

Los datos bibliográficos abiertos enlazados y su comportamiento en la recuperación de información

Eder Ávila Barrientos*

Artículo recibido:

13 de junio de 2019

Artículo aceptado:

9 de septiembre de 2019

Artículo de revisión

RESUMEN

Se analiza el comportamiento de los datos bibliográficos abiertos enlazados (DBAE) y su interacción en el proceso de recuperación de información. Mediante la hermenéutica y el análisis del discurso, se examinó la literatura que pone de manifiesto modelos, estructuras, estudios de caso, pruebas de concepto y análisis estructurales de la implementación de DBAE en la recuperación de información. Los hallazgos identificados permiten estimar que la explosión de los datos en el ambiente digital pone de manifiesto la necesidad de contar con métodos que fomenten la recuperación de información de una manera gráfica, intuitiva e interactiva para el usuario final. Los registros bibliográficos

* Instituto de Investigaciones Bibliotecológicas y de la Información, UNAM, México
eder@iibi.unam.mx

semánticos y los grafos bibliográficos son dos métodos que pueden fomentar este tipo de recuperación. En un contexto datificado, es necesario tener mecanismos que hagan posible la integración y vinculación de los datos que están disponibles en los recursos y contenidos que son localizados en el ambiente web. Cuando DBAE es aplicado en el dominio de la información documental, es posible generar datos bibliográficos con la capacidad de vincularse abierta y significativamente mediante un sistema fundamentado en un esquema interoperable.

Palabras clave: Recuperación de Información; Datos Bibliográficos Abiertos Enlazados; Registro Bibliográfico Semántico; Grafo Bibliográfico

The linked open bibliographic data and its behavior in the information retrieval

Eder Ávila Barrientos

ABSTRACT

The behavior of the linked open bibliographic data (LOBD) and its interaction in the process of information retrieval, is analyzed through hermeneutics and discourse analysis. We examined the literature that reveals models, structures, case studies, concept tests and structural analysis of the implementation of LOBD in information retrieval. The identified findings allow us to estimate that the explosion of data in the digital environment reveal the need to have methods that encourage information retrieval in a graphic, intuitive and interactive way for the end user. Semantic bibliographical records and bibliographical graphs are two methods that can promote this type of recovery. In a datified context, it is necessary to have mechanisms that make possible the integration and linkage of data which are available in the resources and contents located in the web environment. When LOBD is applied in the domain of documentary information, it is possible to generate bibliographic data with the capacity of being openly and significantly vinculated through a system based on an interoperable scheme.

Keywords: Information Retrieval; Linked Open Bibliographic Data; Semantic Bibliographic Record; Bibliographic Graph

INTRODUCCIÓN

Bibliographic Framework Initiative (BIBFRAME) y Library Reference Model (LRM) son dos modelos que tienen el propósito general de establecer pautas para la conformación de datos bibliográficos enlazados, fomentar la interoperabilidad global de estos datos y contribuir con ello al desarrollo de la web semántica. Taniguchi (2018: 452) realizó un mapeo de los elementos de ambos modelos en donde expuso sus similitudes y diferencias con miras a su fusión sistemática. De esta manera, la sistematización de los datos bajo este enfoque pone de manifiesto nuevos comportamientos en la recuperación de información (RI).

En el ámbito de las bibliotecas los datos bibliográficos tienen una función importante para la búsqueda y accesibilidad de los recursos de información, pues son elementos que se relacionan íntegramente con los atributos que presentan dichos recursos. Estos atributos son representados en los registros bibliográficos.

La descripción y el acceso a los recursos de información requieren de modelos de datos que fomenten su interoperabilidad y recuperación. En este sentido, Linked Open Bibliographic Data (LOBD) es un término utilizado para describir un modelo para la estructuración, vinculación y acceso a los datos bibliográficos que están disponibles en los registros que forman parte de los catálogos de acceso en línea, repositorios, descubridores de información y motores de búsqueda bibliográfica. Este término es traducido al español como datos bibliográficos abiertos enlazados (DBAE).

El propósito de este trabajo consiste en analizar el comportamiento de los datos bibliográficos abiertos enlazados y su interacción en el proceso de recuperación de información. Resultan escasos los estudios teóricos que permitan profundizar en la naturaleza de DBAE y su función en el ámbito de la RI. Sin embargo, los hallazgos localizados han permitido identificar el potencial de estos datos para proponer nuevos modelos para recuperar información. Por ejemplo, Ullah *et al.* (2018: 66) presentaron una revisión holística del estado actual de los datos abiertos enlazados en el ámbito de la catalogación. En este trabajo se identificaron los potenciales de Linked Open Data (LOD) y Linked Open Vocabularies (LOV) para realizar descripciones bibliográficas que puedan ser publicadas, vinculadas y consumibles en la web.

Wenige y Ruhland (2018: 267) sugieren que los datos abiertos enlazados son una valiosa alternativa para la recuperación y acceso en el contexto de las bibliotecas. De esta manera proponen la implementación sistemática de los principios técnicos de LOD, mediante el uso del Simple Knowledge Organization System (SKOS). El análisis de los datos es una de las primeras fases de

la implementación de Linked Open Data en cualquier entorno. De hecho, la figura de los datos abiertos enlazados no puede abordarse sin tomar en cuenta los procesos de gestión y análisis que conlleva su estructuración. Cristovao y Fernandes (2018: 204) desarrollaron un estudio cuyo objetivo fue la construcción de un modelo de recuperación de información basado en datos enlazados, el cual permitiría la representación de los resultados mediante el uso de redes complejas. Si bien el uso de estos elementos es una alternativa para la consulta y visualización de los datos, los grafos son métodos adaptables para representar gráficamente las conexiones establecidas entre los datos de esta naturaleza.

La proyección de los catálogos de las bibliotecas a un escenario semántico y la estructuración del registro bibliográfico y su capacidad para vincularse en el entorno web son dos factores que motivan la aparición de nuevos modelos de recuperación. Al respecto, Bermès (2015: 24) diseñó un modelo de plataforma fundamentada en RDF y URI. En este modelo, los datos son ligados mediante los siguientes conceptos: obra, documento, persona, lugar, colección, evento y recurso. En este trabajo se analiza la disponibilidad, generación y modelado de los datos enlazados en el ambiente de las bibliotecas. A su vez, Le Boeuf (2015: 46) explica la conformación del prototipo de OpenCat, un proyecto que fue desarrollado e impulsado por la Biblioteca Nacional de Francia con los objetivos siguientes:

- Utilizar las técnicas de la web semántica para recuperar en una sola lista los documentos de varias fuentes;
- Proponer enlaces directos a documentos digitalizados de libre acceso siempre que sea posible, y
- Proponer un nuevo enfoque para la navegación temática mediante la representación gráfica intuitiva.

Por otra parte, Elsayed y Mesbah (2018: 5) argumentan que “publicar registros bibliográficos como datos enlazados, permitirá vincular entidades en los catálogos de la biblioteca, y relacionar dichas entidades en la nube de datos enlazados abiertos”. No obstante, para que este propósito se convierta en una realidad es necesario definir patrones de interoperabilidad entre los sistemas de información utilizados por las bibliotecas. “La interoperabilidad entre los conjuntos de datos bibliográficos es importante para la utilización de datos globales, no solo dentro del campo de la biblioteca sino también externamente entre los consumidores de datos que desean compilar datos de fuentes complementarias” (Talleras, 2017:150).

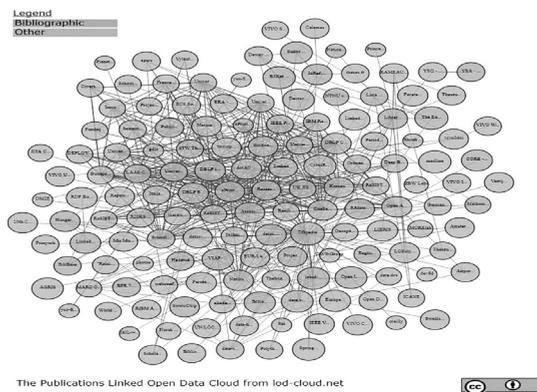


Figura 1. Grafo de subnube del dominio de publicaciones perteneciente a The Linked Open Data Cloud

Fuente: <https://lod-cloud.net/clouds/publications-lod.svg>

En la *Figura 1* puede apreciarse la representación de los conjuntos de datos correspondientes al área de publicaciones, en la cual se encuentran proyectados los conjuntos de datos bibliográficos pertenecientes a bibliotecas, vocabularios y actividades relacionadas con la publicación de datos enlazados en ambientes de información documental. En el mes de marzo de 2019, este dominio registró un total de 1 239 conjuntos de datos con 16 147 vínculos entre ellos. Algunos de los conjuntos que reúne son los siguientes: Virtual International Authority File (VIAF), Linked Data Service der Universitätsbibliothek Mannheim, British National Bibliography (BNB) - Linked Open Data y The European Library Open Dataset.

Las grandes bibliotecas universitarias, bibliotecas nacionales y consorcios de bibliotecas han sido los pioneros en sentar las bases de los datos enlazados en la biblioteca de dos maneras: desarrollando modelos de datos y esquemas para representar información de recursos como datos enlazados y mediante la publicación y agregación de metadatos de recursos como datos enlazados. (Landis, 2019: 7)

De esta manera, se considera que los registros bibliográficos pueden adquirir una estructuración que fomente su capacidad para vincular datos, recursos y contenidos disponibles no sólo en las bibliotecas sino también en el ambiente web. Para ello, es necesario que los datos bibliográficos tengan valores de integridad y estandarización compatibles con los principios de Linked Open Data. “Se trata de interconectar datos de forma dinámica, proporcionando la complementación y significado en los resultados de búsqueda. Esto se reflejará en catálogos enriquecidos, con información adicional, favoreciendo la

selección de la fuente deseada por el usuario” (Giusti Serra y Santarém Segundo, 2017: 182). Proyectos como Share-VDE (<http://www.share-vde.org/sharevde/clusters>) así lo constatan. Además, es necesario fundamentar teóricamente la naturaleza de los datos bibliográficos abiertos enlazados para establecer una metodología susceptible de aplicarse en el óptimo manejo de estos datos e ir más allá de la adaptación y utilización de vocabularios y la generación de propuestas procedimentales aplicadas en escenarios especializados. Por lo tanto, como parte de este trabajo, es conveniente responder dos preguntas elementales:

- ¿Cuál es la función y el comportamiento de los datos bibliográficos abiertos enlazados en el proceso de recuperación de la información?
- ¿Qué características tiene la recuperación de la información mediante la explotación de estos datos?

Por consiguiente, se realiza un análisis de la literatura relacionada con la implementación de DBAE en la RI. El método utilizado es la hermenéutica y análisis del discurso para examinar la literatura que consiste en modelos, estructuras, estudios de caso, pruebas de concepto y análisis estructurales de la implementación de estos datos en el proceso de recuperación de información. Los documentos analizados fueron extraídos de bases de datos especializadas en bibliotecología y estudios de la información y ciencias de la computación, así como de catálogos y descubridores académicos y especializados. Se implementaron estrategias de búsqueda relacionadas con los siguientes términos: Linked Data, Linked Open Data, Information Retrieval y su equivalente en español. Se seleccionaron aquellas obras publicadas de 2013 a 2019, pues al tratarse de un tema tecnológico el factor cronológico fue tratado con rigurosidad temporal. Aquellas obras anteriores a esta cronología fueron seleccionadas tomando en cuenta su grado de contribución y relevancia respecto al análisis del objeto de estudio planteado en la investigación.

Derivado del análisis teórico efectuado, se presentan los componentes que forman parte de un marco para la recuperación de información fundamentado en los principios de DBAE, el cual se encuentra influenciado por la formulación de registros bibliográficos semánticos y grafos bibliográficos, elementos que permiten obtener un panorama inicial de la recuperación de información con estas características.

LOS DATOS BIBLIOGRÁFICOS ABIERTOS ENLAZADOS

El estudio de la aplicación de DBAE en el ámbito de la RI es reciente; sin embargo, se han identificado estudios que ponen de manifiesto consideraciones de índole pragmática acerca de su manejo, principalmente en los contextos bibliotecarios y computacionales. Los datos enlazados (también identificados como Linked Data) son estructuras para la vinculación y publicación de datos que están disponibles en la web. “El nombre original de la web de los datos enlazados fue Grafo Gigante Global (GGG). Este fue el nombre que Tim Berners-Lee acuñó para la web semántica cuando desarrolló su idea en el 2001” (Powell y Hopkins, 2015: 65).

En lo que respecta a la construcción y publicación de datos enlazados, este proceso se lleva a cabo mediante la aplicación de cuatro principios generales (Berners-Lee, 2006: párr. 3):

1. Utilizar URI para nombrar a las cosas disponibles en la web;
2. Emplear el protocolo de transferencia de hipertexto (por sus siglas en inglés HTTP) para que las personas puedan buscar esos nombres;
3. Cuando alguien busque un URI, proporcionar información importante utilizando RDF (Marco de Descripción de Recursos, del inglés Resource Description Framework) y SPARQL (Lenguaje para la consulta de RDF), y
4. Incluir enlaces a otros URI para que la gente pueda descubrir más cosas.

Al respecto, Rietveld (2016: 2) destaca que RDF y Linked Data ofrecen un modelo para representar estructuras semánticas en un grafo; un modelo de datos que no está vinculado a un esquema fijo y que permite definir conjuntos de datos adaptables y flexibles para vincular URI y diversos vocabularios.

Por lo tanto, la asignación de URI a los datos bibliográficos permitirá descubrir los recursos de información y contenidos que tienen una vinculación en el ambiente web, a partir de la descripción de las semejanzas de los atributos bibliográficos, temáticos y de autoridad que los caracterizan. Estos atributos son representados a manera de datos en los registros bibliográficos. Al respecto, Svenonius (2018: 716) manifiesta que “los elementos de datos en los registros bibliográficos se pueden clasificar en dos categorías: los que describen la entidad en cuestión y los que relacionan la entidad con otras entidades”.

De acuerdo con Hagler y Simmons (1982: 55) los datos bibliográficos son de tres tipos:

- Datos que identifican unívocamente a un documento en particular para distinguirlo de otros;
- Datos que revelan una asociación de dos o más documentos, y
- Datos que describen algunas características del contenido intelectual del documento.

Para vincular abiertamente a los recursos y contenidos disponibles en la web, es necesario que los registros de datos bibliográficos se encuentren libres de restricciones técnicas, legales y económicas. En este escenario, es pertinente mencionar la función que Linked Open Data tiene en el proceso de reutilización de los datos.

LOD propicia la generación de estructuras de datos enlazados que sean publicadas mediante el uso de licencias de datos abiertos. El fundamento principal de LOD consiste en publicar y vincular datos que estén libres de restricciones mediante el uso de estándares y protocolos que fomenten la conformación de un ecosistema abierto e interoperable de datos enlazados. Como parte de dicho ecosistema, es importante reconocer la autoría de los datos que han sido procesados como parte de su creación y reutilización. Entonces, para la publicación y vinculación de LOD en el entorno web es pertinente cumplir con cinco criterios básicos (Berners-Lee, 2010: párr. 56):

- Los datos deben estar disponibles en la web (en cualquier formato), pero liberados con una licencia abierta;
- Datos estructurados legibles por computadora (por ejemplo, colocarlos en una hoja de cálculo, en lugar de escanear o capturar en modo imagen las tablas que los contienen);
- Utilizar formatos no propietarios para codificarlos (por ejemplo, CSV en lugar de Excel);
- Utilizar estándares abiertos del W3C como RDF y SPARQL para que las personas puedan identificar cosas, y
- Vincular nuestros datos con los datos de otras personas para proporcionar un contexto.

Así pues, los datos bibliográficos abiertos enlazados están estructurados y representan los atributos bibliográficos de los recursos de información mediante el uso de una norma de catalogación, descripción o a través de formatos de codificación y esquemas de metadatos. Son publicados con licencias abiertas, como Creative Commons y el grupo de licencias de Open Data. No obstante, “algunos datos de bibliotecas tienen un uso restringido, basado en

políticas locales, contratos y otros condicionantes poco claros o difíciles de comprobar, que pueden entorpecer su edición como datos abiertos. Los temas relacionados con los derechos varían mucho de un país a otro y dificultan la colaboración para la publicación de datos abiertos” (W3C, 2011).

En ese caso, las licencias Open Data son instrumentos de acceso abierto que los profesionales de la información pueden aplicar para propiciar la libre reutilización e interoperabilidad de DBAE, tomando en consideración que “el intercambio de datos abiertos puede ser increíblemente beneficioso para la sociedad: facilita la colaboración científica y la reproducibilidad” (Creative Commons, s. a.: párr. 2). Definir la propiedad de los registros bibliográficos es un asunto complicado, pues el gran intercambio de datos entre bibliotecas en los últimos 50 años ha derivado en la copia y duplicación de los registros en diferentes catálogos. Los catalogadores modifican o mejoran localmente esas copias y las agregan a otros catálogos nacionales o internacionales. La falta de certeza en el ámbito de la licencia de los registros dificulta el intercambio de datos en una comunidad que es cautelosa en cuestiones legales.

Por lo tanto, la implementación de DBAE en el ámbito de la recuperación de información requiere tomar en cuenta las cuestiones de libre uso de los datos pues la reutilización, integridad e interoperabilidad de los datos son conceptos que están inmersos en un marco de libre acceso a la información. Bajo esta premisa, la recuperación de la información deberá considerarse de carácter universal e integradora.

RECUPERACIÓN DE INFORMACIÓN

La recuperación de información ha sido analizada de manera persistente en el campo de la bibliotecología y los estudios de la información. Otras disciplinas, como las ciencias de la computación y la informática, también se han dedicado a su estudio y a comprender su aplicación en diferentes entornos informativos. Para Ingwersen (1992: V) “la recuperación de la información abarca los problemas relativos al efectivo almacenamiento, acceso y búsqueda de la información que requieren los individuos”. La explosión de la información y su incesante crecimiento fomentado por el uso de las TIC, la expansión de internet y el impacto de la web en todos los aspectos de la vida humana ha provocado la generación de problemáticas relacionadas con la óptima recuperación de información. De hecho, los datos bibliográficos abiertos enlazados han nacido como un modelo que se propone para obtener resultados significativos al momento de recuperar información.

Autores como Mortimer Taube, Hans Peter Luhn, Calvin Mooers y Gerard Salton (Chu, 2003: 5) fueron pioneros en estudiar el fenómeno de la recuperación de la información desde una perspectiva informática y documental. Algunos de sus fundamentos fueron aplicados en la generación de sistemas permutados que fomentaran la óptima recuperación de la información. Estos sistemas siguen vigentes en las aplicaciones para la búsqueda y recuperación de información en bibliotecas, a pesar de la evolución de los soportes, formatos y atributos que caracterizan a los nuevos recursos de información. Además, la aparición de los Sistemas de Recuperación de Información (SRI) fue un factor clave para motivar la precisión y optimización de los resultados de búsqueda.

Al respecto, Salvador Oliván y Arquero Avilés (2006: 40) definen a la RI como “la aplicación del conjunto de técnicas, métodos y actividades para buscar, localizar y recuperar de una manera eficiente en los diversos SRI la información relevante que requiere el usuario, y satisfacer así su necesidad de información”. Entonces, la recuperación de la información es un proceso que está fuertemente influenciado por la descripción y representación normalizada de los recursos. Además, debido a la constante expansión del universo de información, es evidente que la RI es vista como un proceso en constante evolución, pero con una problemática permanente. Aunado a ello, Blair (1990: 22) manifiesta que el principal problema de la RI consiste en cómo representar a los documentos para que puedan ser recuperados por las personas que los encuentran útiles. Esta problemática aumenta cuando los nuevos recursos de información aparecen en escena, pues sus atributos son más complejos de identificar y representar.

De manera que el progreso de la web ha influido notablemente en la generación de nuevos modelos para la recuperación de la información. Por ejemplo, los motores de búsqueda como Google han adoptado sugerentes métodos y algoritmos que tienen el objetivo de optimizar la RI de una manera integradora en función de las estratosféricas fuentes de información que se localizan en la web. De hecho, “la historia de la teoría de la recuperación de la información en el siglo XX está reflejada en el desarrollo de diversos y especializados motores de búsqueda y sistemas de bases de datos” (Haynes, 2004: 80). Estos motores tienen el ambicioso propósito de buscar los elementos alfanuméricos que aparecen en las estructuras de diferentes recursos y contenidos que se localizan en la web, sin importar el dominio de conocimiento al que pertenecen o el tema que abarquen.

Entonces, la pertinencia y relevancia de la RI son factores que dependen de la efectividad del lenguaje de interrogación que se utilice para las consultas en un determinado sistema de información. Además, los registros

bibliográficos son un componente que permite buscar y recuperar los datos que están almacenados en dicho sistema. De esta manera, la efectividad del lenguaje y su aplicación en las estrategias de búsqueda son elementos clave para la relevancia de los resultados recuperados.

Una vez que los datos bibliográficos han sido codificados binariamente, pueden recuperarse mediante un sistema informático desarrollado expro-feso para cumplir con este propósito. Sin embargo, los datos bibliográficos tienen que estar debidamente descritos, organizados y representados para propiciar su óptima búsqueda y recuperación. De esta manera, los estándares y normas de las bibliotecas deben adaptarse a un entorno de interoperabilidad que fomente la flexibilidad de los datos bibliográficos respecto a su interacción y vinculación con otros sistemas ajenos al catálogo. Tal y como señala Martínez Arellano (2012: 8):

Si se acepta la premisa de que actualmente el usuario desea tener acceso a toda la serie de recursos que existen sobre una temática, sin importar si éstos se encuentran en la biblioteca o fuera de ella, es imprescindible la necesidad de moverse rápidamente hacia la transformación del catálogo en un sistema de localización que abarque todo tipo de recursos independientemente del lugar donde éstos se encuentren.

La transformación del catálogo en otro sistema de localización pone de manifiesto la utilización de modelos de RI diferentes a los tradicionales modelos textuales utilizados por las bibliotecas. En este nuevo escenario, los registros bibliográficos tienen el potencial de conectarse con otras fuentes disponibles en el entorno web, propiciando una RI de características universales, integradoras y vinculadoras, de ahí que datos e información sean considerados dos elementos que se encuentran en constante interacción en un proceso de recuperación.

La RI pone de manifiesto el “número de registros o documentos que son recuperados por diferentes formas de búsqueda, usando secuencias de palabras y palabras múltiples” (Warner, 2010: 134). Esta diferenciación permite establecer un patrón de comportamiento de la RI respecto a la utilización de DBAE en un contexto de información documental específico, pues un modelado de recuperación fundamentado en DBAE permitirá obtener registros que se vinculan significativamente con los atributos de los recursos y contenidos y no necesariamente por el número de palabras que se encuentren en sus estructuras. Para Chu (2003: 17) “el principal problema de la representación y recuperación de la información sigue siendo cómo obtener la información adecuada, para el usuario correcto en el momento adecuado, a pesar de

la existencia de múltiples variables”. En efecto, para la óptima recuperación de información, es necesario contar con métodos de representación efectivos que propicien una secuencia exacta al momento de recuperar información. Así pues, la recuperación de la información mediante DBAE es vista como parte de un proceso integrador (*Figura 2*) que además contempla la descripción, representación, vinculación y acceso a los recursos y contenidos. Se estima que estas etapas permitan la recuperación y consulta de los datos que se vinculan con diferentes recursos y contenidos con similitudes en sus atributos de autoridad, título y tema.

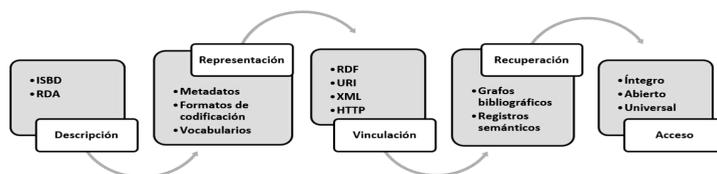


Figura 2. Etapas elementales de la RI fundamentada en DBAE
Fuente: elaboración propia, 2019

La descripción de los datos bibliográficos se realiza tomando en cuenta los principios de organización de las normas y estándares bibliotecarios, por ejemplo, los incluidos en ISBD y RDA. Al respecto, Dunsire (2014: 866) manifiesta que “ISBD tiene un papel importante en las aplicaciones de datos enlazados al proporcionar datos agregados que están disponibles en las diferentes áreas de descripción”. Los modelos de datos enlazados bibliotecarios han decidido tomar en consideración estos estándares para la conformación de vocabularios, los cuales son “un grupo de términos que tienen alguna relación entre sí y que describen los elementos que forman parte de un dominio de conocimiento” (Powell y Hopkins, 2015: xxvi).

Además, el uso de vocabularios es necesario para llevar a cabo una estructuración uniforme de los datos bibliográficos. Al respecto, Linked Open Vocabularies es una plataforma que reúne a los vocabularios LOD que son utilizados por diferentes comunidades enfocadas al uso y manejo de datos enlazados. Cuando los datos bibliográficos son descritos y representados, es necesaria su vinculación individual mediante la asignación de URI y su estructuración mediante RDF y XML para que su acceso en la web mediante el protocolo HTTP sea viable.

Por consiguiente, DBAE deberá fomentar una recuperación de información mediante el uso de registros semánticos y grafos, los cuales deberán ofrecer un acceso íntegro, abierto y universal a los datos y los respectivos recursos de información a los que pertenecen, pues “el modelo de búsqueda y navegación

de las bibliotecas también es transferible a contextos de datos enlazados. En la web semántica, los vocabularios controlados permiten el uso y descubrimiento de la información en colecciones y dominios” (Jai O’Dell, 2016: 50). De esta manera, la interacción de DBAE en el proceso de recuperación de información pone de manifiesto la vinculación semántica de los datos bibliográficos con los recursos y contenidos de atributos similares. En este proceso se observan patrones de comportamiento que se analizan a continuación.

ANÁLISIS DEL COMPORTAMIENTO DE LOS DATOS BIBLIOGRÁFICOS ABIERTOS ENLAZADOS EN LA RECUPERACIÓN DE INFORMACIÓN

El uso de DBAE para recuperar información conlleva su aplicación en la descripción bibliográfica de un recurso, lo cual propicia el desarrollo de dos métodos principales para representar y recuperar información respectivamente: por un lado, la generación de un registro bibliográfico semántico, el cual permita establecer una estructura de datos bibliográficos capaz de vincularse en el ambiente web de manera interoperable mediante la utilización de los principios de Linked Data; por otro, la formulación de grafos bibliográficos para obtener una RI fundamentada en una visualización integral de las vinculaciones que existen entre los datos bibliográficos pertenecientes a un determinado dominio de información documental.

De este modo, DBAE puede contribuir al descubrimiento de patrones de comportamiento entre los datos bibliográficos que pertenecen a las diversas obras, manifestaciones y expresiones de información, a través de:

- Identificación de las vinculaciones de los datos enlazados bibliográficos en diversas fuentes de información localizadas en la web;
- Sistematización del control de autoridades y descubrimiento de relaciones existentes en las obras intelectuales;
- Apertura de las herramientas de búsqueda y acceso a la información de bibliotecas y otras unidades de información, y
- Enriquecimiento de la nube de datos abiertos enlazados y fomento a una mayor interoperabilidad entre sistemas y aplicaciones que gestionan información bibliográfica.

Aunado a ello, DBAE pone de manifiesto la interacción del registro bibliográfico semántico con el grafo bibliográfico mediante una interfaz que permita su consulta de manera interna y externa al entorno web, pues “la transición

de los datos enlazados a una interfaz de mapa de conocimiento es una forma en que los datos enlazados pueden presentarse para uso del usuario” (Faith y Chrzanowski, 2015: 133).

Por lo tanto, la recuperación de información mediante DBAE depende de los siguientes factores:

- Contar con un vocabulario LOD que sea interoperable y aplicable con los diversos esquemas y formatos disponibles en el contexto bibliográfico. Un vocabulario de estas características proporciona el aditivo semántico mediante el cual los datos adquieren significado en los diferentes contextos de la web;
- Implementar sistemas e interfaces que permitan la consulta de DBAE de una manera estructurada, usable, integral e intuitiva para el usuario final, y
- Trabajar en un entorno informático basado en el modelo de vinculación semántica, esto mediante la utilización de bases de datos enfocadas a grafos. NEO4J, GraphDB y Titan GDB son ejemplos de software que puede ser utilizado para cumplir con este objetivo.

De lo anterior, se infiere que la RI caracterizada por el uso de DBAE contempla la aplicación de modelos que permitan una consulta gráfica y visualmente intuitiva de las vinculaciones de significado existentes entre los datos. Esta modalidad de recuperación ha sido tratada con anterioridad, pero de manera sintáctica mediante la implementación del modelo vectorial. Al respecto, Salton y McGill (1983: 131) afirman que “los sistemas de análisis semántico podrían complementar las unidades sintácticas mediante el uso de roles semánticos adjuntos a las entidades que constituyen una descripción de contenido. Los sistemas de análisis semántico utilizan diversos tipos de conocimiento ajenos a los documentos, a menudo especificados por grafos semánticos pre-construidos”. Cuando los datos se representan en un grafo, es posible identificar las vinculaciones entre ellos mediante la definición de variables previamente establecidas, es decir, mediante la asignación de clases y propiedades que permitan estructurar y describir los atributos de los datos. De modo que la RI mediante DBAE no se trata únicamente de un proceso para la obtención de datos, sino de comprender el significado de las vinculaciones que se establecen entre ellos. A continuación se abordan los componentes que caracterizan a este binomio.

Registro bibliográfico semántico

El registro bibliográfico es el resultado de la descripción bibliográfica de los recursos de información, y “consiste en los datos bibliográficos relativos a un ítem ensamblado de acuerdo con un formato particular. Un formato es un contenedor con compartimentos, diseñado para ajustarse a los datos de tal manera que cada elemento de datos sea reconocible y pueda ser recuperado” (Hagler y Simmons, 1982: 109). Este registro es el producto de un análisis intelectual y normativo que es realizado por los catalogadores, el cual consiste en la identificación, extracción y representación de los datos bibliográficos que describen los atributos de autoridad, físicos y de contenido de los recursos. Para Anderson (1989: 3) el registro bibliográfico representa “la suma de todas las áreas y elementos usados para describir, identificar o recuperar cualquier ejemplar físico o su contenido informativo”. En efecto, un registro de estas características se encuentra estructurado en áreas que contienen elementos descriptivos. En los campos correspondientes a dichos elementos se registran los datos bibliográficos que representan a los atributos del recurso.

El estudio desarrollado por IFLA sobre los requisitos funcionales de los registros bibliográficos (FRBR, por sus siglas en inglés) define al registro bibliográfico como “la suma de datos asociados a las entidades descritas tanto en los catálogos de bibliotecas como en las bibliografías nacionales. Forman parte de este conjunto de datos los elementos de datos descriptivos definidos en las Descripciones Bibliográficas Internacionales Normalizadas (ISBDs)” (IFLA, 1998: 40).

La definición de registro bibliográfico plasmada en FRBR hace énfasis en los datos bibliográficos que se obtienen a partir de la descripción y representación de los atributos de las entidades de información. Desde este instante, el registro bibliográfico es visto como una estructura individual, pero integradora de datos, que tiene propiedades y significados establecidos. Esta variable otorga la posibilidad de estudiar los registros bibliográficos como objetos sustanciales con naturaleza propia y establece una evolución de la aplicación y utilidad del registro en un dominio de información documental.

Por consiguiente, un registro bibliográfico semántico es una estructura de elementos descriptivos y datos bibliográficos que está codificada y representada mediante formatos y esquemas que hacen posible su lectura, intercambio y consulta mediante el uso de una computadora. Un registro de estas características puede vincularse en el ambiente web, pues los datos que contiene están codificados mediante URI, los cuales hacen posible su comunicación mediante el protocolo HTTP. Este registro tiene la facultad de representar los datos que forman parte de los recursos de información tradicionales y aquellos que han nacido como parte de un proceso digital.

Los registros bibliográficos semánticos son parte de un proceso de adaptación global relacionado con el crecimiento exponencial del universo de información y su inherente interacción en el ambiente web, son el producto de sofisticados métodos desarrollados para alcanzar el anhelado control bibliográfico universal y su desarrollo está relacionado con su codificación mediante el uso de tecnologías computacionales semánticas.

De esta manera, las principales funciones del registro bibliográfico semántico en el proceso de recuperación de la información son:

1. Identificar, representar y vincular los datos bibliográficos que constituyen y caracterizan a los recursos de información;
2. Proporcionar los puntos de acceso (autor, título y tema) que propicien la identificación, conexión y acceso al recurso y sus diferentes manifestaciones y expresiones;
3. Propiciar la exploración y obtención de conjuntos de datos abiertos enlazados de índole bibliográfica, que remitan a recursos de información libres de restricciones económicas, legales y técnicas, y
4. Motivar el acceso universal a la información mediante el registro de datos bibliográficos normalizados e interoperables.

“La necesidad de una búsqueda integrada en varios dominios sugiere que nuestros datos deben separarse de los sistemas de bases de datos locales y publicarse en formatos compatibles con la web. La capacidad de acceder sin problemas a la información distribuida reducirá la necesidad de reproducir y actualizar datos redundantes en muchas bibliotecas” (Seeman y Goddard, 2015: 332). Los datos bibliográficos disponibles en los catálogos y sistemas de las bibliotecas deben adaptarse a los formatos interoperables de la web para fomentar su publicación y vinculación con otras fuentes externas al contexto bibliotecario.

De esta manera, la recuperación de información mediante DBAE tendrá un carácter integrador y universal, pues además de identificar los datos bibliográficos que corresponden a los recursos, permitirá la consulta de las vinculaciones que estos datos tienen con otros recursos externos a su contexto. Este tipo de recuperación se construye a partir de métodos gráficos para la consulta e interacción con los datos que son procesados a partir de las descripciones colocadas en el registro bibliográfico semántico.

Grafo bibliográfico

“Los grafos son modelos conceptuales abstractos que representan algunos aspectos del mundo que han sido detectados, observados o inferidos” (Powell y Hopkins, 2015: 9). En un nivel aplicativo, los grafos utilizan nodos para representar a los datos que forman parte de un dominio de información, por ejemplo, una colección o un conjunto de datos bibliográficos con temáticas o tipologías específicas. También emplean aristas para vincular a los nodos, y adquieren un significado específico cuando se les aplica una propiedad que corresponde a un determinado vocabulario o a una ontología. La búsqueda y recuperación de datos bibliográficos mediante la utilización de grafos es expuesta por Zhu y Yan (2016: 1094), quienes explican la posibilidad de “modelar datos bibliográficos utilizando diferentes esquemas, para mostrar la existencia de múltiples relaciones entre dos entidades bibliográficas”. Se propone la generación de interfaces de consulta de grafos mediante la utilización de términos que se relacionen con los datos bibliográficos que caracterizan a los recursos. Anutariya y Dangol (2018) diseñaron una herramienta web para extraer y visualizar conjuntos de datos abiertos enlazados mediante la utilización de grafos que recuperan las consultas diseñadas en lenguaje SPARQL. El reto consiste en diseñar estrategias de consulta de datos bibliográficos abiertos enlazados, que sean intuitivas y que favorezcan al individuo en su demanda y proceso de recuperación de información.

En este sentido, la generación de datos bibliográficos abiertos enlazados es un proceso que puede iniciarse atendiendo la naturaleza de los recursos de información y la identificación de los datos que los caracterizan. Al respecto Godby, Wang y Mixter (2015: 2) manifiestan que “una discusión más concreta puede comenzar con la imagen de un documento que debería ser familiar para los bibliotecarios, editores y buscadores de información, y su reformulación como una red de datos estructurados”. Por ejemplo, la identificación de las manifestaciones, expresiones y ejemplares que se vinculan con una determinada obra intelectual. Las obras son plasmadas en diversos tipos de recursos, por ejemplo, libros, audiovisuales, artículos de investigación, imágenes, documentos, etc. Estos recursos tienen diversas manifestaciones y expresiones y son una influencia intelectual para los individuos que se disponen a crear nueva información. Por ejemplo, la obra *El peregrino kamanita* de Karl Adolph Gjellerup fue publicada originalmente en 1917. A través del tiempo, esta obra ha sido influencia de diversos estilos literarios y adaptada en diferentes manifestaciones y expresiones como audiolibros, libros digitales con ilustraciones, reseñas, etc. Los patrones de influencia entre obras pueden identificarse y analizarse mediante la implementación de DBAE en

dominios de información documental. Por este motivo, DBAE no puede ser considerado sólo un modelo para establecer relaciones sintácticas entre los datos, sino también como una propuesta para identificar las vinculaciones semánticas que caracterizan a estos datos.

Entonces, un grafo bibliográfico representa los datos y las vinculaciones que se establecen en un determinado dominio y pretende explicar cuál es la naturaleza de la vinculación entre ellos y qué significado tienen para el dominio de información al que pertenecen. Esta singularidad ofrece la posibilidad de establecer un proceso de recuperación de información integral y proporcional al contexto en donde se localizan los datos. La aplicación de DBAE en el ámbito de la recuperación de la información no se trata exclusivamente de proponer un método para obtener información. Se relaciona con la necesidad de contar con un modelo interoperable y sistematizado que propicie la identificación de los datos bibliográficos que tienen una vinculación de significado con los recursos y contenidos que proliferan en el ambiente de la web, pues “el enriquecimiento semántico de los datos aumentará su descubrimiento, accesibilidad y reutilización” (Silvello *et al.*, 2015: 169).

Esto nos lleva a considerar que los datos bibliográficos de una obra tienen el potencial de integrar los atributos que los caracterizan para vincularse con diferentes manifestaciones y expresiones. Para que este método de visualización y vinculación sea funcional dentro de un sistema, es necesario contar con una estructura interoperable y compatible con los principios técnicos de Linked Open Data.

Además, los registros bibliográficos semánticos deberán contener los datos bibliográficos principales de las obras. Los registros de las obras fungirán como nodo central para la recuperación de sus diferentes manifestaciones y expresiones dentro de un contexto en específico. Entonces, los puntos de acceso de los registros bibliográficos semánticos serán trascendentales para cumplir con este cometido.

De esta manera, el comportamiento de DBAE en la RI se ejemplifica en la *Figura 3*. En esta representación, el catálogo en línea de una biblioteca es la fuente de los datos bibliográficos correspondientes a la *Iliada* de Homero, los cuales forman parte de la manifestación *libro* de la obra. El registro de esta obra se encuentra en LIBRUNAM. Además, en otra fuente disponible en la web, como Internet Archive, se pueden identificar los datos de la expresión *audiolibro* de la misma obra (https://archive.org/details/iliada_1505_librivox). A su vez, en Youtube se tiene acceso a una videograbación digital que lleva por título *La Ira de Aquiles* (<https://www.youtube.com/watch?v=PoDvcXIuTPo>). Tomando en cuenta esta ejemplificación, se estima que las bases de datos orientadas a grafos permitan realizar consultas federadas

que recuperen registros y en consecuencia recursos disponibles en diversas fuentes, aunque estas se ubiquen en diferentes espacios de la web.

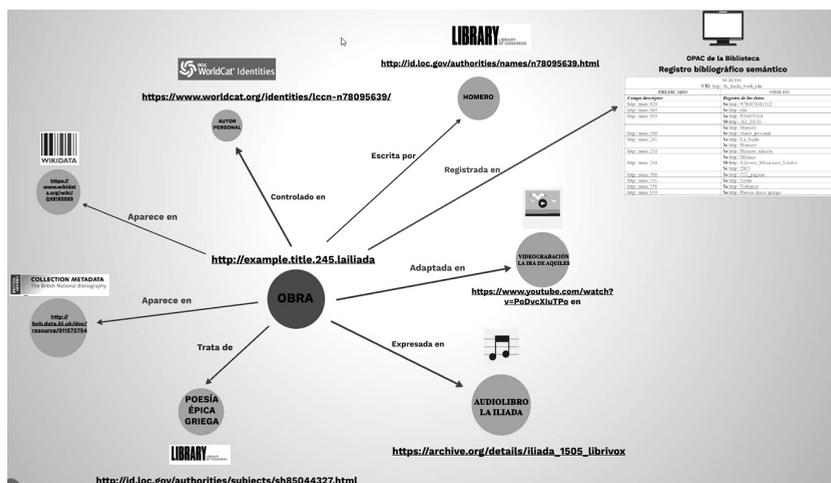


Figura 3. Representación gráfica del comportamiento de DBAE en la recuperación de información
Fuente: elaboración propia, 2019

El nodo central de la obra *Iliada* se vincula con los datos que se localizan en diversas fuentes de la web. Cuando estos datos se conectan tienen la facultad de conformar estructuras de datos con un significado establecido, evidenciando el ambicioso potencial de remitir a recursos y contenidos disponibles en la web.

Obsérvese que las oraciones que representan a las vinculaciones entre los datos son de índole explicativa y tienen el propósito de aumentar la comprensión y precisión de la recuperación de los recursos y contenidos que se han obtenido como parte de una estrategia de búsqueda básica. En este grafo se exponen aquellas fuentes que están disponibles en la web y que contienen datos, recursos y contenidos que se vinculan con los atributos principales de la obra.

Al respecto, existen proyectos consistentes que han liberado datos enlazados de índole bibliográfica, como es el caso de WorldCat, el servicio de datos enlazados de L. C., la Bibliografía Nacional Británica y Wikidata. Sin embargo, ¿cómo pueden vincularse todos estos datos en un entorno web interoperable? ¿De qué manera la vinculación de los datos bibliográficos impacta en la recuperación de información del individuo?

La función principal de DBAE en el proceso de recuperación de la información consiste en ofrecer un modelo visual e interactivo para la consulta de los datos y sus respectivas vinculaciones, que favorezca la identificación de patrones que se establecen entre los datos, contenidos y recursos disponibles

en diferentes dominios de información. En este sentido, “la similitud semántica estima cuantitativa o cualitativamente la intensidad de la vinculación semántica entre unidades de lenguaje, conceptos o instancias, a través de una descripción numérica o simbólica obtenida según la comparación de información que apoya formal o implícitamente su significado” (Harispe *et al.*, 2015: 12). Por lo tanto, las vinculaciones de significado entre los datos bibliográficos disponibles en los registros de las bibliotecas dan la posibilidad de descubrir patrones ocultos que forman parte de la generación de nueva información y de aquella información que ha sido resguardada por largos periodos de tiempo; por ejemplo, conocer la influencia que una determinada obra ha tenido sobre las expresiones y manifestaciones que son plasmadas en los recursos de información, o identificar el curso que una línea de investigación ha tenido a través del tiempo, en función de los productos que se generan como parte del proceso de investigación.

En la actualidad, resulta insuficiente tener acceso a los registros de los datos bibliográficos y a los recursos de información a los que representan, pues la complejidad de las demandas de información de los individuos requiere de consultas sofisticadas que les permitan obtener una visión global e integradora de los registros, recursos y contenidos que se vinculan semánticamente con un dato en particular. El reto consiste en “establecer un nuevo sistema para el acceso y uso de datos bibliográficos que sean compatibles y funcionen dentro de la web” (Coyle, 2013: 57).

Si bien, en la actualidad, los navegadores web permiten consultar la información que está disponible en diversas fuentes y que se relaciona con los términos empleados en las estrategias de búsqueda, la mayoría de los resultados que se obtienen carecen de un impacto en la compleja demanda informativa del individuo, sobre todo si ésta se relaciona con un aspecto académico, científico o para la toma de decisiones.

Las estrategias de búsqueda terminológica fundamentadas en un proceso sintáctico permiten recuperar recursos que contengan dicho término en sus estructuras. Es decir, en los actuales sistemas de información se rastrean las palabras que aparecen en la estructura textual de los recursos o de los registros, dejando de lado el factor significativo del campo de conocimiento al que pertenecen.

De esta manera, el comportamiento de DBAE en la RI se caracteriza por hacer uso de datos bibliográficos estructurados y normalizados, con el propósito de descubrir el significado de los datos y sus respectivas vinculaciones. Se trata de un proceso en el cual intervienen las siguientes consideraciones:

- Para el establecimiento de la recuperación de la información mediante DBAE, es necesario contar con un sistema con características de interoperabilidad e integridad de datos que utilice los principios de Linked Open Data.
- Además, los principios bibliotecológicos para la organización de los recursos de información deben adaptarse y flexibilizarse para la conformación de vocabularios, lenguajes y esquemas que permitan generar registros bibliográficos de índole semántica.
- El grafo bibliográfico es un componente característico y esencial de la recuperación de información mediante la aplicación de DBAE. Se trata de un elemento que establece una notable diferenciación entre los diversos modelos existentes para recuperar información.

Por consiguiente, la generación de aplicaciones para el procesamiento de DBAE en diversos dominios de información documental será un factor trascendental que beneficiará su sistematización. El progreso y adaptación de DBAE en el ámbito de la recuperación de la información requiere de integrar las estructuras normativas con los principios de Linked Open Data, ya que resulta irrelevante contar con enormes cantidades de datos si éstos no se encuentran debidamente estructurados y vinculados.

CONSIDERACIONES FINALES Y FUTURAS INVESTIGACIONES

Con base en la revisión de la literatura y los hallazgos obtenidos, se estima que la complejidad de las demandas informativas de los individuos se encuentra en constante transformación. Los nuevos comportamientos en la búsqueda y recuperación de la información así lo constatan.

La recuperación de información mediante la aplicación de DBAE es un proceso que pone de manifiesto la creación de registros bibliográficos semánticos y grafos bibliográficos, componentes que fomentan de modo integral la recuperación de recursos y contenidos de una manera gráfica, intuitiva e interactiva.

Se ha identificado que el comportamiento de DBAE en la recuperación de información es un proceso que se caracteriza por la explotación de datos bibliográficos estructurados, los cuales tienen la capacidad de establecer vinculaciones semánticas entre recursos y contenidos disponibles en diversas fuentes del ambiente web.

Además, la recuperación de información mediante la sistematización de DBAE será visualmente intuitiva para el individuo que desea consultar de una

manera integradora los datos bibliográficos de una obra conocida universalmente, pues detrás de cada uno de los datos que se presentan existirá un URI asignado que permitirá identificarlos de una manera unívoca y en correspondencia al contexto del grafo y al dominio al que pertenecen.

Por lo tanto, la aplicación de DBAE en el ámbito de la RI requiere de su respetiva sistematización. Bajo esta premisa, será relevante recuperar aquellos datos que son significativos para la demanda informativa de los individuos. En este artículo, desde el punto de vista sistémico, se ha brindado una visión del uso de los datos para apoyar demandas de información relacionadas con el descubrimiento de patrones ocultos en dominios de información documental.

La sistematización de DBAE fomentará el establecimiento de un modelo de recuperación de información caracterizado por la identificación, localización y el descubrimiento de los datos bibliográficos que describen y vinculan a recursos y contenidos con atributos informativos similares. Esto supone la generación de innovadores sistemas de información, los cuales deberán caracterizarse por utilizar principios de interoperabilidad e integridad que fomenten la reutilización de los datos.

Por esta razón, se requieren futuras investigaciones que permitan llevar a cabo la sistematización de DBAE y el desarrollo de aplicaciones descentralizadas que propicien la explotación de los datos bibliográficos abiertos enlazados y su potencial para recuperar información de una manera universal, integradora y libre de restricciones.

REFERENCIAS

- Anderson, Dorothy. 1989. *Standard practices in the preparation of bibliographic records*. London: IFLA, Universal Bibliographic Control and International Marc Programme.
- Anutariya, Chutiporn y Reshma Dangol. 2018. "VizLOD: Schema extraction and visualization of linked open data". Trabajo presentado en la 15th International Joint Conference on Computer Science and Software Engineering, Thailand, 11-13 de julio.
- Blair, David. 1990. *Language and representation in information retrieval*. Netherlands: Elsevier.
- Bermès, Emmanuelle. 2015. "Following the user's flow in the digital Pompidou", en *Linked data and user interaction*, editado por Frank Cervone y Lars Svensson, 19-30. Germany: De Gruyter.
- Berners-Lee, Tim. 2006. "Linked Data".
<https://www.w3.org/DesignIssues/LinkedData.html>
- Berners-Lee, Tim. 2010. "Linked Data" (fragmento de la página agregado en 2010).
<https://www.w3.org/DesignIssues/LinkedData.html>
- Chu, Heting. 2003. *Information representation and retrieval in the digital age*. New Jersey: Information Today.

- Creative Commons. S. a. "Open Data".
<https://creativecommons.org/about/program-areas/open-data/>
- Cristovao, Henrique y Jorge Fernandes. 2018. "Recuperação de informação em dados ligados: um modelo baseado em mapas conceituais e análise de redes complexas". *Transinformação* 2 (30): 193-207.
<http://doi.org/10.1590/2318-08892018000200005>
- Coyle, Karen. 2013. "Library linked data: an evolution". *Italian Journal of Library and Information Science* 1 (4): 53-61.
<http://doi.org/10.4403/jlis.it-5443>
- Dunsire, Gordon. 2014. "The Role of ISBD in the Linked Data Environment". *Cataloging & Classification Quarterly* 8 (52): 855-868.
<https://doi.org/10.1080/01639374.2014.933753>
- Elsayed, Esraa y Saleh Mesbah. 2018. "Proposal for Implementing Linked Open Data on Libraries Catalogue". Trabajo presentado en la 84th IFLA General Conference and Assembly, Kuala Lumpur, Malaysia, 24-30 de agosto.
- Faith, Asleight y Michelle Chrzanowski. 2015. "Connecting RDA and RDF Linked Data for a Wide World of Connected Possibilities". *Pennsylvania Libraries: Research & Practice* 2 (3): 122-135.
<https://doi.org/10.5195/palrap.2015.106>
- Giusti Serra, Liliana y José Santarém Segundo. 2017. "O catálogo da biblioteca e o linked data". *Em Questão* 2 (23): 167-185.
<http://doi.org/10.19132/1808-5245232.167-185>
- Godby, Jean, Shenghui Wang y Jeffrey K. Mixer. 2015. *Library Linked Data in the Cloud: OCLC's Experiments with New Models of Resource Description*. United States of America: Morgan & Claypool.
- Hagler, Ronald y Peter Simmons. 1982. *The Bibliographic record and information technology*. United States of America: ALA.
- Harispe, Sébastien, Sylvie Ranwez, Stefan Janaqi y Jacky Montmain. 2015. "Semantic measures for the comparison of units of language, concepts or instances from text and knowledge representation analysis". *Synthesis Lectures on Human Language Technologies* 1 (8): 1-254.
<http://doi.org/10.2200/S00639ED1V01Y201504HLT027>
- Haynes, David. 2004. *Metadata for information management and retrieval*. London: Facet Publishing.
- Ingwersen, Peter. 1992. *Information retrieval interaction*. London: Taylor Graham.
- IFLA (International Federation of Library Associations and Institutions). 1998. *Requisitos funcionales de los registros bibliográficos: Informe final*. España: Ministerio de Cultura, Sociedad Anónima de Fotocomposición.
- Jai O'Dell, Allison. 2016. "Authority control for the web: integrating library practice with Linked Data", en *Linked Data for Cultural Heritage*, editado por Ed Jones y Michele Seikel, 41-53. Estados Unidos de América: ALA.
- Landis, Cliff. 2019. "Linked open data in libraries", en *New top technologies every librarian needs to know*, editado por K. J. Varnum, 3-15. Chicago: ALA Neal-Schuman.
- Le Boeuf, Patrick. 2015. "Customized OPACs on the Semantic Web: The OpenCat Prototype", en *Linked data and user interaction*, editado por Frank Cervone y Lars Svensson. 31-47. Germany: De Gruyter.

- Martínez Arellano, Filiberto Felipe. 2012. "Retos de la catalogación y de los catálogos". *Investigación bibliotecológica* 58 (26): 7-12.
<http://doi.org/10.22201/iibi.0187358xp.2012.58.35263>
- Powell, James y Matthew Hopkins. 2015. *A librarian's guide to graphs, data and the semantic web*. United States of America: Chandos.
- Rietveld, Laurens. 2016. *Publishing and consuming Linked Data: optimizing for the unknown*. Berlin: AKA.
- Salton, Gerard y Michael McGill. 1983. *Introduction to modern information retrieval*. United States of America, McGraw-Hill.
- Salvador Oliván, José Antonio y Rosario Arquero Avilés. 2006. "Una aproximación al concepto de recuperación de información en el marco de la ciencia de la documentación". *Investigación bibliotecológica* 41 (20): 13-43.
<http://doi.org/10.22201/iibi.0187358xp.2006.41.4113>
- Seeman, Dean y Lisa Goddard. 2015. "Preparing the way: creating future compatible cataloging data in a transitional environment". *Cataloging & Classification Quarterly* 3/4 (53): 331-340.
<https://10.1080/01639374.2014.946573>
- Silvello, Gianmaria, Georgeta Bordea, Nicola Ferro, Paul Buitelaar y Toine Bogers. 2015. "Semantic representation and enrichment of information retrieval experimental data". *International Journal of Digital Libraries* 18: 145-172.
<http://0.1007/s00799-016-0172-8>
- Svenonius, Elaine. 2018. "Bibliographic Entities and their Uses". *Cataloging & Classification Quarterly* 8 (56): 711-724.
<https://doi.org/10.1080/01639374.2018.1524284>
- Talleras, Kim. 2017. "Quality of linked bibliographic data: the models, vocabularies, and links of data sets published by four national libraries". *Journal of library metadata* 2 (17): 126-155.
<https://www.tandfonline.com/doi/full/10.1080/19386389.2017.1355166>
- Taniguchi, Sochi. 2018. "Mapping and Merging of IFLA Library Reference Model and BIBFRAME 2.0". *Cataloging & Classification Quarterly* 5/6 (56): 427-454.
<https://doi.org/10.1080/01639374.2018.1501457>
- Ullah, Irfan, Shah Khusro, Asim Ullah y Muhammad Naeem. 2018. "An Overview of the Current State of Linked and Open Data in Cataloging". *Information Technology and Libraries* 4 (37): 47-80.
<https://doi.org/10.6017/ital.v37i4.10432>
- W3C. 2011. "Informe Final del Grupo Incubador de Datos Vinculados de Bibliotecas". http://www.larramendi.es/LAM/Incubator/lld/XGR-lld-20111025.html#Rights_ownership_is_complex
- Warner, Julian. 2010. *Human information retrieval*. Cambridge: MIT Press.
- Wenige, Lisa y Johannes Ruhland. 2018. "Retrieval by recommendation: using LOD technologies to improve digital library search". *International Journal on Digital Libraries* 2/3 (19): 253-269.
<https://doi.org/10.1007/s00799-017-0224-8>
- Zhu, Yongjun y Erjia Yan. 2016. "Searching bibliographic data using graphs: A visual graph query interface". *Journal of Informetrics*, 10: 1092-1107. doi:
<https://doi.org/10.1016/j.joi.2016.09.005>

Para citar este texto:

Ávila Barrientos, Eder. 2020. “Los datos bibliográficos abiertos enlazados y su comportamiento en la recuperación de información”. *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 34 (82): 203-227.

<http://dx.doi.org/10.22201/iibi.24488321xe.2020.82.58130>

DOI: <http://dx.doi.org/10.22201/iibi.24488321xe.2020.82.58130>