

## Vinculación de contenidos en repositorios institucionales con perfiles semánticos de metadatos

Georgina Araceli Torres Vargas<sup>1</sup>; Eder Ávila Barrientos<sup>2</sup>

Recibido: 7 octubre de 2020 / Aceptado: 19 octubre de 2020

**Resumen.** El objetivo de este trabajo consiste en analizar el proceso para la vinculación de contenidos en repositorios institucionales mediante el uso de perfiles semánticos de metadatos. Se estima que la relación de contenidos en el contexto digital pone de manifiesto una serie de problemáticas, pues la práctica que se requiere para alcanzar este objetivo es la construcción de redes de colaboración a través de plataformas digitales. A simple vista los repositorios institucionales se muestran como el mecanismo idóneo para implementar dichas redes, sin embargo, esto no ha sido posible, ya que los recursos de información documental de que se dispone en el entorno digital cuentan con esquemas de metadatos que cumplen la función básica de describir y representar las características del recurso y se requiere de perfiles semánticos de metadatos. Este es el aspecto que se trata en el presente texto. Aunado a ello, la vinculación de contenidos en repositorios institucionales mediante perfiles semánticos de metadatos se trata de un proceso analítico intelectual que tiene como propósito principal ofrecer un mecanismo para descubrir los atributos que se encuentran relacionados significativamente en un contexto digital. En los repositorios institucionales, se encuentra el ambiente idóneo para la aplicación de estos perfiles, pues son sistemas utilizados por comunidades científicas y académicas para divulgar y desarrollar sus procesos de investigación.

**Palabras clave:** repositorios institucionales; metadatos; datos enlazados, contenidos digitales.

### [en] Linking contents in institutional repositories with semantic metadata profiles

**Abstract.** The objective of this work is to analyze the process for linking content in institutional repositories through the use of semantic metadata profiles. It is estimated that the content relationship in the digital context reveals a series of problems, since the practice required to achieve this objective is the construction of collaboration networks through digital platforms. At first glance, institutional repositories appear to be the ideal mechanism to implement these networks, however, this has not been possible, since the documentary information resources available in the digital environment have metadata schemes that fulfill the function basic description and representation of the characteristics of the resource and requires semantic metadata profiles. This is the aspect that is dealt with in this text. In addition to this, the linking of contents in institutional repositories through semantic metadata profiles is an intellectual analytical process whose main purpose is to offer a mechanism to discover the attributes that are significantly related in a digital context. In the institutional repositories, there is the ideal environment for the application of these profiles, as they are systems used by scientific and academic communities to disseminate and develop their research processes.

**Keywords:** Institutional repositories; metadata; linked data; digital contents.

**Sumario.** 1. Introducción. 2. Linked Data. 3. Repositorios institucionales. 4. Vinculación de contenidos en el entorno digital. 5. Perfiles semánticos de metadatos. 6. Discusión. 7. Conclusiones. Referencias bibliográficas.

**Cómo citar:** Torres Vargas, G. A.; Ávila Barrientos, E. (2020). Vinculación de contenidos en repositorios institucionales con perfiles semánticos de metadatos. *Cuadernos de Documentación Multimedia*, 31, e72157. <http://dx.doi.org/10.5209/cdmu.72157>

### 1. Introducción

En el campo de la bibliotecología y los estudios de la información, es escasa la generación de estudios que aborden la temática de los perfiles semánticos de metadatos

para relacionar los recursos de información documental integrados en los repositorios institucionales.

La relación entre metadatos y repositorios institucionales pone de manifiesto una serie de consideraciones procedimentales, técnicas y prácticas que señalan la ma-

<sup>1</sup> Investigadora. Instituto de Investigaciones Bibliotecológicas y de la Información. Universidad Nacional Autónoma de México. Correo electrónico: [gatv@unam.mx](mailto:gatv@unam.mx)

<sup>2</sup> Investigador. Instituto de Investigaciones Bibliotecológicas y de la Información. Universidad Nacional Autónoma de México. Correo electrónico: [eder@iibi.unam.mx](mailto:eder@iibi.unam.mx)

nera de efectuar determinados procesos para representar los datos descriptivos que caracterizan a los recursos de información documental mediante la utilización de metadatos. En este sentido, el esquema de metadatos que es utilizado con mayor frecuencia es Dublin Core.

Dunsire (2008:52), manifiesta que la cosecha de metadatos disponibles en repositorios institucionales otorga la posibilidad de generar servicios de información agregados, ya que mediante la descripción e identificación de recursos es viable ofrecer al usuario un método automático para la disseminación selectiva de la información.

La compatibilidad entre atributos del recurso y los elementos de metadatos se establece a partir de un análisis descriptivo del recurso de información digital. De esta manera, el objetivo de este trabajo recae en formular un análisis descriptivo de los elementos que permitan la vinculación de contenidos en repositorios institucionales, mediante perfiles semánticos de metadatos. La vinculación de contenidos en repositorios institucionales mediante perfiles semánticos de metadatos es un proceso de análisis intelectual que basa su desarrollo en la aplicación de principios normativos. La posibilidad de obtener una descripción significativa del recurso a través de los perfiles de metadatos recae en la identificación de los atributos descriptivos del recurso y su respectiva vinculación con los elementos de metadatos disponibles en un determinado esquema.

“El uso de datos enlazados e identificadores globales para describir a los objetos de los repositorios, se ha convertido gradualmente en una solución más atractiva para visibilizar la investigación académica del silo de los repositorios institucionales (Downey, 2019: 133)”.

Bajo esta premisa, la vinculación de contenidos a través de los perfiles semánticos de metadatos sugiere la generación de servicios de información digitales, que tendrían la posibilidad de apoyar los procesos de investigación. Además, “mediante la extracción y lectura de metadatos es posible personalizar su búsqueda e identificar nuevas formas de conexiones mediante el establecimiento de grafos semánticos (Russo, 2015: 40)”.

Por lo tanto, el objetivo de este trabajo consiste en analizar el proceso para la vinculación de contenidos en repositorios institucionales mediante el uso de perfiles semánticos de metadatos. Como metodología, se ha utilizado a la hermenéutica y análisis del discurso aplicados a la literatura especializada en el objeto de estudio de la investigación. La relación de contenidos en el contexto digital pone de manifiesto una serie de problemáticas. La práctica que se requiere para alcanzar este objetivo es la construcción de redes de colaboración a través de plataformas digitales.

## 2. Linked Data

La web semántica es una web extendida y dotada de mayor significado, apoyada en lenguajes universales, que van a permitir que los usuarios puedan encontrar respuestas a sus preguntas de forma más rápida y sencilla

gracias a una información mejor definida. Al dotar a la web de más significado y por lo tanto, de más semántica, se pueden obtener soluciones a problemas habituales en la búsqueda y recuperación de información gracias a la utilización de una infraestructura común mediante la cual, es posible compartir, procesar y transferir información de forma sencilla.

Linked Data, conocido también como datos enlazados, se trata de un método para la creación, publicación y vinculación de datos estructurados que estén disponibles en diversas fuentes de la web. Linked Data forma parte de las principales propuestas para la conformación de una web semántica, y engloba a una serie de principios que explican el proceso para la generación de datos enlazados.

En este sentido, Berners-Lee (2009: párr. 3) definió cuatro principios básicos para la publicación de Linked data:

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

De esta manera, es viable publicar datos que cumplan sólo los tres primeros principios, pero el hecho de no aplicar el cuarto los convierte en menos visibles y, como consecuencia, menos reutilizables. Linked Data forma parte esencial de la concepción de la Web semántica, por lo tanto, posee una serie de componentes que permiten la construcción de este espacio digital. A continuación, se describen estos elementos:

- **XML.** Sintaxis superficial para documentos estructurados.
- **XML. Schema.** Lenguaje para definir la estructura de los documentos XML.
- **RDF.** Modelo de datos para los recursos y las relaciones que se puedan establecer entre ellos.
- **RDF Schema.** Vocabulario para describir las propiedades y las clases de los recursos RDF con una semántica para establecer jerarquías de generalización entre dichas propiedades y clases.
- **SPARQL.** Lenguaje de consulta sobre RDF que permite hacer búsquedas sobre los recursos de la Web Semántica utilizando distintas fuentes de datos.
- **OWL. Web Ontology Language.** Añade más vocabulario para describir propiedades y clases como: relaciones entre clases, cardinalidad, igualdad, tipologías de propiedades más complejas, caracterización de propiedades o clases enumeradas.
- **SKOS.** Simple Knowledge Organization System. Sistema simple de organización del cono-

cimiento. Se trata de un modelo para la representación de la estructura básica y el contenido de esquemas de conceptos como tesauros, esquemas de clasificación, listas de encabezamientos de materia, taxonomías, folksonomías y otros vocabularios controlados similares. Al tratarse de una aplicación de RDF, SKOS permite la creación y publicación de conceptos en la Web, así como vincularlos con datos en este

mismo medio e incluso integrarlos en otros esquemas de conceptos.

- **Triples RDF.** La construcción de triples con RDF se basa en la idea de declarar recursos usando la expresión en la forma sujeto-predicado-objeto. Esta expresión se conoce en la terminología RDF como triple o triplete. Un triplete RDF contiene tres componentes, todos con referencia en un URI:

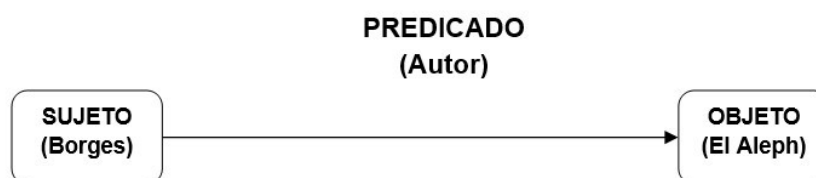


Figura 1. Representación gráfica de un triplete  
Fuente: elaboración propia, 2019.

- **Sujeto.** Se trata de una referencia URI, puede ser una persona, o un nodo, es el ente al cual nos referimos principalmente;
- **Predicado.** Es la propiedad o relación que se desea establecer acerca del sujeto;
- **Objeto.** Es el valor de la propiedad o del otro recurso con el que se establece la relación.

El uso de Linked Data hace posible publicar información migrada desde diferentes fuentes en las que se encuentra almacenada, como por ejemplo bases de datos, repositorios, hojas de cálculo, catálogos en línea, siempre y cuando se utilice el lenguaje RDF, para permitir enlazar esta información con otros datos, de forma que aumente la conectividad entre dicha información y se facilite así la búsqueda, recuperación y la agregación de datos estructurados.

Los principios de Linked Data podrían aplicarse a la estructura de los repositorios institucionales, mediante una arquitectura que contemple los principios señalados anteriormente y la utilización de las normas y lenguajes computacionales que tiene como objetivo la generación de datos enlazados. La vinculación de contenidos mediante la generación de datos enlazados pone de manifiesto el seguimiento de estos principios y de la normatividad identificada.

### 3. Repositorios institucionales

Los repositorios institucionales, recogen la producción intelectual de una institución universidades, es decir, publicaciones de carácter académico. El repositorio institucional en su concepción inicial se muestra como una colección digital, cuyo objetivo es concentrar y preservar la producción intelectual de una comunidad universitaria o de investigación, la cual está conformada por académicos, investigadores y alumnos (Jones; McColl, 2006: 11).

Un repositorio institucional es relevante en tanto que pretende contener los productos de la vida académica e intelectual de las universidades, que se genera cada

vez en mayor medida en formato digital. Para Lynch es una responsabilidad primordial de las universidades administrar estas riquezas: tanto hacerlas accesibles como preservarlas. (Lynch, 2003).

Tomando en cuenta las diversas posturas conceptuales de un repositorio institucional, puede considerarse como un sistema de información digital que almacena, organiza, busca, recupera, accede, preserva y comparte recursos de información documental en formato digital. En la actualidad hay una tendencia por poner en acceso abierto la producción de las universidades, aunque también existen repositorios de acceso restringido, sobre todo de aquellos que contienen datos de investigaciones en proceso y que todavía no pueden compartirse más que entre los investigadores participantes.

Asimismo, un aspecto sumamente trascendental para los repositorios se refiere a su capacidad para comunicarse o ser interoperables con otros, y de esta forma asegurar el intercambio de información digital, no solo entre los usuarios de un repositorio, sino entre usuarios de numerosos repositorios alojados en diferentes partes del mundo esto es, a manera de repositorios distribuidos.

La cosecha de metadatos en repositorios es una práctica habitual para el intercambio de información digital. Esta puede definirse como un proceso de agregación de registros de metadatos de múltiples proveedores en una única base de datos.

La cosecha de metadatos utiliza el protocolo OAI-PMH para establecer la interoperabilidad entre repositorios. De esta manera el intercambio de información digital se lleva a cabo a través de una comunicación entre repositorios con propósitos afines. Por lo tanto, es una alternativa ideal para el intercambio de información digital en repositorios. Esta cosecha puede remitir a los metadatos de un determinado registro o bien proporcionar el acceso en forma al recurso de información digital.

En el entorno de los repositorios institucionales distribuidos, es interesante plantear la cosecha de datos que puedan enlazarse y reflejar relaciones entre los contenidos académicos generados por diversas instituciones de educación superior, tanto nacionales como internacionales.

#### 4. Vinculación de contenidos en el entorno digital

La construcción de relaciones de conocimiento en el entorno digital es un fenómeno derivado de la implementación de las tecnologías de la información y comunicación en las actividades académicas, científicas y profesionales de la sociedad.

En la actualidad, el entorno digital está caracterizado por la utilización de dispositivos inteligentes y aparatos digitales que permiten generar datos e información en cantidades enormes. Gran parte de esta información se encuentra en recursos de información documental que son puestos a disposición de diferentes comunidades de usuarios mediante distintos medios como repositorios institucionales, bibliotecas digitales y plataformas multimedia.

Un entorno digital se puede concebir como un espacio representado por entidades digitales de información, en donde se cuenta con servicios, aplicaciones y una amplia gama de plataformas que permiten al usuario identificar la información que necesita. El ejemplo más concreto de entorno digital se encuentra en la *World Wide Web*.

Actualmente, la web está conformada por enormes cantidades de información distribuida en diferentes espacios y fuentes. Mucha de esta información es documental y se encuentra disponible en:

- Bibliotecas digitales
- Repositorios
- Descubridores de información
- Portales temáticos
- Bases de datos
- Servicios web y aplicaciones

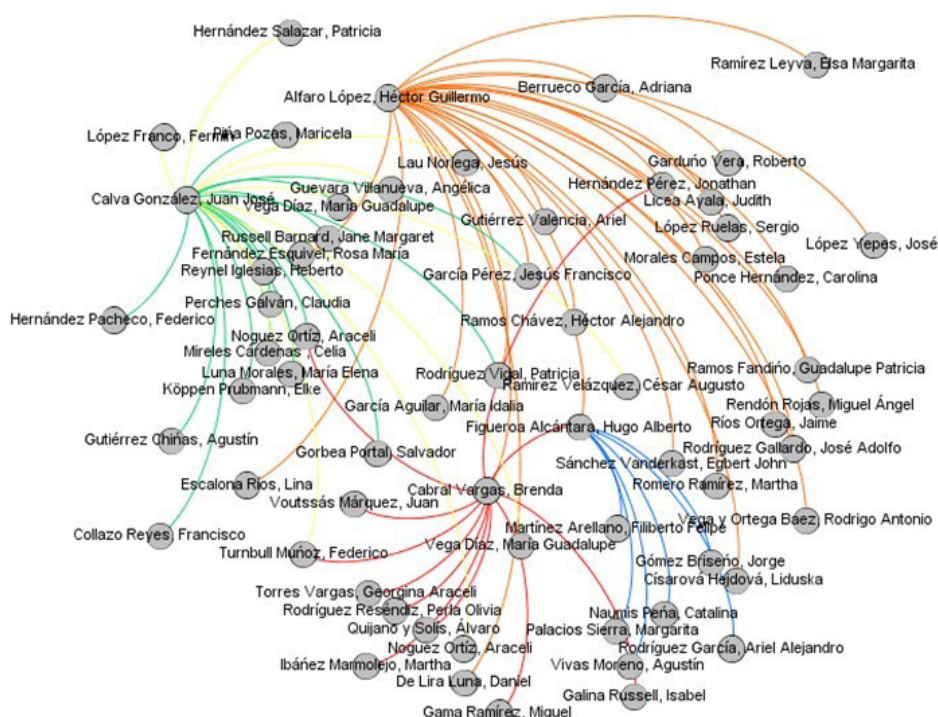
Las redes de colaboración científica y académica utilizan los medios digitales para compartir la información que generan en diferentes contextos, pero con intereses comunes. De hecho, la vinculación de conocimientos a través de estas redes es un proceso común que hace posible establecer la comunicación científica entre comunidades de expertos.

En este contexto, las publicaciones científicas y académicas se caracterizan por contar con un aparato crítico que sustenta su discurso teórico plasmado en sus investigaciones<sup>3</sup>. Este aparato crítico también cumple con la función de relacionar conocimientos afines entre especialistas de una misma disciplina o un campo del conocimiento en común. El análisis de los aparatos críticos permite identificar los patrones de comportamiento en la gestación de vínculos de conocimiento.

La vinculación de conocimientos en el ambiente digital se presenta como un fenómeno cada vez más multidisciplinario. En este contexto, el análisis de redes de conocimiento permite descubrir con detalle los tipos de relaciones que se establecen entre los componentes de un determinado modelo de comunicación científica. Estos componentes pueden ser personas, obras, temas, manifestaciones, expresiones y tipos de recursos de información.

La vinculación de datos en el entorno digital es un proceso que permite identificar las relaciones que se establecen entre diferentes autores y los temas de interés que comparten.

Un tema que guarda estrecha relación con la conformación de datos enlazados es el método de redes sociales. Este método deviene del área de las ciencias sociales y permite definir relaciones entre actores, los nodos que se establecen en tales relaciones y la cercanía o lejanía entre las mismas. Las redes sociales se pueden observar más claramente mediante grafos.



En bibliotecología y estudios de la información, el método de redes sociales ha sido muy utilizado para aná-

<sup>3</sup> El aparato crítico es el conjunto de notas y citas que complementan al texto en cuestión.



lisis de corte bibliométrico, pero puede ser útil en el establecimiento de datos enlazados.

Con ayuda del método de redes sociales, es posible observar la estructura de datos y sus relaciones (Hanneman, 2001:16). A continuación, se ejemplifica con las

relaciones que se dan entre profesores del Posgrado en Bibliotecología y Estudios de la Información de la Universidad Nacional Autónoma de México, a través de su participación en la dirección de tesis de maestría y doctorado, entre 2014-2018.

```
<?xml version="1.0" encoding="UTF-8"?>
<metadata
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/">
  <dc:title>Biblioteca digital académica en bibliotecología y estudios de la información</dc:title>
  <dc:creator>Torres Vargas, Georgina Araceli</dc:creator>
  <dc:subject>Bibliotecas digitales académicas</dc:subject>
  <dc:subject>Recursos de información electrónicos</dc:subject>
  <dc:subject>Tecnología de la información</dc:subject>
  <dcterms:abstract>Con el Proyecto papit it 400312 "Biblioteca Digital en Bibliotecología y Estudios de la Información" (bidei), del cual deriva esta publicación, se pretende delimitar las pautas para el desarrollo de una biblioteca digital que albergue colecciones digitales para la comunidad de investigadores en el área de bibliotecología y estudios de la información del Instituto de Investigaciones Bibliotecológicas y de la Información (iibi), de la Universidad Nacional Autónoma de México. En este proyecto participan investigadores nacionales e internacionales así como estudiantes del posgrado en Bibliotecología y Estudios de la Información de la UAM. En la primera fase del proyecto se realizó un acercamiento a los aspectos que se involucran en la constitución de la biblioteca digital que se pretende concretar. De acuerdo con las características de la comunidad a la que servirá, la biblioteca digital en cuestión será una biblioteca digital académica perfilada al apoyo de lo que hoy se conoce como la e-ciencia. Las políticas de información, las colecciones, los servicios y las formas de acceso a los contenidos se delimitan en este contexto.</dcterms:abstract>
  <dc:publisher>UNAM</dc:publisher>
  <dc:publisher>Instituto de Investigaciones Bibliotecológicas y de la Información</dc:publisher>
  <dc:contributor>Georgina Araceli Torres Vargas</dc:contributor>
  <dc:contributor>Ariel Alejandro Rodríguez García</dc:contributor>
  <dc:contributor>Ana Lúcia Terra</dc:contributor>
  <dc:contributor>Miguel C. Gama Ramírez</dc:contributor>
  <dc:contributor>Adriana Monroy Muñoz</dc:contributor>
  <dc:contributor>Francisco José Valentín Ruiz</dc:contributor>
  <dc:contributor>Juan Miguel Sánchez Vigil</dc:contributor>
  <dc:contributor>Eder Ávila Barrientos</dc:contributor>
  <dcterms:created>2013</dcterms:created>
  <dc:type xsi:type="dcterms:DCMIType">Monografía de investigación</dc:type>
  <dc:type xsi:type="dcterms:DCMIType">Text</dc:type>
  <dc:format>Digital</dc:format>
  <dc:format xsi:type="dcterms:INT">application/pdf</dc:format>
  <dc:identifier>http://bit.ly/2rb0e8X</dc:identifier>
  <dc:language xsi:type="dcterms:ISO639-2">spa</dc:language>
  <dc:language xsi:type="dcterms:ISO639-2">por</dc:language>
  <dc:relation>http://iibi.unam.mx/publicaciones/281/biblioteca%20digital%20academica%20en%20bibliotecologia%20y%20estudios%20de%20la%20informacion%20torres%20vargas%20georgina%20araceli.html</dc:relation>
  <dc:relation>http://iibi.unam.mx/publicaciones/281/biblioteca%20digital%20academica%20en%20bibliotecologia%20politicas%20ana%20lucia%20terra.html</dc:relation>
  <dc:relation>http://iibi.unam.mx/publicaciones/281/biblioteca%20digital%20academica%20en%20bibliotecologia%20desarrollo%20miguels%20gama%20ramirez.html</dc:relation>
  <dc:relation>http://iibi.unam.mx/publicaciones/281/biblioteca%20digital%20academica%20en%20bibliotecologia%20fotografia%20adriana%20monroy%20munoz.html</dc:relation>
  <dc:relation>http://iibi.unam.mx/publicaciones/281/biblioteca%20digital%20academica%20en%20bibliotecologia%20del%20libro%20electronico%20francisco%20jose%20valentin%20ruiz.html</dc:relation>
  <dc:relation>http://iibi.unam.mx/publicaciones/281/biblioteca%20digital%20academica%20en%20bibliotecologia%20linked%20open%20data%20eder%20avila%20barrientos.html</dc:relation>
  <dc:rights>http://creativecommons.org/licenses/by-nc-sa/3.0/deed.es_MX</dc:rights>
  <dcterms:educationlevel>Nivel superior</dcterms:educationlevel>
</metadata>
```

Figura 2. Representación de las relaciones entre tutores del posgrado en Bibliotecología y Estudios de la Información. Elaboración propia, 2019

A simple vista este grafo no da muchos elementos de análisis, porque tan solo refleja la compartición de trabajo de dirección de tesis. En el caso de los repositorios sería necesario enlazar datos entre profesores y tesis; tesis y temas de investigación; profesores con otros profesores que hayan escrito artículos en conjunto, entre otras reacciones. En el contexto de la web, las vinculaciones entre datos es un fundamento que pretende establecer relaciones significativas entre los datos disponibles en diferentes contextos, es decir, mediante la idea semántica. Berners Lee, Hendler y Lassila (2001: 5), han manifestado que el funcionamiento de la web semántica requiere de computadoras que tengan acceso a colecciones estructuradas de información, y a conjuntos de reglas de inferencia que les permitan utilizar dicha información a través de un razonamiento automatizado. Hasta ahora eso no ha sido posible.

Lim (2010: 61) considera que, para materializar la idea de la web semántica, las ontologías web deberían ser cuidadosamente definidas considerando la complejidad de los temas que se tratan en los contenidos dispuestos en la web. La web semántica, no es sólo una cuestión tecnológica; por el momento implica también la intervención del hombre y su razonamiento, ya que implica la interpretación de contenidos. Es por eso por lo que la estructura semántica de la información web se cambia y se desarrolla sobre la base de la interacción social establecida entre los usuarios de Internet.

En este contexto, las ontologías web juegan un papel trascendental al momento de otorgar un significado conceptual, temático y semántico a los datos que representan el conocimiento plasmado y contenido en los

recursos generados por comunidades científicas y académicas.

El término ontología, al igual que el término concepto, se ha definido desde diferentes puntos de vista. Las ontologías pueden ser vistas como mediadoras en la representación del conocimiento por medio de conceptos (Jakus et al., 2013: 29). La representación de conocimientos por medio de ontologías en el entorno digital utiliza estándares y normas para estructurar los contenidos que formarán parte de la ontología que representará a un determinado contexto de conocimientos.

W3C (*World Wide Web Consortium*), ha estandarizado una serie de vocabularios de ontología que poseen las ventajas de los formalismos de representación de conocimiento y los métodos para el modelado conceptual utilizado en bases de datos. Tales actividades de estandarización fomentaron la creación de una nueva ontología para la conformación de la web semántica (Ding et al., 2007: 81).

En la web semántica, los vocabularios definen los conceptos y relaciones (también denominados "términos") utilizados para describir y representar un área de conocimiento. Los vocabularios se utilizan para clasificar los términos que se pueden utilizar en una aplicación particular, caracterizar relaciones posibles y definir posibles restricciones en el uso de dichos términos (W3C, 2015: párr. 1).

Por lo tanto, vocabularios y ontologías forman parte de los componentes básicos de la web semántica y en su aplicación permiten vincular datos con significado para el usuario final. Así pues, en el ámbito de la web semántica, "no existe una división clara entre

lo que se conoce como “vocabularios” y “ontologías”. La tendencia es utilizar la palabra “ontología” para una colección de términos más compleja y bastante formal, mientras que el “vocabulario” se usa cuando tal formalismo estricto no se utiliza o sólo en un sentido muy elemental. Los vocabularios son los componentes básicos de las técnicas de inferencia en la Web Semántica” (W3C, 2015: párr. 2).

Cuando se refiere a los principios de vocabulario y ontología en el contexto de la web semántica, es necesario tomar en cuenta la presencia de la normalización de la información. Si en los inicios de la web se pensó que la normalización terminológica ya no se requería, hoy se hace presente en el ámbito de las ontologías y es un aspecto sumamente importante para el desarrollo de vinculaciones de contenido en el entorno digital.

La normalización de la información en el contexto digital pone de manifiesto la utilización de normas, principios y estándares que permitan desarrollar de manera eficiente el proceso de búsqueda y recuperación de información digital.

En el caso de los repositorios institucionales, se ha optado por utilizar esquemas de metadatos para organizar y propiciar el acceso a los recursos de información que coleccionan.

La mayoría de la información disponible en la Web carece de una normalización que permita su eficiente acceso. Por este motivo, los repositorios institucionales han implementado estrategias como los metadatos para describir, representar, recuperar y acceder a los recursos digitales de información que albergan.

Los principios de la normalización de la información digital engloban una serie de fundamentos que en su conjunto pueden ser aplicados en un proceso intelectual y sistemático, a través de las siguientes etapas:

1. Análisis del recurso de información
  - a. Identificación de atributos
  - b. Representación de atributos (modelado de metadatos)
  - c. Descripción de atributos (bibliográfica y temática)
2. Sistematización
  - a. Definición de sistema de información
  - b. Implementación del sistema
  - c. Evaluación de la sistematización
  - d. Usabilidad
  - e. Interoperabilidad
  - f. Accesibilidad
3. Acceso a la información documental
  - a. Evaluación del proceso de búsqueda y recuperación de información en el sistema digital.

Por lo tanto, la normalización de la información involucra la utilización de estándares, normas y principios que se reúnen en aparatos normativos como manuales, reglas, listas, vocabularios y herramientas que tienen como objetivo atender un aspecto específico como la catalogación, indización, elaboración

de resúmenes, clasificación, control de autoridades y tratamiento documental del recurso de información.

La normatividad general y específica de un repositorio institucional, se plasma en un instrumento integral identificado como políticas. La política de un repositorio institucional debe mencionar los principios de organización, recuperación, búsqueda y acceso a los recursos documentales contenidos en el repositorio.

## 5. Perfiles semánticos de metadatos

Los perfiles descriptivos de metadatos son estructuras que permiten representar los atributos bibliográficos e intelectuales de los recursos de información documental, a través de un conjunto de metadatos descriptivos previamente establecidos. Estos perfiles son el resultado del análisis de los atributos que caracterizan a dichos recursos.

En este sentido, cada recurso de información documental es susceptible de contar con un perfil de metadatos en específico, el cual permita su descripción y representación en el entorno digital mediante la funcionalidad que un repositorio institucional puede proporcionar para cumplir estos objetivos. Berners-Lee (1997: párr.4), considera que los metadatos son información inteligible para la computadora sobre recursos web u otras cosas. En un amplio sentido los metadatos son datos que permiten representar los atributos de un recurso de información, para su posterior búsqueda, recuperación, acceso y vinculación en diferentes plataformas disponibles en el entorno digital.

Los principios de la aplicación de perfiles de metadatos son lineamientos que deben tomarse en cuenta para la descripción de los recursos de información que estarán almacenados en el repositorio. Estos principios toman en consideración los siguientes elementos:

- Tipo de recurso de información que se va a describir.
- Esquema de metadatos utilizado para describir el recurso.
- Nivel de descripción del recurso.
- Análisis del recurso de información
- Identificación de los atributos del recurso (bibliográficos, descriptivos, de autoría, de acceso, específicos).
- Descripción de los atributos del recurso mediante la utilización del esquema de metadatos elegido.
- Revisión de la descripción efectuada.
- Evaluación de la descripción.

Los metadatos descriptivos tienen como objetivo la descripción e identificación de recursos digitales de información. Proporcionan elementos para la organización, descripción, identificación, búsqueda, recuperación y acceso al recurso de información.

Tabla I. Elementos Dublin Core para la conformación del perfil semántico de metadatos.

Fuente: elaboración propia, 2020.

Elementos Dublin Core para la conformación del perfil semántico de metadatos		
Nombre del elemento	Etiqueta XML	Aplicación
Título	<title>	Nombre dado al recurso. Es el nombre por el cual el recurso es conocido. Debe plasmarse tal y como aparece en el recurso.
Creador	<creator>	Nombre de la persona o entidad corporativa responsable de la creación del contenido intelectual del recurso.
Tema	<subject>	Término, descriptor o encabezamiento que representa y permite identificar el contenido intelectual del recurso. Se recomienda utilizar lenguajes documentales para la descripción temática del recurso. Los temas pueden estar representados de la siguiente manera: nombres de personas, corporativos, de forma, de contenido, geográficos, históricos.
Descripción	<description>	Representación abreviada que refleja el contenido temático del recurso. La descripción de un recurso puede incluir los siguientes elementos: tabla de contenido, resumen, notas de contenido.
Editor	<publisher>	Nombre de la entidad responsable de publicar bajo un formato determinado el recurso de información. La imagen de editor engloba a personas, organismos o servicios.
Colaborador	<contributor>	Persona o entidad corporativa que contribuye intelectualmente en el contenido del recurso de información. Los colaboradores pueden ser personas, organismos y cumplir determinadas funciones: productores, editores, ilustradores, traductores, músicos, etc.
Fecha	<date>	Fecha asociada con el ciclo de vida del recurso de información. La fecha se relaciona directamente con la creación o disponibilidad del recurso de información. En este elemento pueden utilizarse fechas relacionadas con los siguientes datos: vigencia, disponibilidad, expedición o emisión, modificación, registro de derechos, dictaminación.
Tipo	<type>	Se refiere a la naturaleza o género del contenido del objeto. Por ejemplo: novela, ensayo, reposte técnico, artículo de investigación.
Formato	<format>	Se refiere a la manifestación física o digital del recurso de información. El formato incluye el medio o las dimensiones del recurso. Este elemento es utilizado para determinar el software, hardware o aplicación necesaria para consultar y visualizar el recurso.
Identificador	<identifier>	Cadena de caracteres o signos que sirven para identificar el recurso de una manera unívoca. Los ejemplos más comunes de identificadores son los siguientes: ISBN, DOI, URL.
Fuente	<source>	Referencia al recurso o fuente de donde se deriva el recurso actual. Por ejemplo, el capítulo de un libro digital.
Idioma	<language>	Idioma en el que se expresa el contenido intelectual del recurso.
Relación	<relation>	Referencia a otro recurso o material relacionado con el mismo. Por ejemplo, el álbum musical que incluye la pista descrita. O bien, la ponencia de una memoria digital descrita. Este es un elemento de suma importancia para establecer vinculaciones entre recursos de información documental. Y es un elemento altamente relevante para la vinculación semántica de los recursos.
Cobertura	<coverage>	Alcances espaciales, temporales o jurisdiccionales del contenido intelectual del recurso.
Derechos	<rights>	Mención acerca de los derechos del recurso de información. Este elemento contiene un identificador que vincula al recurso con una declaración acerca de sus derechos patrimoniales.

El elemento de relación es altamente relevante para la construcción de vínculos significativos entre recursos de información documental. Estos vínculos pueden ser de tipo temáticos, referentes al creador del recurso, relativos al título de la obra y al colaborador del desarrollo del contenido intelectual del recurso.

En la actualidad, los elementos de metadatos de Dublin Core son capaces de integrarse con la norma RDF que hace posible el establecimiento de vinculaciones

significativas entre datos. En este sentido, la construcción de perfiles semánticos de metadatos requiere de la integración de los elementos Dublin Core y los aspectos normativos de RDF. Además, esta integración permite flexibilizar los metadatos utilizados actualmente. De acuerdo con Foulonneau y Riley (2008: 187), “... *la flexibilidad de los metadatos está ganando importancia. Las funciones previstas de las aplicaciones de la información en las instituciones del patrimonio cultural*

*cambian con el tiempo. Las fuentes potenciales de metadatos se están diversificando.”*

El estudio de los metadatos semánticos para el desarrollo de procesos de búsqueda y recuperación de información ya ha sido abordado con anterioridad, García y Celma (2005: 1) han identificado la comple-

jidad de establecer entornos semánticos de metadatos, al afirmar que, una vez en un espacio semántico, la integración de datos es un factor crucial cuando se dispone de varias fuentes de información, pues se facilita enormemente su interrelación.

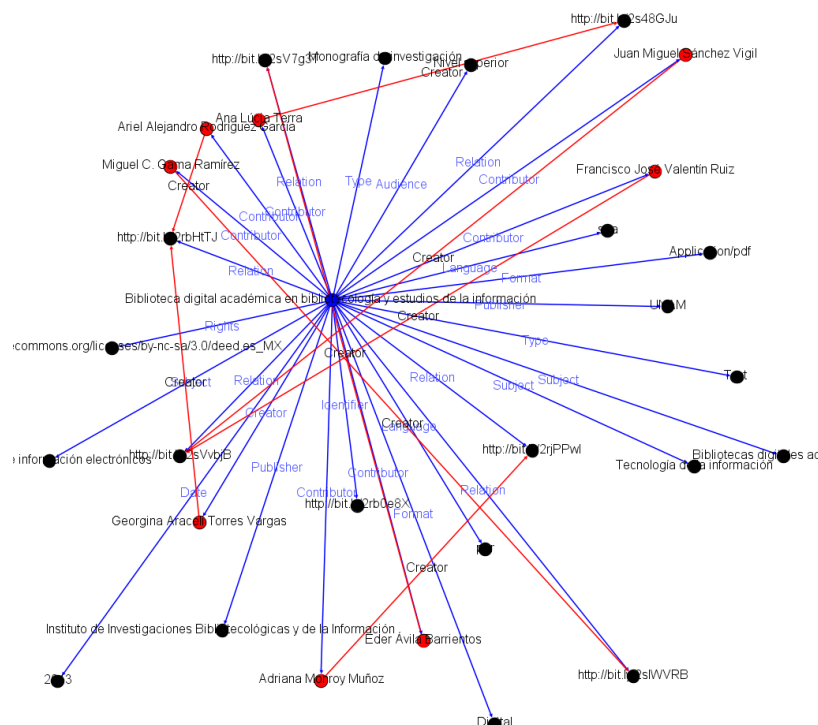


Figura 3. Codificación XML basada en Dublin Core y RDF que ejemplifica la construcción de un perfil semántico de metadatos para monografías digitales. Elaboración propia, 2020.

Como puede apreciarse en la ilustración no. 3, la construcción de un perfil semántico de metadatos consiste en la aplicación e integración de los elementos de la norma RDF y el esquema de metadatos Dublin Core. Este perfil de metadatos bien puede colocarse en la interfaz de descripción de recursos de información documental disponible en el software que se utiliza para la construcción de un repositorio institucional.

En la ilustración no. 4, puede apreciarse una representación gráfica que ejemplifica la función de un perfil de metadatos semántico correspondiente a una monografía digital. Como puede observarse, cada elemento descriptivo de Dublin Core (color azul) permite describir, representar y vincular los atributos que conforman a un recurso de información documental (nodos negros).

A su vez, pueden descubrirse las relaciones intelectuales que se presentan en el contenido del recurso (nodos rojos). Lo importante de esta función, recae en la posibilidad de identificar la relación significativa entre los datos bibliográficos y de autoridades que caracterizan al recurso de información digital.

## 6. Discusión

Un perfil semántico de metadatos es un mecanismo normativo que permite representar, describir y relacionar

los atributos de un recurso de información digital a través de elementos previamente definidos. La aplicación de estos perfiles en el ambiente de los repositorios institucionales requiere de un dominio normativo, tecnológico y metódico.

La formulación de un análisis descriptivo de los elementos que permiten la vinculación de contenidos en repositorios institucionales, mediante perfiles semánticos de metadatos pone de manifiesto la utilización de normas de índole documental e informática como es el caso de Dublin Core y RDF.

Sin embargo, la aplicación de los perfiles semánticos de metadatos en repositorios institucionales debe contemplar todas las variables que rodean a este contexto, por ejemplo, tipo de usuarios, tipos de colecciones que almacena el repositorio, misión, visión y propósitos de este.

Es decir, contar con un marco de políticas para la implementación de perfiles semánticos de metadatos por tipo de colección almacenada en el repositorio. De no contemplar estos elementos, solo se aplicarían principios normativos sin un fundamento consistente y sin capacidad para atender la demanda informativa del usuario final.

Los esquemas de metadatos han evolucionado a la par del entorno digital y los recursos de información, por este motivo, es necesario plantear alternativas que



hagan posible el desarrollo de estructuras integrales para representar a los datos que forman parte de los recursos, además que permitan vincularlos con diferentes fuentes disponibles en la web.

De esta manera, la transición hacia una nueva generación de metadatos debe contemplar la interoperabilidad de los datos y de los diferentes sistemas en donde se representan, así como la incorporación de los principios de los datos enlazados.

Al respecto Smith-Yoshimura (2020: 1) manifiesta que *los metadatos son fundamentales para la transición a las próximas generaciones de bibliotecas y sistemas de descubrimiento. Los buenos metadatos creados hoy se pueden reutilizar fácilmente en un entorno de datos vinculados en el futuro*. No obstante, esta transición traerá consigo cambios en la manera de catalogar y describir a los recursos, por lo que el profesional de la información deberá capacitarse y adaptarse en términos de desarrollar una organización de la información de carácter integrador y universal, pues la aplicación de los datos enlazados para vincular contenidos tiene un carácter unificador que fomenta el descubrimiento de nuevos hallazgos basados en el procesamiento y organización de los datos.

## 7. Conclusiones

La vinculación de contenidos en repositorios institucionales mediante perfiles semánticos de metadatos se trata de un proceso analítico intelectual que tiene como propósito principal ofrecer un mecanismo para descubrir los atributos que se encuentran relacionados significativamente en un contexto digital.

En los repositorios institucionales, se encuentra el ambiente idóneo para la aplicación de estos perfiles, pues son sistemas utilizados por comunidades científicas y académicas para divulgar y desarrollar sus procesos de investigación.

La vinculación de contenidos en el ambiente digital consiste en un rastreo de las relaciones significativas que rodean y conforman a un determinado recurso de información documental. Este recurso es susceptible de relacionarse en una colección específica que está disponible en un repositorio institucional y, en consecuencia, vincularse con fuentes externas a éste mediante la utilización de principios normativos e informáticos similares.

Los perfiles semánticos de metadatos aspiran a convertirse en un mecanismo para vincular contenidos en el entorno de los repositorios institucionales y de esta manera ofrecer servicios de información digitales con un factor innovador.

## Referencias bibliográficas

- Berners Lee, Tim. (2006). Linked Data. <<https://www.w3.org/DesignIssues/LinkedData.html>>. (Consulta: 23/02/2019)
- Berners Lee, Tim (2009). Metadata architecture. <<https://www.w3.org/DesignIssues/Metadata.html>>. (Consulta: 23/02/2019)
- Berners Lee, Tim.; Hendler, James. y Lassila, Ora. (2001). The semantic web. *Scientific American*, no. 5, vol. 284, p. 34-43.
- Ding, Li. et al. (2007). Using ontologies in the semantic web: a survey. // Sharman, Raj; Kishore, Rajiv; Ramesh, Ram (Coordinadores). *Ontologies: A handbook of principles, concepts, and applications in information systems*. New York: Springer, p. 79-113.
- Downey, Moira. (2019). Assessing Author Identifiers: Preparing for a Linked Data approach to name authority control in an institutional repository context. // *Journal of Library Metadata*, no. 1/2, vol. 19, p. 117-136. <<https://doi.org/10.1080/19386389.2019.1590936>>(Consulta: 25/10/2019).
- Dunsire, Gordon. (2008). Collecting metadata from institutional repositories. // *OCLC Systems & Services: International digital library perspectives*, no. 1, vol. 24, p. 51-58, <<https://doi.org/10.1108/10650750810847251>>(Consulta:20/11/2019).
- Foulonneau, Muriel y Riley, Jenn. (2008). *Metadata for Digital Resources*. USA: Chandos Publishing, 2008.
- García, Roberto y Celma, Oscar. (2005). Semantic Integration and Retrieval of Multimedia Metadata. // 5th International Workshop on Knowledge Markup and Semantic Annotation, p. 69-80. <<http://ceur-ws.org/Vol-185/semAnnot05-07.pdf>> (Consulta: 20/04/2019).
- Hanneman, Robert. (2001). Introducción a los métodos del análisis de redes sociales. // *Revista redes* <<http://revista-redes.rediris.es/webredes/textos/Introduc.pdf>> [Consulta: 24/03/2020]
- Jakus, Grega. (2013). *Concepts, Ontologies, and Knowledge Representation*. New York: Springer, 2013.
- Jones, Richard, Andrew, Theo y MacColl, John. (2006). *The institutional repository*. USA: Chandos, 2006.
- Lim, Yon Soo. (2010). Semantic Web and Contextual Information: Semantic Network Analysis of Online Journalistic Texts. // Breslin, John G. et al. (editores) *Recent Trends and Developments in Social Software*. Lecture Notes in Computer Science. Springer: Berlin, 2010. p. 52-62. <[https://link.springer.com/content/pdf/10.1007%2F978-3-642-16581-8\\_6.pdf](https://link.springer.com/content/pdf/10.1007%2F978-3-642-16581-8_6.pdf)>
- Lynch, Clifford. (2003). Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. // *Portal Libraries and the academy*, no. 226, <<http://www.arl.org/resources/pubs/br/br226/br226ir.shtml>> [Consulta: 20/04/2020]
- Russo, Vanessa. (2015). Semantic Web: metadata, linked data, open data. *Science and philosophy*, no. 2, vol. 3. <<https://core.ac.uk/download/pdf/322555618.pdf>> [Consulta: 30/10/2019].
- Smith-Yoshimura, Karen. (2020). *Transitioning to the Next Generation of Metadata*. USA: OCLC Research, 2020.
- World Wide Web Consortium - W3C (2015). *Vocabularies*. <<https://www.w3.org/standards/semanticweb/ontology>> [Consulta: 25/02/2020]

