si no van acompañadas de su correspondiente conjunto de metadatos. *Europeana* lo tuvo claro desde sus mismos inicios: el potencial de uso y reaprovechamiento del patrimonio cultural debía basarse en la publicación y reutilización de los datos semánticos de los objetos, y de los esquemas de datos necesarios para su interpretación. <http://labs.europeana.eu/api/linked-open-data-introduction>

Wikimedia ha desarrollado desde 2012 *Wikidata*, "una base de datos secundaria libre, cola-

borativa y multilingüe, que recopila datos estructurados para dar soporte a *Wikipedia*...".

<https://www.wikidata.org/wiki/Wikidata:Introduction/les>

El uso de *Wikidata* crece exponencialmente: es el origen de los grafos que *Google* presenta como fichas destacadas en sus respuestas, y de la mayoría de las *infoboxes* de los artículos de *Wikipedia*. Por definición, *Wikidata* trabaja con *open data*, lo que hace posible la reutilización de datos sin limitaciones.

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

Al hilo de todo esto, los museos flamencos han sido de los primeros en iniciar un proyecto de integración y aportación de los datos de unos 25.000 objetos culturales de sus colecciones a *Wikidata*. El proyecto *Flemish art collections, Wikidata and Linked Open Data*, iniciado en 2015, es coordinado por el *Packed vzw (Expertisecentrum Digitaal Erfgoed)*, y en el mismo participan cuatro *wikimedistas* expertos, seis museos, y dos colecciones virtuales.

<https://goo.gl/Nk9Ziv>

"Wikidata es una base de datos secundaria libre, colaborativa y multilingüe, que recopila datos estructurados para dar soporte a Wikipedia"

Como uno de los resultados entregables del proyecto, los participantes publicaron en octubre de 2015 un *whitepaper*, que es el objeto de estos párrafos.

<https://goo.gl/o7jovQ>

El documento se estructura en cuatro partes principales y un anexo. La primera parte se dedica a explicar, a nivel introductorio, los conceptos de *Wikimedia*, *Wikipedia* y *Wikidata* (en particular, los dos primeros se confunden a menudo). La parte dedicada a *Wikidata* es la más extensa y, realmente, resulta ser una concisa y necesaria explicación de su contenido. Tras explicar lo que es un ítem y sus componentes, entra en materia sobre las características que deben tener los datos. Éstos deben ser:

- libres, y más concretamente con licencia Crea-

- *tive Commons Zero*, para facilitar su uso, reutilización y manipulación sin trabas legales;
- significativos, por ejemplo al relacionarse con entidades con suficiente interés para tener entrada en los proyectos de Wikimedia;
- referenciados, permitiendo la verificación de datos y publicación de múltiples valores de acuerdo a diferentes fuentes;
- consultables, aplicando herramientas para llevar a cabo los procesos de publicación y mantenimiento tan automáticamente como sea posible;
- editables, como el resto del contenido de Wikimedia, lo que implica que deben estar abiertos para integrar mejoras;
- con mantenimiento sostenido, en lugar de descargados una vez y olvidados (Wikidata busca una relación a largo plazo).

Cada una de estas cuestiones es explicada con mayor detalle en el propio documento. Una vez acabada esta sección, la segunda parte del documento desgrana los beneficios y los costes que estos proyectos tienen para las instituciones participantes. En particular, se debe llamar la atención sobre el apartado dedicado a la renuncia a la exclusividad de los datos, donde se cita:

“Los museos renuncian explícitamente a cualquier forma de exclusividad sobre los datos que publican en la plataforma Wikidata. Una vez que los datos se han publicado bajo una licencia Creative Commons Zero, esta es

irrevocable. A través de Wikidata, los museos donan datos a la sociedad. Al hacerlo, descartan cualquier posible modelo de ganancia o beneficio por la puesta a disposición para su reutilización por terceros. Específicamente significa por ejemplo que un museo no puede obtener ingresos por la venta de licencias sobre estos datos” (traducción del equipo editorial).

“A través de Wikidata, los museos donan datos a la sociedad”

No deja de resultar curioso, a la par que ridículo, que pueda haber instituciones que se sientan tentadas de hacer negocio con algo como “Autor=Goya, Título=La familia de Carlos IV”, máxime cuando son datos públicos, grabados y almacenados merced al uso de dinero público. El análisis muestra que los beneficios para la sociedad son muchos, mientras que el coste es, precisamente, el uso de fondos públicos para producir los datos; un coste que, además, es técnicamente mínimo.

Es importante llamar la atención sobre la licencia elegida: *CC0 1. Universal Public Domain Dedication*.
<https://creativecommons.org/publicdomain/zero/1.0>

The screenshot shows the Wikidata page for the whitepaper. The main content area contains the following text:

Other languages: العربية • مصرى • čeština • English • British English • suomi • français • italiano • Nederlands • українська

The whitepaper below was written in October 2015, for the project *Linked Open Data publication with Wikidata*. In this project, several Flemish museums contribute collection data to Wikidata. This whitepaper may be interesting for other parties as well – especially other cultural institutions who consider contributing data to Wikidata.

Feedback is welcome. Questions and remarks can be placed on the talk page.

This paper was written by Sandra Fauconnier (User:Spinster), Bert Lemmens & Barbara Dierickx (PACKED vzw). It is published under a Creative Commons Attribution Share-Alike license.

Editors' notes

This whitepaper is the first project deliverable of the project *Linked Open Data publication with Wikidata (D1. Whitepaper open data management in Wikidata)* and was described in the project plan as follows:

PACKED vzw and Wikimedia develop a shared vision on how data managers in museums may publish collection data on Wikidata, and how they can update this collection data at regular intervals. This includes:

- how collection data is modeled in Wikidata;
- how the upload and export to Wikidata works.

This vision is recorded in a whitepaper that is proposed to the steering committee of this project. The aim of the whitepaper is to match the objectives of Wikimedia and museums, and Wikidata.

The thumbnail image is titled "The adoration of the Magi (Q5861089) by Peter Paul Rubens (Q5599), collection Royal Museum of Fine Arts Antwerp (KMSKA) (Q1471477)".

https://www.wikidata.org/wiki/Wikidata:Flemish_art_collections,_Wikidata_and_Linked_Open_Data/Whitepaper

Se trata de una licencia apropiada para los productos culturales libres, en todos los sentidos, y como tal es señalada por *Creative Commons*: "La persona que ha asociado una obra a esta declaración, la entrega al dominio público..." (traducción del equipo editorial). Y es que, a veces, se olvida que no todas las licencias CC son libres ni suponen dominio público...

La tercera parte propone un ejemplo de estructura de datos para obras culturales, indicando metadatos obligatorios y opcionales, el elemento correspondiente dentro del esquema general de *Wikidata*, y unas indicaciones para cada elemento. Incluye, además, una orientación sobre los formatos de fichero que deben usarse para la importación y carga de datos.

La conclusión, última parte del documento, es un breve (demasiado breve) resumen de posición.

Un último valor añadido es que el documento incorpora como anexo informativo las referencias a las guías de uso de datos abiertos que aplican tres instituciones de gran prestigio, como *The Museum of Modern Art (MoMA)*, *The Tate Gallery*, y *The Smithsonian Cooper-Hewitt, National Design Museum*. En su conjunto, un documento que sirve como punto de partida para plantear proyectos de *cultural linked open data*, y para detenerse a pensar si realmente merece la pena afrontar el descrédito y rechazo que supone la mercantilización del patrimonio cultural.

Tramullas, Jesús (2015). "Retos y oportunidades para las GLAM de Australia". *Anuario ThinkEPI*, v. 10, pp. 123-124.
<https://doi.org/10.3145/thinkepi.2016.24>

Fauconnier, Sandra; Lemmens, Bert; Direickx, Barbara (2015). *Wikidata: Flemish art collections, Wikidata and Linked Open Data/Whitepaper*.
https://www.wikidata.org/wiki/Wikidata:Flemish_art_collections,_Wikidata_and_Linked_Open_Data/Whitepaper

Nota original en *Blok de BiD*:
<http://www.ub.edu/blokdebid/es/node/689>

* * *

El estado de los datos vinculados en bibliotecas en 2015

Xavier Agenjo-Bullón

Fundación Ignacio Larramendi
xavier.agenjo@larramendi.es

Francisca Hernández-Carrascal

Digibis
francisca.hernandez@digibis.com

Agenjo-Bullón, Xavier; Hernández-Carrascal, Francisca (2017). "El estado de los datos vinculados en bibliotecas en 2015". *Anuario ThinkEPI*, v. 11, pp. 321-323.

<https://doi.org/10.3145/thinkepi.2017.60>

Publicado en *Blok de BiD* el 15 de junio de 2016



Los cuatro capítulos que componen este informe tienen como finalidad actualizar el anterior análisis de Mitchell sobre *linked data*, publicado también en *Library technology reports (LTR)* en 2013.

Las revisiones sobre el estado de la cuestión ofrecen una oportunidad de reflexionar sobre el camino andado, sobre las tendencias que van perfilándose y sobre los interrogantes que surgen de la aplicación de *library linked data*. Este es el ejercicio que realiza Erik Mitchell en esta primera entrega del año 2016 de *Library technology reports* y, en nuestra opinión, su aportación más importante.

En el capítulo 1, *The current state of linked data in libraries, archives, and museums*, revisa el estado de la adopción de *linked data* en bibliotecas, archivos y museos, y en los objetivos y finalidades que estas instituciones aducen para implantar esta tecnología.

En el capítulo 2, *Projects, programs, and research initiatives*, Mitchell reúne los proyectos, programas e iniciativas de investigación que considera más significativos por su tamaño, sin pretensiones de ser exhaustivo, centrándose más en los aspectos de su definición general y en las características de su adopción que en los aspectos de diseño técnico o de modelo de datos.