

---

# Los datos bibliotecarios en la nube de datos: análisis de los datasets GLAM presentes en el LOD cloud diagram

---

*Library data in the data cloud: an analysis of the GLAM datasets in the LOD cloud diagram*

---

Ana B. RÍOS HILARIO (1), Blanca GIL URDICIAIN (2)

(1) Departamento de Biblioteconomía y Documentación, Universidad de Salamanca, anarihi@usal.es

(2) Departamento de Biblioteconomía y Documentación, Universidad Complutense de Madrid, mbgil@pdi.ucm.es

## Resumen

Se analiza la representación de los conjuntos de datos pertenecientes a galerías, bibliotecas, archivos y museos (GLAM) en la nube de los datos abiertos y vinculados. En primer lugar, se establecen las bases teóricas en las que se sustenta el análisis: por un lado, se define el concepto de linked open data (LOD); posteriormente, se delimitan las fuentes objeto de estudio, en concreto: el diagrama de la nube LOD en el que se exponen la totalidad de los conjuntos de datos; el catálogo de datos enlazados, en donde se obtiene información individualizada de cada caso específico y el informe titulado conjunto de datos, vocabularios de valores y conjuntos de elementos de metadatos, cuya tipología nos sirve para categorizar los datasets bibliotecarios. El análisis propiamente dicho se centra: primero, en el estudio de la muestra de los datos examinados sobre la totalidad de la nube; segundo, en la categorización de dichos conjuntos en función de la clasificación previamente establecida; finalmente, se describen los casos más representativos centrándose en la tecnología de datos vinculados empleada. Se concluye exponiendo los principales resultados obtenidos y estableciendo una serie de propuestas de mejora.

**Palabras clave:** Datos abiertos y vinculados (LOD). Galerías, bibliotecas, archivos y museos (GLAM). Diagrama de la nube LOD. Catálogo de datos enlazados Mannheim. Conjunto de datos, vocabularios de valores y conjuntos de elementos de metadatos (VO-CABDATASET).

## 1. Introducción

La web semántica permite pasar de una web de documentos a una web de datos en la cual unos datos se conectan o enlazan con otros. Se abre así la posibilidad de enlazar conjuntos de datos (*datasets*) con otros conjuntos y en última instancia datos con datos de acuerdo con una serie de principios y modelos de interrogación bien establecidos (*Linked data*, 2014). Estamos por lo tanto ante un nuevo fenómeno, el que hace referencia a los datos abiertos y vinculados,

## Abstract

The representation of the data sets belonging to galleries, libraries, archives and museums (GLAM) cloud of linked open data is analyzed. First, theoretical basis on which the analysis is based are set: on one hand, the concept of open data linked (LOD) is defined; subsequently object of study sources are delimited, namely: the cloud diagram LOD in which all datasets are displayed; Mannheim Linked Data Catalog where individual information of each specific case is obtained and the report entitled Datasets, Value Vocabularies, and Metadata Element Set whose typology serves to categorize librarians datasets. The analysis itself focuses: first study sample of data examined of the entire cloud; second, in the categorization of such set according to the predetermined classification; finally, the most representative cases focusing on technology linked data used are described. It is concluded stating the main results and establishing a series of proposals for improvement.

**Keywords:** Linked Open Data (LOD). Galleries, libraries, archives and museum (GLAM). Cloud diagram LOD. Mannheim Linked Data Catalog. Datasets, Value Vocabularies, and Metadata Element Set (VO-CABDATASET).

más conocido por su acrónimo inglés LOD (*linked open data*).

Desde hace más de 20 años, en el entorno de las instituciones de memoria o de las entidades también conocidas con las siglas de LAM o GLAM (*galleries, libraries, archives and museum*) se están produciendo toda una serie de cambios, a menudo vertiginosos que han tenido como consecuencia un replanteamiento total en los modos de gestión y funcionamiento de dichas instituciones.

Centrándonos en el caso específico de las bibliotecas hemos contemplado como a finales de los 80 principios de los 90 se producía de modo masivo la automatización de dichos centros. El fin de siglo trajo consigo la aparición de nuevos modelos conceptuales, como el FRBR (*Functional Requirements for Bibliographic Records*), y la explosión del término metadato para referirse de modo específico a la descripción de los documentos digitales. En el nuevo milenio aparecen nuevos conceptos asociados a esa nueva tipología documental como son biblioteca digital, virtual o el más “novedoso y determinante” que es el término repositorio. Siguiendo con este rápido recorrido, llegamos al comienzo de la nueva década hablando de datos abiertos y enlazados.

Tomando como base las anteriores premisas el objetivo principal de este artículo se centra en estudiar los conjuntos de datos abiertos y vinculados pertenecientes al ámbito bibliotecario en el sentido más amplio del término. De este propósito general se establecen a su vez los siguientes objetivos específicos:

- Definir los conceptos claves objeto de estudio en lo referente a los datos abiertos y vinculados y delimitar las fuentes de información empleadas específicamente para la realización del análisis.
- Establecer una categorización de los distintos *datasets* GLAM presentes en el *LOD cloud diagram* (diagrama de la nube LOD) siguiendo la tipología establecida en el documento *Datasets, Value Vocabularies, and Metadata Element Sets* (*Conjunto de datos, vocabularios de valores y conjuntos de elementos de metadatos*; VOCABDATASET) realizado por el World Wide Web Consortium (W3C).
- Realizar un análisis estadístico que posibilite conocer la representación de los conjuntos bibliotecarios dentro de la nube en general y de modo más específico la representación de las categorías establecidas.
- Detallar los casos más representativos de cada categoría expuestos en el *Mannheim Linked Data Catalog* (*Catálogo de datos enlazados Mannheim*) centrándonos especialmente en la tecnología de datos vinculados empleada.

Dejando constancia de la imposibilidad de recoger y analizar la totalidad de *datasets* del ámbito GLAM publicados como datos vinculados en la actualidad, decidimos tomar como fuente de información principal el *LOD cloud diagram* debido a su representatividad y vigencia. El último

rastreo de la web *linked data* se efectuó en abril de 2014 y su difusión se hizo pública en agosto de ese mismo año. Para el estudio y comprensión de la “nube” nos ha sido de mucha utilidad la consulta del documento titulado *Adoption of the linked data best practices in different topical domains* (Schmachtenberg, Bizer y Paulheim, 2014a) y su versión abreviada disponible en el recurso *State of the LOD cloud 2014* (Schmachtenberg, Bizer y Paulheim, 2014b). Así mismo, para el análisis específico de los *datasets* acudimos al *Mannheim Linked Data Catalog* (2014). Otro instrumento clave para el desarrollo del estudio ha sido el *Datasets, Value Vocabularies, and Metadata Element Sets* (Isaac; Waites; Young; Zeng, 2011), fruto del informe final realizado por el *Library Linked Data Incubator Group* (*Grupo Incubador de Datos Vinculados de Bibliotecas*).

Para cumplir con los objetivos propuestos hemos aplicado una triple metodología. En primer lugar, para sentar las bases teóricas se acudió a las propias fuentes anteriormente indicadas. Para la cuantificación de los *datasets* se llevó a cabo un análisis estadístico; y, finalmente, se aplicó el método descriptivo para definir los casos más representativos.

El artículo se estructura en dos partes. En la primera se define el concepto de *linked open data* y se explican las fuentes de información en las que se base el análisis. En la segunda se presenta el análisis propiamente dicho: tras la categorización y contabilidad de los conjuntos de datos bibliotecarios se muestran los casos más relevantes. Finalmente, en las conclusiones se exponen los principales resultados obtenidos y se realizan una serie de propuestas de mejora. El artículo incluye un apéndice final en el que figuran categorizados según las clases establecidas todos los conjuntos de datos bibliotecarios.

## 2. Conceptos claves entorno a la definición de linked open data (LOD)

Antes de pasar a realizar el estudio específico y objetivo general de este artículo, previamente debemos definir y delimitar una serie de conceptos claves sin los cuales sería difícil comprender el análisis propiamente dicho.

### 2.1. Linked open data (LOD)

En los últimos años, la palabra “data” aparece con frecuencia en artículos científicos y de divulgación, generalmente acompañado de otro término que delimita su significado como puede ser *open data*, *data mining* y el más reciente, *big data*. Cada vez es más amplio el número de

personas y organizaciones que están contribuyendo a este “diluvio de datos” al optar por compartir su información con los demás (Heath y Bizer, 2011). Sirvan de ejemplo instituciones tan importantes como Amazon, medios de comunicación como la BBC o el New York Times, los organismos públicos como el gobierno de los Estados Unidos y las iniciativas de investigación en diversas disciplinas científicas.

Por supuesto, en esta maraña de datos también están presentes los datos bibliotecarios cuya misión se centra en el intercambio de datos bibliotecarios y en el potencial para la creación de estos datos globalmente interconectados; el intercambio y utilización conjunta de datos con instituciones no bibliotecarias; la creciente confianza en el crecimiento de la web semántica, y en el mantenimiento de una gráfica cultural global de la información que sea fiable y persistente (Bauer y Kaltenböck, 2012). Como ejemplos más representativos de datos bibliotecarios podemos citar la presencia de bibliotecas como la Library of Congress, e iniciativas como Euro-peana y la Digital Public Library of America (DPLA).

El adjetivo “open” asociado a otros términos, ha dado lugar a los denominados “movimientos abiertos”. El citado movimiento está siendo objeto de atención por la comunidad bibliotecaria en los últimos años, ya que reporta numerosos beneficios ante el alto crecimiento de información. El concepto de *open* se incluye dentro de una expresión más amplia que es la de conocimiento abierto. Según el *Open Definition Advisory Council* (2014) este término haría referencia tanto al “contenido incluido en música, películas y libros; los datos de carácter científico, histórico, geográfico; o cualquier otro tipo información gubernamental y de otras administraciones públicas”.

Los datos enlazados es la forma que tiene la web semántica de vincular los distintos datos que están distribuidos en la Web, de tal manera que los datos se enlazan del mismo modo que lo hacen las páginas web. En la *Guía breve de Linked Data* (2015) se especifica que “la web semántica no se trata únicamente de la publicación de datos en la web, sino que éstos se pueden vincular a otros, de forma que las personas y las máquinas puedan explorar la web de los datos, pudiendo llegar a información relacionada a la que se hace referencia desde otros datos iniciales”. Si queremos una definición más precisa del término podemos acudir a la página oficial *linked data* en la que se define este concepto como “la utilización de la Web para conectar los datos relacionados que no estaban vinculados previamente, o el uso de la Web para

disminuir los obstáculos en la conexión de los datos actualmente vinculados mediante otros métodos”.

El término datos vinculados se refiere, por lo tanto, a un conjunto de buenas prácticas para la publicación y la interconexión de datos estructurados en la Web. Para hablar de LD los datos deben publicarse de acuerdo con los principios diseñados para facilitar los vínculos entre los conjuntos de datos, elementos y vocabularios controlados (Berners-Lee, 2006). Estas prácticas fueron presentadas en 2006 por Tim Berners-Lee y se han dado a conocer como los principios de *linked data* (LD). Tales principios son los siguientes:

1. Usar URIs (*uniform resource identifiers*) para identificar los recursos de forma unívoca.
2. Usar URIs http para que la gente pueda acceder a la información del recurso.
3. Ofrecer información sobre los recursos usando RDF.
4. Incluir enlaces a otros URIs, facilitando el vínculo entre los distintos datos distribuidos en la web.

Teniendo en cuenta las anteriores premisas podemos decir que, *linked open data* es un modo de publicar datos estructurados lo que permite que los metadatos se conecten y enriquezcan, de tal modo que diferentes representaciones de un mismo contenido puedan encontrarse y enlazarse entre recursos relacionados (Europeana Linked Open Data, 2015).

Según Mitchell (2013, p. 12) “LOD comprende dos conceptos distintos, el primero es que los datos publicados en la Web deberían conectarse fácilmente con información (“enlazados”) y al hacer esto debería ser accesible tanto para las computadoras como para los humanos (“datos”). El segundo concepto clave es que para que los datos se vinculen y se reutilicen, deben estar abiertos y libres de las restricciones de derechos y copyright (“abiertos”).

Aunque en algunos contextos se suele asimilar ya *linked data* con *linked open data*, no todos los datos enlazados son datos abiertos –y no tienen por qué serlo– sólo lo son si se liberan bajo acuerdos de licencia que permitan su libre reutilización. Por su parte, mientras *linked data* alude a la interoperabilidad técnica de los datos, *open data* hace referencia a su interoperabilidad legal.

## 2.2. La nube de los datos abiertos y enlazados

El *LOD cloud diagram* (*Diagrama de la nube LOD*) expone los conjuntos de datos que se han publicado en el formato de datos vinculados y que son recogidos por los colaboradores del proyecto Linking Open Data y de otra serie de personas y organizaciones (Figura 1). El propósito principal de este proyecto es “extender la Web como un bien común de datos mediante la publicación de varios conjuntos de datos abiertos como RDF en la Web y mediante el establecimiento de enlaces RDF entre elementos de datos de diferentes fuentes” (Linking Open Data, 2014). El diagrama se sustenta por los metadatos recopilados por los colaboradores del *Data Hub*, así como en los metadatos extraídos de un rastreo de la web *linked data* que se realizó en abril de 2014. El *Data Hub* a su vez podemos definirlo como como “un registro de datos en el que se puede compartir información sobre paquetes de datos de cualquier tipo y describirlos de forma colaborativa” (Isaac, Waites, Young y Zeng, 2011).

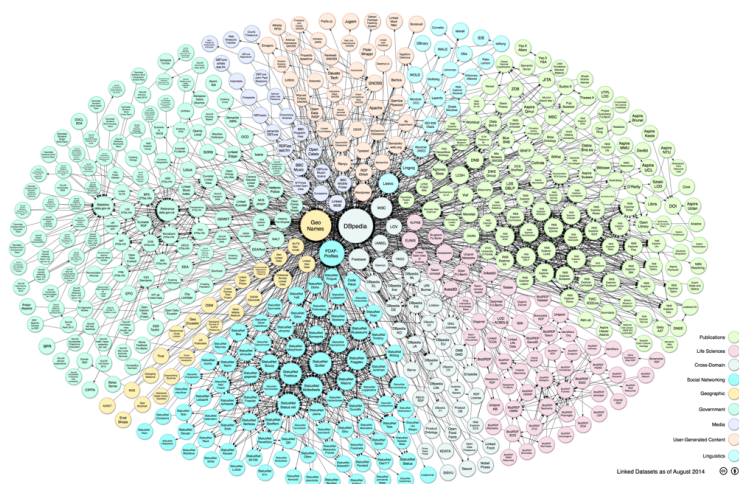


Figura 1. Diagrama de la nube LOD  
Fuente: <http://lod-cloud.net/>

En dicho diagrama figuran dos documentos independientes que presentan información sobre los *datasets* en función del momento en el que se obtiene la imagen.

- El estado de la nube 2011: Proporciona información sobre los conjuntos de datos vinculados que fueron catalogadas en el *Data Hub* en septiembre de 2011.
- El estado de la nube en 2014: Presenta información sobre la estructura y el contenido del subconjunto rastreado de la nube LOD en abril de 2014.

La interpretación del diagrama es sencilla. La figura 1 muestra los conjuntos de datos publicados en formato *linked data* y que se entrelazan con otros conjuntos de datos de la nube. El tamaño de los círculos se corresponde al número de tripletas (sujeto, predicado, objeto) en cada conjunto de datos. Generalmente, los números son proporcionados por los editores de los conjuntos de datos y en ocasiones son cálculos aproximados.

Los conjuntos de datos referentes a la categoría *Publicaciones*, donde se incluyen los datos bibliotecarios en su más amplio sentido, se identifican por el color verde claro y se sitúan en el margen derecho del diagrama.

Las flechas muestran la existencia de al menos 50 enlaces entre dos conjuntos de datos. Un enlace, según el propósito de la nube, es una tripleta RDF donde los URIs del sujeto y objeto están en los espacios de nombres de diferentes conjuntos de datos.

La dirección de las flechas indica el conjunto de datos que contiene los enlaces, por ejemplo, una flecha de A a B significa que el conjunto A contiene tripletas RDF que utilizan los identificadores de B. Las flechas bidireccionales generalmente indican que los enlaces se reflejan en ambos conjuntos de datos. El espesor se corresponde con el número de enlaces.

### 2.2.1. El catálogo de datos enlazados

El *Mannheim Linked Data Catalog* (*Catálogo de datos enlazados Mannheim*) es un instrumento que proporciona información sobre los conjuntos de datos disponibles en la Web en el momento de la realización del último estudio del estado de la nube desarrollado en agosto de 2014. La creación de este catálogo, así como el diagrama de la nube LOD, han sido financiados por el proyecto de la Unión Europea *Planet Data*. El contenido de dicha herramienta se ha generado a partir de dos fuentes:

1. En abril de 2014 se realizó un rastreo de la web de LD y se analizaron los conjuntos de datos vinculados hallados que cumplieran con las prácticas de LD.
2. La comunidad *Linked data* recoge la meta-información de los conjuntos de datos disponibles en el catálogo *datahub.io*.

Dicho catálogo es un complemento al documento *Adoption of the linked data best practices in different topical domains* (Schmachtenberg; Bizer; Paulheim, 2014a) “permitiendo al lector desglosar y explorar la información sobre los conjuntos de datos detrás de cada resultado estadístico” (Schmachtenberg, Bizer, Paulheim, 2014b). Pre-

cisamente, en este trabajo nuestro objetivo es el estudio de los *datasets* bibliotecarios etiquetados bajo la categoría de *Publicaciones* y que figuran en la mencionada herramienta.

### 2.3. Conjuntos de datos, vocabularios de valores y metadatos

El W3C Library Linked Data Incubator Group (Grupo Incubador del W3C de Datos Vinculados de Bibliotecas; LLD-XG) se crea en el año 2010 con el objetivo de "contribuir a incrementar la interoperabilidad global de los datos de las bibliotecas en la Web, reuniendo a personas implicadas en actividades de la web semántica — centradas en los datos vinculados— en bibliotecas e instituciones afines, mediante el examen de las iniciativas en curso e identificando futuras vías de colaboración". Fruto de su trabajo durante un año fue la redacción de un informe final que se publica en octubre de 2011 bajo la denominación de *Library Linked Data Incubator Group Final Report (Informe Final del Grupo Incubador de Datos Vinculados de Bibliotecas)*.

Dicho informe se estructura en cuatro apartados:

1. Ámbito del informe.
2. Beneficios para la comunidad científica y académica, organizaciones, profesionales de la información y desarrolladores y firmas comerciales.
3. Estado de la cuestión en torno a problemas existentes con los datos bibliotecarios tradicionales, descripción de los LLD disponibles en la actualidad y las cuestiones relacionadas con los derechos.
4. Recomendaciones para directores de bibliotecas, organismos de normalización, ingenieros de datos y de sistemas y bibliotecarios y archiveros.

El informe final incluye a su vez dos documentos que lo complementan pero que se publican de modo independiente: *Use Cases (Casos de usos; USECASE)* y *Datasets, Value Vocabularies, and Metadata Element Sets (Conjunto de datos, vocabularios de valores y conjuntos de elementos de metadatos; VOCABDATASET)*.

El conjunto de casos de uso, estudios de casos, iniciativas e ideas se organizaron en el USECASE atendiendo a los ocho grupos que se citan a continuación: datos bibliográficos; datos de autoridad; vocabularios; archivos y datos heterogéneos; citas; objetos digitales; colecciones; y usos sociales

El USECASE presenta los casos de uso recopilados y ofrece un breve resumen de cada caso individual. Con el fin de proporcionar información más detallada para el lector interesado, las secciones de los grupos y los resúmenes de casos están relacionados con las páginas wiki que el Grupo Incubadora utilizó para recoger y conservar la información original.

Sin embargo, para el propósito de nuestro estudio el documento más relevante es el titulado *Conjunto de datos, vocabularios de valores y conjuntos de elementos de metadatos*. Su objetivo es la identificación de "un conjunto de recursos útiles para la creación o consumo de datos enlazados en el campo de las bibliotecas" (Isaac, Waites, Young y Zeng, 2011). La lectura de este informe permitirá, por un lado, "proporcionar a la comunidad de datos enlazados la oportunidad de comprender el punto de vista específico, los recursos y la terminología usada por parte del colectivo bibliotecario para sus datos"; y, al mismo tiempo, "servirá de ayuda a los profesionales de la información para comprender el modo en el que los conceptos relacionados con los datos enlazados encajan con sus propias tradiciones profesionales" (Isaac, Waites, Young y Zeng, 2011). Se concibe como un punto de partida para que los profesionales puedan encontrar, comprender y explorar algunos ejemplos de conjuntos de datos, vocabularios controlados y metadatos. El informe comienza por definir los conceptos claves que dan nombre al propio documento. Posteriormente se recopila un listado de los casos de uso citados en el documento precedente sobre cada uno de ellos.

Ya que estos mismos términos son los que vamos a emplear a la hora de categorizar los *datasets* bibliotecarios recogidos en el catálogo *Mannheim*, a continuación pasaremos a definirlos según se detalla en el propio VOCABDATASET (Isaac, Waites, Young y Zeng, 2011).

- *Conjunto de datos*: considerados como colecciones de metadatos estructurados, descripciones de cosas como los libros de una biblioteca. En el mundo bibliotecario el equivalente a un conjunto de datos es una colección de registros bibliográficos. Los registros bibliográficos son afirmaciones sobre cosas, en los que cada afirmación está formada por un elemento ("atributo" o "propiedad") de la entidad y un "valor" para ese elemento. Por ejemplo, un registro de un conjunto de datos sobre un libro determinado puede tener un elemento 'materia' tomado de *Dublin Core* y un valor para la materia tomado de los *Library of Congress Subject Headings* (LCSH).

- *Vocabulario de valores*: Un vocabulario de valores define recursos (como instancias de materias, estilos artísticos o autores) que se utilizan como valores de elementos en los registros de metadatos. Son componentes estructurales que sirven para llenar de contenido los registros de metadatos. Muchas bibliotecas usan obligatoriamente determinados vocabularios para seleccionar los posibles valores de un elemento de metadatos. Así, un vocabulario de valores es una lista de los valores permitidos para un elemento. Es el caso de los tesauros, listas de códigos, listas de términos, esquemas de clasificación, listas de encabezamientos de materia, taxonomías, ficheros de autoridades, nomencladores geográficos, esquemas de conceptos y otros tipos de sistemas de organización del conocimiento. Por ejemplo, el *Fichero Virtual de Autoridades (Virtual International Authority File; VIAF)* define autoridades de nombres (por ejemplo, Mark Twain).
- *Conjuntos de elementos de metadatos o conjuntos de elementos*: Los conjuntos de elementos de metadatos definen las clases y atributos utilizados para describir entidades de interés. En la terminología de datos vinculados, estos conjuntos de elementos se concretan generalmente por medio de RDF Schemas (esquemas RDF) o Web Ontology Language (Lenguaje de Ontologías Web; OWL), que con frecuencia se agrupan bajo el término "vocabulario RDF". Normalmente los conjuntos de elementos de metadatos no describen entidades bibliográficas, sino más bien proporcionan los elementos que se pueden utilizar para describir estas entidades. Términos equivalentes serían vocabularios RDF, esquemas RDF y ontologías (*Library terminology informally explained*, 2011). Por ejemplo, MARC21 define los elementos (campos) para describir registros bibliográficos y de autoridades.

### 3. Los datos bibliotecarios en el LOD cloud diagram

#### 3.1. Análisis estadístico

Antes de comenzar el estudio pormenorizado de la categoría denominada *Publicaciones* es necesario obtener una visión global de la totalidad de conjuntos de datos presentes en el diagrama. Para ello, se ha tomado la información que se proporciona en el propio documento, en el que se clasifican por categorías temáticas los 1.014 conjuntos de datos que se identificaron en el último rastreo (Tabla I).

Categoría	Conj.	%
Gobierno	183	18,05%
Publicaciones	96	9,47%
Ciencias de la vida	83	8,19%
Contenido generado por el usuario	48	4,73%
Dominios cruzados	41	4,04%
Medios	22	2,17%
Geografía	21	2,07%
Web social	520	51,28%
Total	1.014	100%

Tabla I. Conjuntos de datos por categorías temáticas (Fuente: Mannheim, 2014)

Como puede observarse de las ocho materias identificadas, la categoría *Publicaciones* ocupa la tercera posición, con 96 conjuntos de datos identificados, lo que se corresponde con el 9,47% del espacio en la nube. Destacan por encima del resto, los conjuntos de datos relativos a la web social (520), que ostenta la primera posición.

Centrándonos en el análisis específico de la materia objeto de estudio, en primer lugar, tenemos que comenzar definiendo qué *datasets* se han etiquetado bajo dicha denominación.

Tal y como se indica en el propio documento, la categoría denominada *Publicaciones* comprende "conjuntos de datos de bibliotecas, información sobre publicaciones científicas y conferencias, listas de lectura de las universidades, y las bases de datos de citas".

El número de conjuntos que figura en el resultado final difiere del que se muestra en el catálogo Mannheim. En dicho catálogo aparecen 101 conjuntos de datos etiquetados dentro de la categoría *Publicaciones*. Tal diferencia viene dada porque los conjuntos pueden categorizarse en ocasiones en más de una categoría, pero a la hora de realizar el recuento final se ha optado por incluirlos en una única. De tal manera que del total de casos que figuran en el catálogo, un ejemplo aparece repetido y cuatro son contabilizados además en otra categoría. Es así como Uniprot-citation se incluye en el apartado de Ciencias de la vida, y DWS-Group, Data.dcs y Morelab se recogen en la tipología denominada Web Social.

Si hacemos una clasificación del total de los 96 casos contabilizados atendiendo a las subcate-

gorías establecidas en la anterior definición obtenemos los siguientes resultados (Tabla II).

Subcategoría	Número	Porcentaje
Bibliotecas (GLAM)	42	44%
Publicaciones científicas	23	24%
Listas de lectura	17	18%
Bases de datos	14	14%
Total	96	100%

Tabla II. Subcategorías de la categoría Publicaciones

Como puede observarse el mayor número de casos hace referencia a la subcategoría denominada *Bibliotecas* con un total de 42 ejemplos. Entiéndase el término bibliotecas en su más amplio sentido; es decir, abarcaría a todas las instituciones recogidas bajo las siglas GLAM. En un porcentaje similar se encuentran tanto la subcategoría que recoge las *Publicaciones científicas y conferencias* como las listas de lectura ofertadas por las instituciones académicas de enseñanza superior. La última subcategoría referente a las base de datos de citas presenta 15 casos recogidos en el mencionado catálogo.

Atendiendo a la clasificación que recoge el documento VOCABDATASET —que divide los datos vinculados bibliotecarios entre conjuntos de datos, vocabulario de valores y conjunto de elementos de metadatos—, de los 42 conjuntos de datos obtenidos en la anterior tabla como *datasets* bibliotecarios podemos obtener los datos de la Tabla III. Las colecciones bibliotecarias destacan ligeramente sobre los vocabularios de valores. En un porcentaje menor aparecen los conjuntos de metadatos.

Clases	Número	Porcentaje
Colecciones	20	48%
Vocabularios de valores	16	38%
Metadatos	6	14%
Total	42	100%

Tabla III. Clases según la clasificación establecida por el VOCABDATASET

Haciendo nuevamente un desglose del conjunto de casos que se han agrupado bajo la clase denominada *Colecciones*, aparecen en porcentajes similares los fondos de bibliotecas y archivos, bibliografías nacionales y catálogos (Tabla

IV). Aunque todos y cada uno de los casos referentes a esta categoría hacen referencia a “los conjuntos de datos considerados como colecciones de metadatos estructurados” (VOCABDATASET, 2011) los hemos diferenciado entre las diferentes colecciones que figuran en la tabla III, tan sólo hemos dejado un caso sin catalogar por la dificultad de clasificarlo en dichas categorías. Se trata del *Project Gutenberg* que consta de un conjunto de libros electrónicos gratuitos, tal y como figura en la propia definición del caso y que, aunque similar a los datos de una biblioteca, no está definido como tal.

Subclases	Número	Porcentaje
Bibliotecas	5	25%
Archivos	4	20%
Museos	2	10%
Bibliografías	4	20%
Colecciones	1	5%
Catálogos	4	20%
Total	20	100%

Tabla IV. Subclases para Colecciones

La siguiente clase definida en el documento elaborado por el LLD-XG hace referencia a los vocabularios de valores y vocabularios controlados. En esta ocasión, a la hora de subdividir esta clase, hemos empleado los mismos apartados que figuran recogidos en el propio documento. En este caso, aparecen con el mismo número tanto los sistemas de clasificación como los tesauros. Con dos ejemplos se presentan tanto los encabezamientos de materia como los datos de autoridad (Tabla V).

Subclases	Número	Porcentaje
Sistemas de clasificación	6	38%
Encabezamientos de materia	2	12%
Datos de autoridad	2	12%
Tesauros	6	38%
Total	16	100%

Tabla V. Subclases para Vocabulario de valores

Finalmente, si subdividimos la última clase que se presenta en el VOCABDATASET —denominada ‘conjunto de elementos de metadatos’—, es significativo que la mayoría de los

casos hacen referencias a ontologías. Tan sólo se da un único caso definido como metadatos, que es nuevamente el Project Gutenberg pero en este caso presentado como RDF (Tabla VI).

Subclase	Número	Porcentaje
Ontologías	5	83%
Metadatos	1	17%
Total	42	100%

Tabla VI. Subclases para Metadatos

### 3.2. Descripción de los casos más representativos

En este apartado pasaremos a describir brevemente los casos más representativos de cada una de las clases anteriormente identificadas.

Comenzando por los *datasets* incluidos dentro de la primera clase establecida —la referente al apartado *Colecciones*, y dentro de ella la subclase *Bibliotecas*—, destacan por un lado Europea, y los proyectos de datos enlazados de las bibliotecas nacionales de Francia y España.

the Cittie of London 31

**Descripción:**  
This map has been attributed to Augustus Ryther, an engraver who prospered between 1572 and 1592, contributing to Saxton's Atlas of 1579. This plan was produced to satisfy a European market, and contains certain inaccuracies which a native Londoner would not have tolerated. The streets appear very much wider than they were in actuality. Houses are depicted as having large gardens, when these had, in fact, begun to disappear from London two centuries before. The map details the gap at the north end of London Bridge, caused by a fire in 1632.

**Creador:**  
Cartographer: Ryther, Augustus

**Cobertura geográfica:**  
531500, 181500; London, City of London

**Fecha:**  
1633, Part of: Mid 17th century; From: Sat Jan 01 01:00:00 CET 1633 — To: Sat Dec 31 01:00:00 CET 1633

**Fecha de creación:**  
1633

**Tipo:**  
mapas

**Formato:**  
jpeg, 368, 464, Milímetros, Engraving

**Materia:**  
River Thames, London Bridge, London (England) — Maps, 912

**Identificador:**  
00700000000001U00031000, Maps.Grace I

**Es parte de:**  
Grace Collection Of Maps Of London, Collect Britain

**Por el proveedor:**  
The British Library

**Proporcionar Contenido:**  
The European Library

**Providing country:**  
United Kingdom

Etiquetas generadas automáticamente \*

Other items you may be interested in:

Figura 2. Ejemplo de Europea para Londres

Europeana, actualmente contiene metadatos abiertos sobre 2,4 millones de textos, imágenes, vídeos y sonidos recogidos por esta institución. Estos objetos proceden de proveedores de datos que han reaccionado pronto y positivamente a la iniciativa de Europea de promoción de datos más abiertos y nuevos acuerdos de intercambio de datos. Estas colecciones provienen de 8 proveedores directos de Europea que

abarca más de 200 instituciones culturales de 15 países. Cubren una gran variedad de objetos del patrimonio. Para el modelado de los datos siguen un esquema propio: *Europeana Data Model* (EDM). La adopción de este modelo basado en los estándares desarrollados por el W3C, ha permitido a Europea la compatibilidad con el paradigma de la web semántica. EDM está constituido por la reutilización de namespaces como RDF (*Resource Description Framework*), RDFs (*Resource Description Framework Schema*), OAI-ORE (*Open Archives Initiative Object Reuse and Exchange*), SKOS (*Simple Knowledge Organization System*) y DCMI Terms (*Dublin Core Metadata Initiative Terms*) (Figura 2).

*Data.bnf.fr*, recoge los datos de las diferentes bases de datos de la Biblioteca Nacional de Francia, con el fin de crear páginas web sobre obras, autores, materias y lugares, junto con una vista RDF de los datos extraídos. Fue uno de los proyectos pioneros que comenzó su actividad en julio de 2011. Como modelo conceptual utilizan el esquema FRBR de la IFLA y los vocabularios SKOS, FOAF (*Friend Of A Friend*), DC y RDA (*Resource Description and Access*).

DATOS.BNE.es

INICIO AUTORES OBRAS TEMAS Ayuda

Buscar un autor, obra o tema

Cervantes Saavedra, Miguel de / Don Quijote de la Mancha

Don Quijote de la Mancha

En la BNE

Disponibles 86 Versiones

Tema en 440 Obras

Versión destacada (Español) (ver las 863 ediciones)

Vida, y hechos del ingenioso caballero D. Quijote de la Mancha compuesta por Miguel de Cervantes Saavedra; parte II.

**Título:** Vida, y hechos del ingenioso caballero D. Quijote de la Mancha compuesta por Miguel de Cervantes Saavedra; parte II.

**Lugar de publicación:** En Madrid

**Editorial:** escuela de Santa Maria de Arnedo, escuela de Juan Antonio Bonet, mercader de libros, en France de S. Phelipe por Roque Rico de Miranda, impresor de libros.

**Fecha de publicación:** 1616

**Tipo de recurso:** Libro

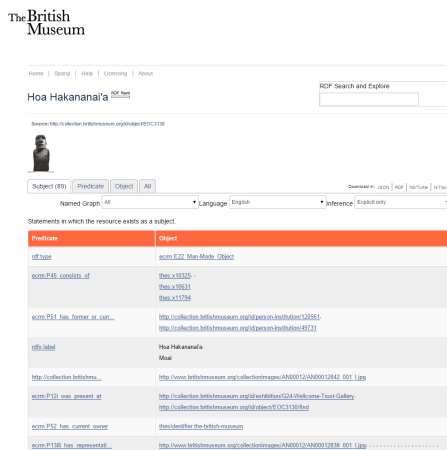
Figura 3. La obra *Don Quijote de la Mancha* en la BNE

*Datos.bne.es* es el proyecto de datos bibliográficos vinculados de la Biblioteca Nacional de España. Incluye 4 millones de registros de autoridad y 2.4 millones de registros bibliográficos que dan lugar a más de 58 millones tripletas. Los datos españoles se modelan mediante vocabularios IFLA (FRBR, FRISAD, ISBD) y otros como RDA y DC. Para la migración y la vinculación se utilizó Marimba, una herramienta especial desarrollada por la propia institución y el Ontology Engineering Group y diseñada para ser utilizada por los bibliotecarios (Figura 3).



En el apartado denominado *Archivos* destaca por su importancia dentro de los datos vinculados la agrupación *Archives Hub Linked Data*. Este ejemplo recoge una muestra de datos de descripciones de fondos de archivos mantenidos en el Archives Hub, un agregador del Reino Unido, y emitidos como *linked data*. Ofrece una perspectiva sobre las personas, las organizaciones, materias y lugares relacionados con los archivos que se describen. Se proporcionan enlaces externos a otros conjuntos de datos, como el VIAF y la LCSH. Asimismo, proporciona una hoja de estilo para convertir los datos EAD (XML para archivos) en RDF XML.

Dentro de la subclase catalogada como *Museos* es relevante la colección del *British Museum*. Este servicio de datos vinculados y SPARQL (*Protocol and RDF Query Language*) proporciona acceso a la misma colección de datos disponibles a través de la web del Museo. El uso del estándar de datos abiertos RDF del W3C permite que los datos de la colección del museo se unan y se refieran a un creciente cuerpo de datos vinculados publicados por otras organizaciones de todo el mundo interesadas en la promoción de la accesibilidad y la colaboración. Los datos se han dispuesto mediante el empleo del CIDOC-CRM (modelo conceptual de referencia) fundamental para la armonización con otros datos del patrimonio cultural (Figura 4).



Predicate	Object
<code>rdf:type</code>	<code>skos:Concept</code>
<code>skos:P43_coexist_of</code>	<code>http://www.oxfordjournals.org/abstract/doi/10.1093/ajph/117.12.1724</code>
<code>skos:P51_has_former_or_current_name</code>	<code>http://collection.britishmuseum.org/id/object/C013138</code>
<code>skos:P52_has_current_name</code>	<code>http://collection.britishmuseum.org/id/object/C013138</code>
<code>skos:P138_has_essential_uri</code>	<code>http://www.britishmuseum.org/collect/objects/AN00012400012401_001.jsp</code>

Figura 4. Hoa Hakananai'a del British Museum en rdf

El *Proyecto Gutenberg*, cuya dificultad en su clasificación ya hemos descrito anteriormente, es la primera y más grande colección de libros electrónicos gratuitos. Sobre este caso no se proporciona información técnica en el catálogo Mannheim desde el punto de vista de ejemplo de aplicación LOD. Su versión en RDF figura como un caso específico dentro de la categoría

metadatos, a la que posteriormente haremos referencia.

Si pasamos a la siguiente tipificación —*Vocabulario de valores*—, comenzando por la subclase *Sistemas de clasificación*, destaca la *Dewey Decimal Classification* (DDC). *Dewey.info* es un espacio experimental para los datos vinculados de la DDC. La intención de este prototipo es convertirse en una plataforma de los datos Dewey en la Web. Incluye como LD los sumarios DDC (los tres primeros niveles) de la vigésima segunda edición, en once idiomas, y todos los números asignables de la decimocuarta edición abreviada, en tres idiomas. La clasificación semántica se codifica en RDF utilizando SKOS y otros vocabularios. Cada clase tiene también una representación HTML (XHTML + RDFa) y varias serializaciones RDF —RDF/XML, Turtle (*Terse RDF Triple Language*) y JSON (*JavaScript Object Notation*)—.

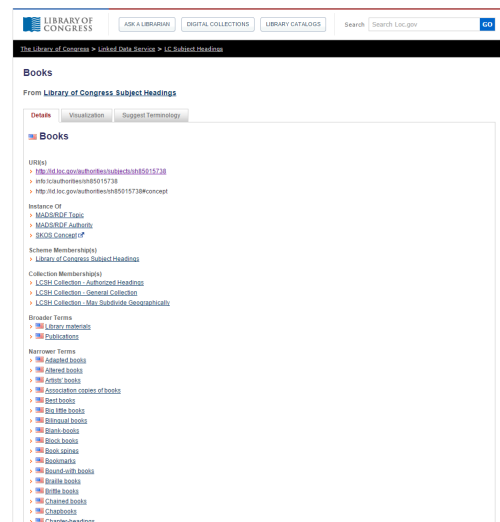


Figura 5. La materia Libros en la LCSH

Siguiendo con esta clase, en la subclase *Encabezamiento de materias* vamos a destacar por su interés los dos casos presentes: por un lado, la *Library of Congress Subject Headings* (LCSH) y la *Lista de Encabezamientos de Materia* (LEM). La LCSH es la lista de materias más utilizada a nivel internacional. Desde el año 1898 la Biblioteca del Congreso ha mantenido de forma activa esta lista para la indización de sus fondos. Los modelos utilizados para su implementación con LOD son los *Metadata Authority Description Schema* (MADS) —el esquema XML para los elementos de autoridad establecida por la propia institución— y SKOS (Figura 5).

Por su parte, mediante esta iniciativa la *Lista de Encabezamientos de Materia* (LEM) transforma su esquema a SKOS, y se adapta a *Europeana*

como uno de los *namespaces* constitutivos del *Europeana Data Model*. Ha establecido siempre que fue posible, un *skos:closeMatch* con los de la LCSH y se está progresando para proporcionar una funcionalidad análoga con RAMEAU (*Répertoire d'autorité-matière encyclopédique et alphabétique unifié*) o SWD (*Schlagwortnormdatei*). Incorpora además un fichero en zip de los registros MARC de autoridades de materia, es decir, aquellos basados en el campo 150 en los que se ha añadido las URI de las LEM y las de la LCSH (Figura 6).

The screenshot shows the LEM interface for the concept 'Libros'. It includes a search bar, navigation options, and a list of related concepts. The 'Conceptos relacionados' section lists terms like 'Libros (es)', 'Bibliografía (es)', 'Bibliografía (es)', 'Edición (es)', 'Ejemplares (es)', 'Ej. libro (es)', 'Imprenta (es)', 'Lectura (es)', and 'Museos del libro (es)'. The 'Conceptos similares en otros vocabularios' section lists 'Libros (g) (LEMA2)', 'Libros (ca) (LEMAC)', and 'Books (en) (LCSH)'. The 'Conceptos más amplios' section lists 'Bibliotecas-Fondos (es)'. The 'Conceptos más concretos' section lists 'Bibliometría (es)', 'Encuadernación (es)', 'Librerías (es)', 'Libros de bolsillo (es)', 'Libros de texto (es)', 'Libros en miniatura (es)', 'Libros ilustrados (es)', 'Libros infantiles y juveniles (es)', 'Libros prohibidos (es)', and 'Libros raros (es)'. The footer includes the logo of the Ministerio de Educación, Cultura y Deporte and various standards logos like SPARQL, RDF, SKOS, and BIBO.

Figura 6. El concepto Libros en la LEM

The screenshot shows the AGROVOC interface for the concept 'Wheat'. It includes a search bar and a list of related concepts. The 'PREFERRED TERM' section lists 'Wheat'. The 'BROADER CONCEPT' section lists 'Cereals'. The 'NARROWER CONCEPTS' section lists 'Hard wheat', 'Soft wheat', and 'Karnel bunt'. The 'HAS DISEASE' section lists 'Aegilops'. The 'HAS MEMBER' section lists 'Wheat flour' and 'Wheat straw'. The 'IN OTHER LANGUAGES' section lists 'Wheat' in various languages including Arabic, Chinese, Czech, French, German, Hindi, Hungarian, Italian, Japanese, Korean, Lao, Malay, Persian, Polish, Portuguese, Russian, Slovak, Spanish, Thai, Turkish, and others.

Figura 7. El concepto trigo en AGROVOC

Finalmente, en el último apartado de esta categoría destacaremos el tesoro AGROVOC publicado por la FAO (Food and Agriculture Organization). *AGROVOC Linked Open Data (LOD)* es un proyecto para convertir el tesoro AGROVOC en una columna vertebral terminológica multilingüe para artículos digitales agrícolas. Alojado por MIMOS, socio en la investigación, provee acceso web a los registros de datos estructurados sobre conceptos agrícolas y, más importante aún, enlaces sobre esos conceptos a

otros tesauros en línea. Actualmente describe más de 4.830.173 de recursos. La versión de LD del tesoro está en RDF/ SKOS-XL, y el conjunto de tripletas se almacena en Allegro-Graph (Figura 7).

Para terminar con este apartado descriptivo concluiremos haciendo referencia a la última clase establecida –*Metadatos*– y al último caso nombrado: *Project Gutenberg in RDF*. Esta iniciativa ofrece los metadatos sobre las obras de dominio público disponibles en el *Proyecto Gutenberg* al que anteriormente ya habíamos hecho referencia dentro de la subclase *Colección*. El proyecto ha sido realizado por el grupo de investigación Research Group Data and Web Science de la Universidad de Mannheim. Los datos estructurados se proporcionan de acuerdo a los principios de LD y se pueden acceder a los mismos a través del sitio web del grupo por medio de tres vías diferentes: HTML, RDF y SPARQL (Figura 8).

The screenshot shows the Project Gutenberg RDF interface for the concept 'Moby Dick'. It includes a search bar and a list of related concepts. The 'Property' section lists 'Moby Dick by Herman Melville'. The 'Value' section lists '1851-85-91 (void:datz)', '<http://wifo5-04.informatik.uni-mannheim.de/gutendata/resource/people/Melville\_Herman\_1819-1891>', 'There is no txt file, you'll have to get the zip.', 'en', 'Project Gutenberg', '<http://www.gutenberg.org/license>', '<http://wifo5-04.informatik.uni-mannheim.de/gutendata/resource/subject/Whaling\_-\_Fiction>', 'Moby Dick', 'dctype Text', and '<http://www.gutenberg.org/dirs/etext91/moby.zip>'. The 'Metadata' section lists '<http://wifo5-04.informatik.uni-mannheim.de/gutendata/data/etext15>', '2015-03-26T12:54:48.095Z', '<http://wifo5-04.informatik.uni-mannheim.de/gutendata/dataset-void/inDataset>', '<http://wifo5-04.informatik.uni-mannheim.de/gutendata/dataset-void/type Text>', 'prv:DataItem', and 'foaf:Document'. The footer includes the logo of the University of Mannheim and the text 'Generated by LOD2.org'.

Figura 8. Moby Dick en el Proyecto Gutenberg

## 4. Conclusiones

Primeramente nos gustaría subrayar las dificultades a las que hemos tenido que hacer frente a la hora de realizar nuestro estudio y las anomalías encontradas tanto en la configuración del *LOD cloud diagram* como en la estructura del catálogo *Mannheim*, con el objetivo de que puedan ser solventadas en las próximas ediciones de ambos instrumentos repercutiendo en una susceptible mejora de los mismos. No obstante, ponemos de manifiesto nuevamente la relevancia y utilidad de este tipo de materiales siendo conscientes igualmente de la complejidad de su elaboración.

Primero, pensamos que la denominación de la categoría *Publicaciones* no es muy acertada; si bien es cierto que los autores de la nube han querido englobar en un número reducido las

categorías que comprende el diagrama y que esta denominación es un concepto muy amplio y en ella se pretende reunir la información científica y cultural. Nuestra propuesta sería o bien cambiar la denominación con un epígrafe similar al anteriormente indicado o utilizar dentro de esta categoría etiquetas más específicas que nos permitirían encontrar directamente las distintas subcategorías. En este sentido, para los conjuntos de datos GLAM, proponemos la misma clasificación que hemos empleado nosotros, cuya definición viene avalada por el W3C.

Por otro lado, la información individualizada de cada *dataset* que nos proporciona el catálogo *Mannheim* es muy dispar y en ocasiones muy escueta. De tal modo que en un número importante de casos hemos tenido que proceder a buscar el enlace del mismo teniendo en cuenta que no es igual la dirección de una biblioteca, por ejemplo, que la dirección al proyecto de datos enlazados de esa misma institución. Es decir, la URL de la Biblioteca Nacional es *bne.es* mientras que su proyecto de datos enlazados es *datos.bne.es*. Incluso hay casos en los que ni siquiera se proporciona información, por lo que hemos tenido que llevar a cabo toda una tarea de investigación para identificar el ejemplo en concreto y así poderlo categorizar. El responsable último de este hecho no es tanto el propio catálogo sino los responsables de los *datasets*. Por lo tanto, abogamos porque se proporcione la mayor información posible por parte de los gestores de los conjuntos de datos en aras a facilitar tanto el uso del propio *dataset* como el análisis del mismo.

Uno de los objetivos que nos planteábamos en el inicio del trabajo era la definición de los conceptos claves tanto en lo concerniente a la tecnología de los datos abiertos y vinculados como a la definición de conceptos más específicos de nuestro ámbito, en concreto, los referentes a los recursos bibliotecarios empleados: colecciones, vocabularios de valores y metadatos. A través de la consulta de la documentación especializada sobre el tema comprobamos que en ocasiones ambos conceptos se confunden y se está perdiendo precisión terminológica. Mediante la explicación de los términos en el apartado teórico y la sistematización de los ejemplos que figura en el apéndice esperamos haber cumplido con nuestro propósito.

Centrándonos ya en los resultados obtenidos en lo referente a los *datasets* culturales, la categoría *Publicaciones* en la que se incluyen los mismos ocupa el tercer puesto dentro de la nube pero su porcentaje no alcanza ni siquiera el 10 % de la totalidad. Sin embargo, dentro de esta categoría los datos de las instituciones GLAM

son casi la mitad del resto de clases que comprende esta división. De acuerdo a la clasificación del LLD-XG el número de los datos es muy parecido tanto para la clase *Colecciones* como para la de *Vocabulario de valores*, con un ligero predominio de los primeros. Por su parte, dentro de la subclase *Colecciones* se obtienen porcentajes similares dándose en igual número los conjuntos de *Bibliotecas*, *Bibliografías* y *Catálogos*. En la clase de *Vocabulario de valores* destacan por igual los *Sistemas de clasificación* y los *Tesauros*, mientras en la última división se da una supremacía de las *Ontologías* frente a los *Metadatos*.

En el último apartado hemos intentando describir los casos más representativos de las categorías previamente establecidas, centrándonos no sólo en su definición sino en la aplicación de las tecnologías de datos vinculados. Esperamos así conseguir que las instituciones culturales que estén interesadas en publicar sus datos como LOD puedan obtener con este informe una primera aproximación a la aplicación de esta tecnología, remitiendo a cada caso particular para obtener una información más exhaustiva.

Sobre este punto diremos que nos hubiera gustado profundizar en la explicación de los recursos empleados a la hora de disponer los datos como LOD. Sin embargo, este hecho no ha sido posible debido a la dificultad anteriormente comentada para encontrar la información referente a la utilización de las diferentes prácticas relativas a este aspecto.

## Referencias

- Bauer, F.; Kaltentböck, M. (2012). *Linked Open Data: The Essentials: a Quick Start Guide for Decision Makers*. Viena, Austria: edition mono/monochrom, 2012.
- Bernes-Lee, Tim (2006). *Linked data – Design Issues*. [S.l.] : Bernes-Lee, 2006. <http://www.w3.org/DesignIssues/LinkedData.html> (2015-03-30).
- Europeana Linked Open Data (2015). European Union: Europeana Foundation: European Creative Project, 2015. <http://data.europeana.eu> (2015-03-30).
- Guía Breve de Linked Data (2015). W3C. Gijón: W3C Oficina España, 2015. <http://www.w3.es/Divulgacion/GuiasBreves/LinkedData> (2015-03-30).
- Heath, T.; Bizer, C. (2011). *Linked data: Evolving the Web into a Global Data Space*. Florida, USA: Morgan & Claypool Publishers, 2011.
- Isaac, Antoine; Waites, William; Young, Jeff; Zeng, Marcia. *Library Linked data Incubator Group: Datasets, Value Vocabularies, and Metadata Element Sets*. [S.l.]: W3C, 2011. <http://www.w3.org/2005/Incubator/ld/XGR-ld-vo-cabdataset-20111025> (2015-03-30)
- Library Linked Data Group Incubator (2011). *Linked Data Incubator Group Final Report*. [S.l.]: World Wide Web Consortium, 2011. <http://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/> (2015-03-30).
- Library Linked Data Group Incubator (2011). *Linked Data Incubator Group Use Case*. [S.l.]: World Wide Web Con-

- sortium, 2011. <http://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/#ref-USECASE> (2015-03-30).
- Library terminology informally explained (2011). [S.I.]: World Wide Web Consortium, 2011. [http://www.w3.org/2001/sw/wiki/Library\\_terminology\\_informally\\_explained#Definitions](http://www.w3.org/2001/sw/wiki/Library_terminology_informally_explained#Definitions) (2015-03-30).
- Linked data: connect Distributed Data across the Web (2014). [S.I.]: [s.n.], 2014. <http://linkeddata.org> (2015-03-30).
- Linking Open Data (2014). [S.I.]: World Wide Web Consortium, 2014. <http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData> (2015-03-30).
- Mannheim Linked Data Catalog (2014). Mannheim: University of Mannheim, 2014. <http://linkeddatacatalog.dws.informatik.uni-mannheim.de> (2015-03-30).
- Mitchel, Erik (2013). Library Linked data: research and adoption. // *Library Technology Reports*. 5:49 (2013) 11-25.
- Open Definition Advisory Council (2014). Cambridge: Open knowledge, 2014. <http://opendefinition.org/> (2015-03-30).
- Schmachtenberg, Max; Bizer, Christian; Paulheim, Heiko (2014a). Adoption of the linked data best practices in different topical domains. Mannheim: University of Mannheim; Unión Europea: Planet Data, 2014. <http://dws.informatik.uni-mannheim.de/fileadmin/lehrstuehle/ki/pub/SchmachtenbergBizerPaulheim-AdoptionOfLinkedDataBestPractices.pdf> (2015-03-30).
- Schmachtenberg, Max; Bizer, Christian; Paulheim, Heiko (2014b). State of the LOD Cloud 2014. Mannheim: University of Mannheim; Unión Europea: Planet Data, 2014. <http://linkeddatacatalog.dws.informatik.uni-mannheim.de/state/#toc1> (2015-03-30).

Enviado: 2015-04-01. Segunda versión: 2015-05-11.  
Aceptado: 2015-05-22.

## Apéndice

	<i>Bibliotecas</i>	<i>Princeton Library FindingAids</i>	<a href="http://findingaids.princeton.edu/">http://findingaids.princeton.edu/</a>
		data.bnf.fr - Bibliothèque nationale de France	<a href="http://data.bnf.fr/">http://data.bnf.fr/</a>
		datos.bne.es	<a href="http://datos.bne.es">http://datos.bne.es</a>
		Europeana Linked Open Data	<a href="http://data.europeana.eu/">http://data.europeana.eu/</a>
		Public Library of Veroia	<a href="http://libver.math.auth.gr">http://libver.math.auth.gr</a>
	<i>Archivos</i>	Airforcehistoryindex	<a href="http://airforcehistoryindex.org/">http://airforcehistoryindex.org/</a>
		Muninn World War I	<a href="http://rdf.muninn-project.org">http://rdf.muninn-project.org</a>
		Archives Hub Linked Data	<a href="http://data.archiveshub.ac.uk/">http://data.archiveshub.ac.uk/</a>
		National Digital Data Archive of Hungary	<a href="http://lod.sztaki.hu/">http://lod.sztaki.hu/</a>
<i>Colecciones</i>	<i>Museos</i>	British Museum Collection	<a href="http://collection.britishmuseum.org/">http://collection.britishmuseum.org/</a>
		Amsterdam Museum	<a href="http://semanticweb.cs.vu.nl/lod/am/">http://semanticweb.cs.vu.nl/lod/am/</a>
		Bibliografías Nacionales	
		British National Bibliography	<a href="http://bnb.data.bl.uk/">http://bnb.data.bl.uk/</a>
		Deutsche Biographie	<a href="http://www.deutsche-biographie.de/">http://www.deutsche-biographie.de/</a>
		Deutsche Nationalbibliografie	<a href="http://www.dnb.de/EN/datendienste/linkedData">http://www.dnb.de/EN/datendienste/linkedData</a>
		LIBRIS	<a href="http://libris.kb.se">http://libris.kb.se</a>
	<i>Colecciones</i>	Project Gutenberg	<a href="http://www.gutenberg.org/">http://www.gutenberg.org/</a>
	<i>Catálogos</i>	Sudoc bibliographic data	<a href="http://punktokomo.abes.fr/2011/07/04/le-sudoc-sur-le-web-de-donnees/">http://punktokomo.abes.fr/2011/07/04/le-sudoc-sur-le-web-de-donnees/</a>
		Hungarian National Library (NSZL) catalog	<a href="http://nektar.oszk.hu/wiki/Semantic_web">http://nektar.oszk.hu/wiki/Semantic_web</a>
	HeBIS	<a href="http://www.hebis.de/">http://www.hebis.de/</a>	
	85ZDB	<a href="http://www.zeitschriftendatenbank.de/services/schnittstellen/linked-data/">http://www.zeitschriftendatenbank.de/services/schnittstellen/linked-data/</a>	

	Sistemas de clasificación	Mathematics Subject Classification	<a href="http://msc2010.org/mscwork/">http://msc2010.org/mscwork/</a>	
		Faceted Application of Subject Terminology	<a href="http://experimental.worldcat.org/fast/">http://experimental.worldcat.org/fast/</a>	
		Dewey Decimal Classification (DDC)	<a href="http://dewey.info">http://dewey.info</a>	
		JITA Classification System of Library and Information Science	<a href="http://www.destin-informatique.com/ASKOSI/Wiki.jsp?page=JITA%20Maintenance">http://www.destin-informatique.com/ASKOSI/Wiki.jsp?page=JITA%20Maintenance</a>	
		ICONCLASS - Multilingual Thematic Classification	<a href="http://www.iconclass.org/help/lod">http://www.iconclass.org/help/lod</a>	
		Glottolog	<a href="http://glottolog.org">http://glottolog.org</a>	
Vocabularios de valores	Encabezamientos de materia	Lista de Encabezamientos de Materia	<a href="http://id.sgcb.mcu.es">http://id.sgcb.mcu.es</a>	
		Library of Congress Subject Headings	<a href="http://id.loc.gov/authorities/">http://id.loc.gov/authorities/</a>	
	Datos de autoridad	IdRef: Sudoc authority data	<a href="http://punktokomo.abes.fr/2011/07/05/idref-des-pages-html-et-rdf-plus-riches/">http://punktokomo.abes.fr/2011/07/05/idref-des-pages-html-et-rdf-plus-riches/</a>	
		Gemeinsame Normdatei (GND)	<a href="http://d-nb.info/standards/elementset/gnd#">http://d-nb.info/standards/elementset/gnd#</a>	
	Tesauros		yso-fi-allars	<a href="http://finto.fi/allars/en/page/Y46304?rdf+xml">http://finto.fi/allars/en/page/Y46304?rdf+xml</a>
			Unesconom	<a href="http://skos.um.es/unesco6/6101/rdfxml">http://skos.um.es/unesco6/6101/rdfxml</a>
			STW Thesaurus for Economics	<a href="http://zbw.eu/stw/versions/latest/about">http://zbw.eu/stw/versions/latest/about</a>
			National Agricultural Library Thesaurus	<a href="http://agclass.nal.usda.gov/agt.shtml">http://agclass.nal.usda.gov/agt.shtml</a>
			AGROVOC	<a href="http://aims.fao.org/vest-registry/vocabularies/agrovoc-multilingual-agricultural-thesaurus">http://aims.fao.org/vest-registry/vocabularies/agrovoc-multilingual-agricultural-thesaurus</a>
			Mis Museos, índice semántico de museos, artistas y obras de arte (GNOSS)	<a href="http://mismuseos.net/comunidad/metamuseo">http://mismuseos.net/comunidad/metamuseo</a>
Ontologías	Ontologías	yso-fi-ysa	<a href="http://www.yso.fi/onto/ysa/Y99974?rdf+xml">http://www.yso.fi/onto/ysa/Y99974?rdf+xml</a>	
		Bible Ontology	<a href="http://bibleontology.com/service/">http://bibleontology.com/service/</a>	
		ASN:US	<a href="http://asn.jesandco.org">http://asn.jesandco.org</a>	
		Traditional Korean Medicine Ontology	<a href="http://tkm.kiom.re.kr">http://tkm.kiom.re.kr</a>	
		FAO geopolitical ontology	<a href="http://www.fao.org/countryprofiles/geoinfo/en/?lang=en">http://www.fao.org/countryprofiles/geoinfo/en/?lang=en</a>	
Metadatos		Project Gutenberg in RDF	<a href="http://www4.wiwiss.fu-berlin.de/gutendata/">http://www4.wiwiss.fu-berlin.de/gutendata/</a>	

Tabla VII. Listado de los conjuntos de datos bibliotecarios del LOD cloud diagram

