

Tendencias multidisciplinarias del uso de los metadatos

*Ariel Alejandro Rodríguez García
Raúl Ariel González Castillo*
(COORDINADORES)



La presente obra está bajo una licencia de:

http://creativecommons.org/licenses/by-nc-sa/3.0/deed.es_MX



Atribución-No Comercial-Licenciamiento Recíproco 3.0 Unported

Eres libre de:



copiar, distribuir y comunicar públicamente la obra



hacer obras derivadas

Bajo las condiciones siguientes:



Atribución — Debes reconocer la autoría de la obra en los términos especificados por el propio autor o licenciente.



No comercial — No puedes utilizar esta obra para fines comerciales.



Licenciamiento Recíproco — Si alteras, transformas o creas una obra a partir de esta obra, solo podrás distribuir la obra resultante bajo una licencia igual a ésta.

Esto es un resumen fácilmente legible del:
[texto legal \(de la licencia completa\)](#)

En los casos que sea usada la presente obra, deben respetarse los términos especificados en esta licencia.



**Tendencias multidisciplinares
del uso de los metadatos**

COLECCIÓN
METADATOS

Instituto de Investigaciones Bibliotecológicas y de la Información

**Tendencias multidisciplinares
del uso de los metadatos**

Coordinadores

**Ariel Alejandro Rodríguez García
Raúl Ariel González Castillo**



**Universidad Nacional Autónoma de México
2017**

Z666.7
T45

Tendencias multidisciplinares del uso de los metadatos / coordinadores Ariel Alejandro Rodríguez García, Raúl Ariel González Castillo. – México : UNAM, Instituto de Investigaciones Bibliotecológicas y de la Información, 2017.

xii, 191 p. – (Metadatos)
ISBN:978-607-02-9946-9

Metadatos 2. Organización de la Información 3. Servicios Web I. Rodríguez García, Ariel Alejandro, coordinador II. González Castillo, Raúl Ariel, coordinador III. ser.

Diseño de portada: Mario Ocampo Chávez
Imagen base de portada: Salvador Mendoza López

Primera edición, 2017
D.R. © UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
Ciudad Universitaria, 04510, Ciudad de México
Impreso y hecho en México
ISBN: 978-607-02-9946-9

Publicación dictaminada

Contenido

| | |
|----------------------------------|-----|
| Introducción..... | vii |
| ARIEL ALEJANDRO RODRÍGUEZ GARCÍA | |
| RAÚL ARIEL GONZÁLEZ CASTILLO | |

MODELOS, ESTRUCTURAS Y PAUTAS DE LOS METADATOS

| | |
|---|---|
| Sound Archives, Metadata and New Web Technologies. | 3 |
| JOSEPHINE SIMONNOT | |

| | |
|---|----|
| Datos enlazados para bibliotecarios. | 11 |
| JOSÉ MANUEL MORALES DEL CASTILLO | |

| | |
|--|----|
| Las ontologías en la representación y organización temática de la información bibliotecológica..... | 25 |
| ADRIANA SUÁREZ SÁNCHEZ | |

USUARIOS, CONSUMIDORES Y PROSUMIDORES DE INFORMACIÓN

| | |
|--|----|
| El uso de metadatos para la gestión del patrimonio digital mundial en la Sociedad de la Información. | 39 |
| LUIS FERNANDO RAMOS SIMÓN | |
| SILVIA COBO SERRANO | |

| | |
|--|----|
| Interoperabilidad global en las comunidades bibliotecaria, archivística y museográfica..... | 51 |
| ARIEL ALEJANDRO RODRÍGUEZ GARCÍA | |

| | |
|--|----|
| Tecnologia, Metadados e Interoperabilidade: aspectos na modelagem de ambientes informacionais digitais. | 61 |
| FABIANO FERREIRA DE CASTRO | |

GOBIERNO ABIERTO Y CIUDADANIA

| | |
|---|----|
| La participación ciudadana en el contexto del gobierno abierto. | 85 |
| ALEJANDRO RAMOS CHÁVEZ | |
| Centro de Información Político-Electoral. | 97 |
| PABLO PRIETO MONTES DE OCA MOISÉS RUBÉN RAMÍREZ ZÁRATE | |

CULTURA DIGITAL Y MEDIOS SOCIALES

| | |
|---|-----|
| Consumidores de información y descripción de documentos: las etiquetas de los usuarios en los catálogos de bibliotecas..... | 117 |
| ANDRÉS FERNÁNDEZ RAMOS | |
| Camino para establecer la relación entre usuarios y metadatos. | 129 |
| PATRICIA HERNÁNDEZ SALAZAR | |

EDUCACIÓN A DISTANCIA Y METADATOS

| | |
|--|-----|
| Tendencias en la literatura sobre metadatos de contenidos educativos: atención al formato, la web semántica y el aprendizaje. | 143 |
| ALMA BEATRIZ RIVERA AGUILERA | |
| La importancia de los metadatos para el aprovechamiento de los recursos educativos en la educación a distancia..... | 161 |
| JORGE LEÓN MARTÍNEZ EDITH TAPIA RANGEL | |
| Competencia en metadatos como parte del proceso educativo en educación a distancia. | 177 |
| MARCO BRANDÃO | |

Introducción

La perspectiva que se tiene sobre el uso de los metadatos es multidimensional. Hay un universo de esquemas y normas que responde a formas particulares de tratar los datos de los recursos de información digital para que sean visibles en las redes y los sistemas de información. De ahí que cada disciplina procure conservar sus métodos y prácticas pero a su vez acepte que sus datos puedan analizarse desde diferentes campos.

Detrás del concepto “uso de los metadatos”, existen definiciones muy diversas, incluso contrapuestas, que permiten superar la tentación de una ciencia o disciplina reina o arquetípica —la Bibliotecología y los Estudios de la Información en nuestro caso— al ser abordadas desde una visión multidisciplinaria, y pasar a un modelo donde las distintas disciplinas planifican sus recursos, necesidades y medios para ser articuladas interdisciplinariamente.

El uso de los metadatos se vuelve difícil de comprender en cada campo disciplinar principalmente por dos razones. La primera de ellas es que la práctica de la catalogación desarrollada en la Bibliotecología ha creado metadatos desde hace muchos años, lo que ha provocado el surgimiento de varias estructuras para la descripción y representación de la información. La segunda es que los

metadatos emiten teorías al interior de cada comunidad, las cuales cumplen con sus propósitos de construcción.

La multidiscipliplina da razón al saber sobre lo concreto de un problema, como el que nos reúne en esta ocasión, los metadatos. Desde que aparecieron éstos en la Bibliotecología, se consideraron como referentes para la descripción de los recursos de información, así como los servicios que potencialmente se ofrecen a los usuarios. De ahí que su fin principal sea facilitar la identificación, localización, recuperación, manipulación y uso de los datos contenidos en los recursos de información.

Las características que tienen los metadatos son muy variadas, van desde su concepción como estándares integradores de información, hasta la configuración de estructuras interoperables de datos con identificadores propios, éstos últimos definidos por la adopción del lenguaje de marcado. Es decir, quien haya aprendido y aplicado alguna estructura de metadatos, reconocerá que por natural que parezca su aplicación, éstos nos dirigen a un mismo propósito: sistematizar los datos de los recursos de información.

El enfoque simplista que diversas disciplinas tienen respecto a los metadatos propicia que se vea solamente su alcance utilitario más que su fundamentación teórico-conceptual, por esto es que la investigación sobre los metadatos, aparentemente, transita de manera pausada. Sin embargo, la realidad advierte que el uso de los metadatos está siendo más favorecida porque se requiere de contar con descripciones más cercanas a los usuarios como lo señala el creador de la web semántica, Tim Berners Lee, quien refiere que los metadatos son piezas fundamentales para la organización-recuperación de información y sus ámbitos de acciones son tan numerosos que incluyen, entre otros, las redes sociales, las colecciones digitales y las tareas afines de colaboración e interoperabilidad de las comunidades, por ejemplo, la bibliotecaria, archivística y museográfica.

Esta obra se forma a partir de las reflexiones recogidas del Seminario de Investigación sobre Metadatos y el esfuerzo colectivo de distintos investigadores y especialistas en Bibliotecología, documentación, educación, ciencias políticas, periodismo, entre otras áreas, cuyo marco de discusión fue analizar los aspectos conceptuales,

teóricos y metodológicos sobre el uso de los metadatos desde una perspectiva multidisciplinaria.

La composición temática seleccionada para esta obra probablemente sea limitada para varios lectores cuando se tiene en mente un interés en particular o alguna necesidad específica de información. Pero la intensión que se busca es que el lector no pierda de vista que el asunto de los metadatos es más que el referente primario del que se tiene respecto a la tesis de “datos acerca de datos”.

De ahí que la obra se constituya en cinco partes. La parte 1 inicia con documentos relacionados con el modelado, las estructuras y pautas sobre el uso de los metadatos y la forma en que diversos especialistas y equipos multidisciplinarios interpretan las directrices para crear o actualizar un esquema de metadatos.

La parte 2 continúa con dos trabajos que se enfocan al estudio del usuario, consumidor y prosumidor de información desde la perspectiva de la gestión del patrimonio mundial digital y los alcances que tiene la interoperabilidad global en las comunidades bibliotecaria, museográfica y archivística, y deja de manifiesto que debe defenderse el libre uso de los metadatos y estándares comunes con el fin de impulsar el patrimonio cultural al alcance de todos.

Las investigaciones relacionadas con el gobierno abierto y ciudadanía que constituyen la parte 3, donde se trazan los argumentos referentes a las diversas iniciativas que el gobierno abierto ha puesto a disposición de los ciudadanos con relación a los procesos de rendición de cuentas y los canales para hacer un gobierno más participativo que responde a las demandas del ciudadano. Además, se deja en claro que las formas sistemáticas y automáticas para la extracción de datos son las mejores prácticas en la descripción de los documentos oficiales.

La cultura digital y los medios sociales, temas de la parte 4, invita a que el profesional dedicado al trabajo con metadatos reflexione sobre los nuevos retos y cambios en su perfil, ya que el etiquetado de la información está siendo la práctica común para establecer la relación entre el usuario, la información y los metadatos.

Tendencias multidisciplinarias...

Las tres investigaciones que constituyen la parte 5 de la obra cubren específicamente el asunto de la educación a distancia con el firme propósito de dar a conocer las modalidades actuales y futuras del proceso de enseñanza-aprendizaje, el cual no sólo permite el acceso a la información, sino que inclusive genera conocimiento nuevo.

Las contribuciones reunidas en esta edición, de diferente manera, favorecen a la composición temática y líneas de acción para que los integrantes del Seminario de Investigación sobre Metadatos que se desarrolla en el Instituto de Investigaciones Bibliotecológicas y de la Información de la UNAM y aquellos interesados en el tema, continúen investigando respecto al objetivo, proceso, prácticas, productos y servicios relacionados con la tesis de “datos acerca de datos”.

Queremos agradecer a todos los que de manera directa e indirecta permitieron la culminación de este proyecto, especialmente a la doctora Georgina Araceli Torres Vargas, directora del IIBI, al personal del Instituto y a los miembros del Seminario de Investigación sobre Metadatos por impulsar y ver culminado el trabajo de muchas semanas. A todos, gracias por su colaboración y participación.

Ariel Alejandro Rodríguez García

Raúl Ariel González Castillo

**MODELOS, ESTRUCTURAS
Y PAUTAS DE LOS METADATOS**

Sound Archives, Metadata and New Web Technologies

JOSEPHINE SIMONNOT

*Centre de Recherche en Ethnomusicologie (CREM)
Université Paris 10, France*

INTRODUCTION

In this paper, I present my experience to preserve, manage and disseminate a cultural heritage: Les Archives Sonores du CNRS-Musée de l'Homme, France, because, in order to keep alive this common world heritage, it must not disappear, and it must be accessible by the specialists and also shared with a large community of users to be enriched for the future generations.

Those archives are supported by the French National Research Center (CNRS), the National Museum of Natural History, the Paris 10 University and the Culture Ministry. Its history follows many steps: it began with the Trocadéro Ethnographic Museum and goes on with the Musée de l'Homme in 1937, under the direction of André Schaeffner. The Ethnomusicology Department, with his successor Gilbert Rouget, established a research team in 1968, in order to improve the research on music and sound. The aim was

also to go on the dissemination with the discographic publication.¹ The other main objective was to preserve this traditional heritage for the next generations. Since the beginning, Gilbert Rouget was convinced that it's important to disseminate this musical diversity and share the knowledge of this world. Publishing books was not enough to understand the musical phenomenon.

To illustrate the history and diversity of the collections deposited during one century, we listened the oldest recording of our sound archives from Mexico wa cylinder around 1932, a woman song: http://archives.crem-cnrs.fr/archives/items/CNRSMH_I_1939_001_002_01/. And a young Mexican Lacandon boy song (“levent”) recorded on the first magnetic tape of our Mexican colletions before 1953: http://archives.crem-cnrs.fr/archives/items/CNRSMH_I_1953_005_001_03/.

The collections content about 5000 published records of traditional music and fields recordings (grabaciones de campo). Today, the database contents more than 60 000 items to inform the various recordings (42 000 from fields recordings).

For many years, we started a long process of digitization: first on CD-R, later on hard disc and now on the server of the National Infrastructure Huma-num, the service for humanities research and cultural data in France. <http://www.huma-num.fr>.

The Bibliothèque Nationale of France (BnF) continues the digitisation and we hope to finish it in 4 or 5 years. It consumes a lot of time because it is necessary to segment the sound files to have an easy access to the content. The treatment of the metadata is also a long process. For example, we can see on this paper document the description done in 1975 in the old catalogue with ethnographic and musical information:

1 All the phonographic publications are available online on the CREM database (http://archives.crem-cnrs.fr/archives/fonds/CNRSMH_Editions/).

Figure 1. Aurore Monod Becquelin archives item, Bachajon, Mexico, 1973. Source: CREM-LESC, CNRS.²

| | | | | | | |
|--|-----|---------------------------------------|------------------------|--|--------|---|
| Becquelin Monod (Aurore) 15 Bachajón 1 Mars 1973 | | COPIE selectine Becquelin Monod A. | Über cassette M. H. | L. D. 1/2 P | Ø 13 | B N (75.15.1.) 1 à XI |
| I | 19 | 1 | 5' | MEXIQUE, État de Chiapas, village de Bachajón. <i>Fr. de Base</i> <i>à l'heure 1/2 p. original</i> <i>7h10 2/10</i> | | Tzeltal - groupe maya- Carneval de Bachajón : musique des Kabinai, hommes de la forêt. Parlé, conversation, puis deux flûtes <i>amay</i> , deux tambours <i>kayob</i> à deux peaux tenus horizontalement et battus avec une baguette, deux conques <i>puj</i> , scalles de tortues <i>hax</i> battues avec deux épis de maïs, hochements et discours rituel <i>pa'to'tan</i> psalmodié par deux He. |
| N.B. cf. Becquelin Monod Aurore "Le Carnaval de Bachajón" J.S.A. 1975 (à paraître). Cf. documentation photo, Mission archéologique française au Mexique : réf. 73 E. | | | | | | |
| PIECE | Vn. | Pa. | Durée | ÉTAT | ETHNIE | CONTENU |

Sometimes, we don't have any information, only some notes on the original box tapes. But the priority is to share the audio document to facilitate the identification of the content.

So, how to give access to this documents, both audio and text? How to play audio on line and navigate inside the sound file? How to preserve data for long term? How to manage the database daily changing and to manage the access according to the intellectual property?

For many years, one of CNRS's objectives has been to improve access and to facilitate data sharing to the entire academic community. We do this through a toolbox of open source software's (HumaNum service). In anthropology, ethnomusicology and linguistics, researchers work on multiple kinds of documents such as sound recordings. The need to preserve and to easily access, visualize and annotate these materials is problematic because of their diverse formats, contents and the increasing quantity of data. With new audio technologies, questions linked to the preservation, the archiving and the availability of these audio materials have arisen. Since 2007, French ethnomusicologists and engineers have joined

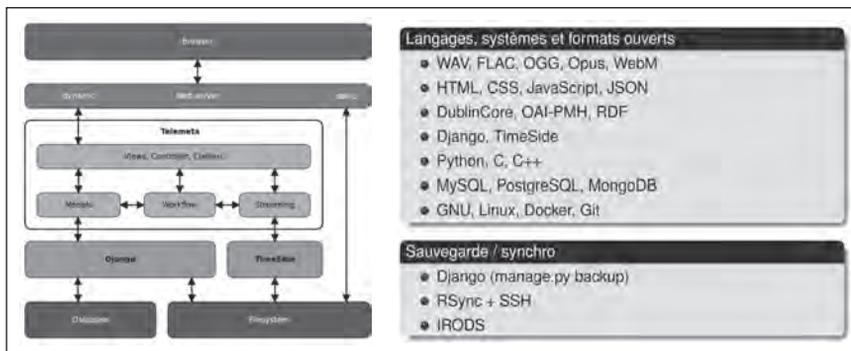
2 Aurore Monod is research director at the CNRS Maya Civilization. More information on her work on: http://archives.crem-cnrs.fr/archives/fonds/CNRS_MH_Monod-Becquelin/.

Tendencias multidisciplinarias...

their efforts to develop a collaborative web platform for managing and improve access to digitized sound archives. This platform, an open-source web audio CMS dedicated to digital sound archives (Telemeta³), is developed through the expertise of the Parisson/Start up. Its architecture is associated with Time Side,⁴ an open-source audio processing framework written in Python and JavaScript languages, which provides decoding, encoding and streaming capabilities together with an embeddable HTML audio player. Consequently, this CMS is able to produce visualization, annotation, segmentation, etcetera.

This tool allows to work on archives and to disseminate the sound collaboratively. A prototype started with the metadata integration in a MySQL database. It started with a prototype on line and the integration of a MySQL database to manage only metadata. It was yet a great improvement because we could browse the catalogue and create new metadata on line! The Telemeta platform launches in 2011 with 7000 sound files.

Figure 2. Telemeta architecture, formats. Source: Parisson.



Telemeta is a web content system to manage a database and the audio files (streaming player, graphic view, compression).

3 <http://telemeta.org>.

4 <https://github.com/Parisson/TimeSide>.

- Manage metadata and the audio/video files.
- Collaborative editing.
- Standardization: data and metadata. For example: the ISO standard for the languages vocabulary, Thesaurus of Geographic Names.
- Audio analysis, visualisation, encoding, embedding (i-frame link).
- Open source software and open formats.
- Access management with User count system.
- Catalogues accessible in open access.
- Right policy: sounds older than 50 years are public domain, they have free access but no re-use (the collector permission is required).
- Last change for the follow the evolution and the contributors work.
- Textual research and advanced research.

Each database has specific needs. An adapted research engine is required to have an easy access to the data: location, population, musical instrument, physical support, etc.

The first step was to describe the content with a structured model and create a specific form to organize the metadata. The collection form contents many fields: numbers, text, tables for ontologies. The next step in 2016-2017 will be to include authorities for personal names. For the interoperability, we must do the relation of the main metadata with the Dublin Core model so the database can be harvested by other web portals.

The catalogue is organized with four levels and you can group the collections as you wish. Items are grouped in collections, collections are grouped in corpus, and corpus are grouped in Funds. These Funds level describe mainly the works of a collector and propose a list of the different corpus of this person.

In the corpus level, you can see the collections, digitized or not digitized. The archivist can put markers on the sound files to show where the sound engineer should do the segmentation. The specificity of this tool is the possibility to add markers on the timeline

Tendencias multidisciplinarias...

as well as notes on the sound file. The annotation is possible if you have a user account and they are available to everybody with the author of the notes. The spectral view is processed by the Time Side engine and offers another visualization of the content. You can copy the i-frame link to imbed the player in your website. It is not necessary to download the sound file and in fact, it is not allowed without permission. Interoperability is possible thanks to the Dublin Core and OAI-PMH protocol. For example: <http://www.rechercheisidore.fr>. A best practice guide is on line: <http://www.hu-ma-um.fr/ressources/guides>

Since the platform is launched, the database increase regularly and we have today more than 40 000 sound files on line, which 24 000 are on free access. And 2000 more visitors than before the platform!

The community of Telemeta users count 3 other teams: <http://telemeta.lam.jussieu.fr>, <http://sabiody.telemeta.org/> and <http://pho-notheque.cmam.tn/>.

This project brings on other programs, international partnership and funding: Europeana Sounds for example, or analysis research program (Music Information Retrieval). So maybe we will all be connected together soon, because it will possible to connect Telemeta platforms with each other! This a rich experience to build common vocabulary, data models, and a task force about Intellectual property and ethical rules for the intangible world heritage.

The Wiki Telemeta project is available at: <http://telemeta.org> and <https://github.com/Parisson/Telemeta>. The benefits of this collaborative platform apply to numerous aspects of the field of ethnomusicology, ranging from musical analysis to comparative history and anthropology of music, as well as to the fields of anthropology, linguistics and acoustics. You can join us and download the tool, contribute on the database and you are welcome to visit the Sound Archives of Musée de l'Homme.⁵

5 <http://archives.crem-cnrs.fr>.

BIBLIOGRAPHY

- Archivos de Aurore Monod Becquelin (1973). Bachajón.
Disponible en: http://www.mae.u-paris10.fr/lesc/IMG/pdf/bibweb_monod.pdf.
- Monod Becquelin, A. y Alain Breton (2002). *La guerre rouge, ou: une politique maya du sacré*. París: CNRS éditions.
- Monod Becquelin, A. (2002). Parlons tzeltal. *Une langue maya du Chiapas*. París: L'Harmattan.

Datos enlazados para bibliotecarios

JOSÉ MANUEL MORALES DEL CASTILLO

Biblioteca Daniel Cosío Villegas

El Colegio de México

En 2001 Berners-Lee, Hendler y Lassila (Berners-Lee, Hendler y Lassila, 2001) publicaron un trabajo en el que nos describen una web en la que al momento de realizar una consulta, dispondríamos de uno o varios agentes software inteligentes que se interconectarían con otros agentes para agregar información obtenida de diferentes fuentes dispersas y proporcionar una respuesta relevante, precisa y confiable (lo que haría que la tarea de ojear listados con docenas de resultados irrelevantes pasara a la historia). En esta evolución de la web, que ellos denominan Web semántica, la información se estructura en un intrincado entramado de conceptos interrelacionados entre sí que son interoperables, reutilizables, comprensibles para los humanos y procesables por máquina.

Son muchos los que se preguntan en qué ha quedado el proyecto, de qué se trata realmente y dónde quedaron todas esas buenas intenciones descritas en el trabajo de Berners-Lee. Pues, aunque no lo parezca, la idea de desarrollar este modelo de web ha ido permeando poco a poco y, si bien es verdad que estamos lejos de poder decir que la web actual es semántica, sí que se han estado dando pasos firmes hacia su progresiva “semantización”.

Tendencias multidisciplinarias...

En esta web, el elemento esencial ya no es el documento, sino el dato y la red de relaciones semánticas que se establece con otros datos, todos ellos interoperables y reutilizables. Por lo tanto, el proceso de semantización es inherente a esta web de datos. Los profesionales de la información también juegan un papel destacado en este nuevo escenario donde los recursos web tienen que ser “catalogados” para poder ser semánticamente accesibles. La consolidación de este modelo podría servir para reivindicar el papel profesional de los bibliotecarios y su reinserción a la web como actores destacados (García Marco, 2013).

No siempre tenemos claro a qué nos estamos enfrentando o qué se espera realmente de nosotros. La idea de este trabajo es tratar de definir de una manera accesible qué es la web semántica o web de datos, qué elementos esenciales la conforman, en qué momento de desarrollo se encuentra el proyecto y cómo encajan en él las bibliotecas y los bibliotecarios. Comencemos.

EL ORIGEN

Lo que movió a Berners-Lee, Hendler y Lassila a idear el modelo de web semántica fue la necesidad de resolver el problema de la sobrecarga de información. Todos hemos utilizado un motor de búsqueda (como Google) y hemos sufrido la situación de recibir como respuesta un listado interminable de resultados, de los cuales un gran porcentaje poco o nada tienen que ver con lo que estamos buscando.

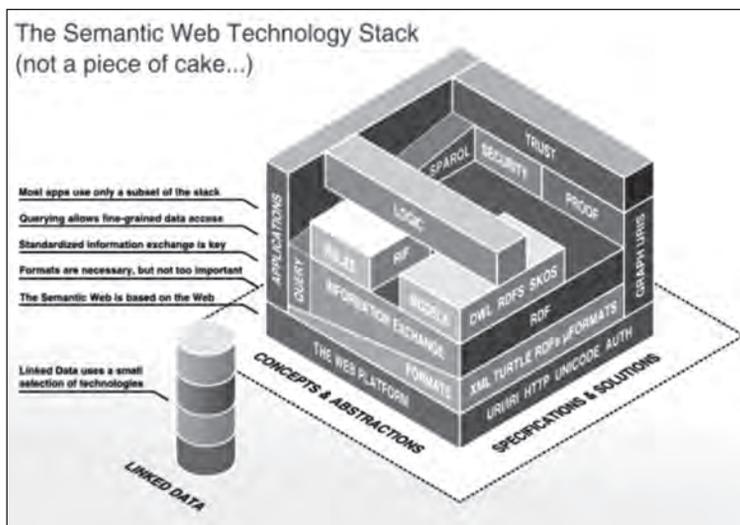
Esto ocurre porque HTML, el lenguaje de etiquetado en el que están descritos la mayoría de los documentos web, no está diseñado para caracterizar contenidos, sino que sirve principalmente para estructurar cómo se va a presentar la información en pantalla. Los motores de búsqueda explotan esta circunstancia realizando correspondencias entre las palabras clave definidas por el usuario en su consulta y cadenas de caracteres definidas dentro de etiquetas con un valor semántico potencial (como las etiquetas *title* o *meta*). Esto explica por qué obtenemos una mezcla de resultados

relevantes (no podemos obviar esa circunstancia) con otros completamente descontextualizados. La solución, pues, implicaba conseguir que la web pudiera determinar exactamente qué es lo que el usuario está buscando y ofrecerle una respuesta precisa y fiable al tomar en cuenta el contexto en el que la consulta se ha realizado. De esta manera, se hará posible una recuperación de información que se genera de manera dinámica a partir de fuentes dispersas (aunque conectadas semánticamente) y que agrega diferentes tipos de datos y formatos (texto, imagen, mapas, gráficas, sonido, video) a través de un interfaz único.

EL MODELO

La construcción de la web semántica se estructura en diferentes capas de desarrollo superpuestas unas sobre otras e interoperables entre sí (es decir, los elementos desarrollados en las capas inferiores pueden ser utilizados en las superiores y viceversa).

Figura 1. Modelo de capas tridimensional (Nowak, 2009)



En la base se ubica la denominada capa sintáctica y es donde se definen elementos como las URI (identificadores uniformes de recursos que permiten referir de manera unívoca a todos los elementos que intervienen en la web semántica como conceptos, personas o cosas), el protocolo de transferencia de información *http*, el sistema de codificación universal UNICODE, y las convenciones sintácticas definidas en metalenguajes como XML (Extensible Markup Language) (W3C, 2003) para garantizar la interoperabilidad, la reutilización y el intercambio de información.

Sobre ellos, en la capa semántica, se establece el modelo de datos que permite describir la información de manera que pueda ser interpretada semánticamente por máquinas. Este modelo de datos, basado en la lógica descriptiva (o lógica basada en la semántica), viene definido en el estándar RDF (Resource Description Framework) (W3C, 2014a), donde se especifica que la información se tiene que definir en forma de tripletas o sentencias Sujeto/Propiedad/Objeto para poder ser comprensible para las máquinas.

Figura 2. Ejemplo de tripleta RDF



Las tripletas RDF se pueden representar mediante grafos orientados y se interpretan de la siguiente manera: las propiedades que caracterizan a los sujetos (entendidos como cualquier tipo de entidad, evento o cosa) y les permiten relacionarse con otras entidades, eventos o cosas, o bien presentan un valor específico (como una cadena de caracteres). De acuerdo con esta convención, vemos que hay dos clases de tripletas: las que tienen como valor una entidad (“la casa está pintada de azul”) y las que tienen como valor una cadena de caracteres (“el coche cuesta \$10 000”).

No obstante, estructurar la información en forma de tripletas no basta para definir una semántica. Es necesario definir claramente el contexto en el que la información deber ser interpretada y para ello hay que recurrir a herramientas más complejas, como

las ontologías, que se definen como esquemas de conocimiento, como lo puedan ser los tesauros, pero que no se limitan a definir una estructura jerárquica. Van más allá, ya que permiten categorizar los conceptos o entidades, definir el tipo de atributos que los caracterizan y establecer cómo se relacionan unos con otros en un determinado contexto para que de esta manera los agentes software sean capaces de obtener inferencias, o lo que es lo mismo, nuevo conocimiento no explícito, a partir de una serie de asertos y un conjunto de reglas (que se definen en la capa de reglas con el estándar RIF (Rule Interchange Format) (W3C, 2013).

A su vez son varios los estándares que sirven para generar esquemas de conocimiento. Dependiendo del que elijamos, podemos definir diferentes niveles de granularidad o profundidad de acuerdo con nuestras necesidades. Por ejemplo, si queremos desarrollar ontologías con el máximo de expresividad tenemos el estándar OWL (Ontology Web Language) (W3C, 2012a), aunque esto implica un mayor costo computacional que merma la eficiencia operativa (dicho de otro modo, a mayor complejidad, mayor dificultad para obtener inferencias de una manera eficiente). Como alternativas, tenemos RDFS (Resource Description Framework Schema) (W3C, 2014b) y SKOS (Simple Knowledge Organization System) (W3C, 2012b) que ofrecen, respectivamente, soluciones para el desarrollo de ontologías ligeras y tesauros (menos potentes que una ontología definida en OWL, pero con los que es más fácil operar). Por lo tanto, la elección de unos u otros estándares depende exclusivamente de la naturaleza del problema que queramos resolver.

Por último, las capas superiores del modelo (la capa lógica, de seguridad de prueba y de confianza) son las que se encargarían de garantizar la fiabilidad y seguridad de la información que circula en la web, procurando que los datos proporcionados por los agentes fueran no sólo precisos, sino también verídicos. Al día de hoy no se han desarrollado estas capas ya que algunos de los niveles inferiores aún se encuentran en fase de consolidación.

LA APLICACIÓN PRÁCTICA DEL MODELO

Como acabamos de ver, el modelo teórico desplegado para caracterizar la web semántica está bien definido, cuenta con una multitud de estándares y vocabularios, pero ¿hasta qué punto hay un desarrollo práctico de éste? Y, si nos centramos en el ámbito bibliotecario, ¿de qué manera puede afectar al desarrollo de la profesión? Tratemos de responder a estas preguntas.

La naturaleza y complejidad del modelo de web semántica le han dado un ritmo de crecimiento y una cadencia propias, alejada de otros problemas que han encontrado una solución tecnológica casi inmediata. En este caso, se requiere de la intervención de múltiples factores y de una conjunción de sinergias que no podrían realizarse en los cortos plazos que los entornos tecnológicos imponen. No obstante, —buenas noticias—, sí que se están dando pasos hacia la web semántica; quizá no en la forma original en que se concibió el proyecto, pero sí de una manera firme hacia algo tangible. Con respecto a los esquemas de conocimiento y la necesidad de decidir entre sacrificar expresividad o facilidad de procesamiento, los desarrolladores han optado por el pragmatismo y la implementación de una web semántica *lightweight* o *ligera* que permite aprovechar algunas de sus principales funcionalidades pero a bajo costo computacional. En otras palabras, nos movemos hacia la versión beta de la Web de datos.

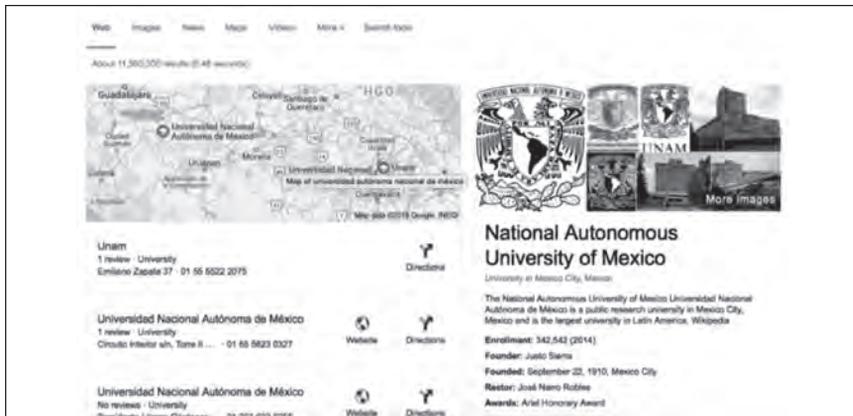
Así, frente al desarrollo de recursos descritos de manera nativa usando RDF (lo cual requiere un esfuerzo considerable), se ha optado por embeber metadatos RDF en el código fuente de otros vocabularios anfitriones. De esa manera poco invasiva, se aprovechan tecnologías que ya están plenamente operativas para enriquecer la descripción de los recursos y, consecuentemente, abrir la puerta a la obtención de resultados más precisos.

Éste es el caso de RDFa (Resource Description Framework in attributes) (W3C, 2015), una sintaxis que permite especificar metadatos de acuerdo al modelo RDF (es decir, atributos con el formato de tripletas sujeto-propiedad-objeto), pero embebidos en el código HTML5 (dentro de sus propias etiquetas). Basados en esta sintaxis,

se han desarrollado varios vocabularios específicos que pretenden facilitar el proceso de semantizar de origen los recursos de la web (democratizando el proceso de catalogación).

El ejemplo más claro es el vocabulario de microdatos Schema.org, desarrollado en 2012 por Google, Yahoo! y Microsoft, cuyo objetivo es que sus buscadores puedan proporcionar respuestas enriquecidas formadas a partir de la agregación de datos de diferentes fuentes (como Wikipedia y otras fuentes de datos estructuradas). Por ejemplo, si buscamos una universidad en Google, además del conocido listado de recursos, el buscador genera en el momento un agregado de datos que incluye mapas, logos o enlaces a biografías de egresados de esa universidad, lo que enriquece la experiencia del usuario con información adicional potencialmente relevante.

Figura 3. Ejemplo de resultados enriquecidos con datos enlazados.



DATOS ENLAZADOS PARA BIBLIOTECARIOS

El ejemplo anterior es un ejemplo claro del uso de datos enlazados en la web. La cuestión que puede surgir ahora es cómo podemos aprovechar esto en las bibliotecas. En principio, para poder trabajar con datos enlazados es necesario disponer de un conjunto de

datos, y eso es precisamente algo en lo que las bibliotecas son especialmente ricas. Sin ir más lejos, tenemos el catálogo bibliográfico como una invaluable fuente de datos semiestructurada que es posible compartir sin incurrir en ninguna violación de derechos de autor (ya que estaríamos compartiendo exclusivamente metadatos, no el contenido de las obras).

Para publicar conjuntos de datos enlazados (y enlazables), debemos tener claras algunas cuestiones esenciales. En primer lugar, debemos de aplicar el test de las 5 estrellas definido por Tim Berners-Lee para datos enlazados (Berners-Lee, 2010), una serie de criterios que nos dan una idea de la calidad de los datos que vamos a publicar:

1. Los datos deben ser legibles por máquina.
2. Estar definidos en formatos no propietarios.
3. Expresados de acuerdo al modelo de datos de RDF.
4. Deben definir vínculos entre sí y con otros conjuntos de datos.
5. Deben estar publicados bajo licencias abiertas, como dominio público o bajo una licencia Creative Commons Zero.

De acuerdo al número de criterios que cumplamos, será el número de estrellas que calificará nuestro conjunto (5 es la situación ideal).

El siguiente paso consistiría en enlazar nuestros datos tanto entre sí como con otros conjuntos de datos. Para ello, y dependiendo del tamaño del conjunto, podemos hacerlo de manera manual (cuando trabajamos con conjuntos pequeños) o de manera semiautomatizada (para conjuntos mayores) utilizando un software especializado como SILK (University of Mannheim, 2016)).

Por último, lo que hay que hacer es publicar el conjunto de datos para que estén disponibles para todos aquellos que quieran reutilizarlo. Se puede hacer, por ejemplo, mediante un simple archivo estático RDF/XML que contenga todos los elementos del conjunto; embebiendo el conjunto en un archivo HTML mediante RDFa, o utilizando bases de datos específicas que permiten el almacenamiento de tripletas RDF (Heath & Bizer, 2011).

Es importante resaltar en este punto que la publicación de un conjunto de datos a su vez implica la adquisición de un compromiso claro por parte de la institución para garantizar su veracidad, mantenimiento y actualización continua. De otro modo, no tendría sentido realizar este esfuerzo.

En el ámbito de las bibliotecas, los archivos y los museos, existe un activo movimiento de profesionales que promueven la publicación de este tipo de conjuntos de datos abiertos y enlazables (LODLAM, 2010), donde realmente la única exigencia que se hace a las bibliotecas interesadas en participar es que contribuyan poniendo a disposición del público su conjunto de datos de una manera legible por máquina. De este modo, se estará posibilitando que cualquier persona pueda trabajar con estos datos en bruto y estructurarlos de manera adecuada para su explotación y aprovechamiento por parte de toda la comunidad.

LA WEB DE DATOS Y LAS BIBLIOTECAS

La aparición de los datos abiertos enlazados tiene más implicaciones de las que en principio podríamos imaginar dentro del ámbito de las bibliotecas. De entrada, la necesidad de estructurar los datos usando la estructura que define RD hace necesario que se superen modelos tradicionales que se han quedado obsoletos ante este nuevo panorama, como es el caso de las reglas de catalogación angloamericanas AACR2 y el formato MARC. En el caso de las reglas de catalogación se utiliza RDA (Resource Description and Access) (Picco & Ortiz Repiso, 2012), un estándar que proporciona un nuevo marco para la descripción de recursos de información de acuerdo a los modelos conceptuales definidos en FRBR (Functional Requirements for Bibliographic Records) (Tillett, 2005) y FRAD (Functional Requirements for Authority Data) (Sardo, 2004) para la descripción de registros bibliográficos y registros de autoridades, respectivamente.

En esencia, lo que se define en ambos modelos es un conjunto de entidades y las relaciones que pueden establecer entre sí (lo cual

Tendencias multidisciplinares...

encaja perfectamente con la filosofía de representación de la realidad de RDF). En concreto, se definen tres grandes grupos de entidades (con sus correspondientes atributos) que engloban las diferentes formas en las que los recursos de información se pueden presentar (distinguiendo entre obra, expresión, manifestación e ítem), las personas o entidades corporativas que de alguna manera se relacionan con estos recursos, y las materias que representan el contenido de dichos recursos (conceptos, objetos, eventos y lugares).

Como vemos, la integración entre los modelos de datos definidos por RDF y FRBR se puede hacer de manera natural, por lo que la futura inclusión de los registros catalogados con RDA en la web semántica no será traumática y permitirá contar de inmediato con una masa crítica de recursos semánticamente accesibles para el desarrollo de aplicaciones y servicios de diversa índole.

El otro gran reto pendiente es la superación del formato MARC por un modelo de codificación más flexible. Con este espíritu surgió Bibframe (Bibliographic Framework) (Library of Congress, 2015), una iniciativa de la Library of Congress que pretende proporcionar un esquema para describir y conectar datos bibliográficos entre sí. De entrada, MARC y Bibframe (<http://bibframe.org/tools/>) presentan notables diferencias que van más allá de lo meramente estético. Bibframe no sólo rompe con la obsoleta codificación alfanumérica de campos y subcampos al utilizar el lenguaje natural para hacerlos fácilmente comprensibles e interpretables de un sólo vistazo, sino que también proporciona nuevas maneras de diferenciar entre los contenidos y sus manifestaciones, de caracterizar entidades bibliográficas (obra, instancia, autoridad y anotaciones), de identificarlas de manera unívoca y de explicitar las relaciones que se definen entre ellas (como las relaciones obra-obra, obra-instancia, obra-autoridad).

No obstante, esta idea de ruptura con los esquemas tradicionales de la catalogación no debe hacernos pensar que todo el trabajo que durante décadas han realizado las bibliotecas caerá en saco roto, ya que gracias a herramientas de mapeo será posible hacer la correspondencia entre campos MARC y Bibframe, y así aprovechar la ingente cantidad de registros que ya existen al usar RDA como un puente entre ambos.

A MODO DE CONCLUSIÓN

La web semántica se perfila cada vez más como una realidad palpable. Si bien los resultados son algo modestos si los comparamos con las expectativas creadas en un principio, no se puede negar que son estas tecnologías las que están marcando el camino hacia donde se dirige la web.

Se ha optado por sacrificar la expresividad (es decir, la capacidad de expresar semántica) en aras de conseguir descripciones más sencillas procesables por máquina a un bajo coste computacional, lo cual me parece un buen trato si con ello conseguimos implantar una web semántica *lightweight* que sirva como base para alcanzar mayores cotas de desarrollo a corto o medio plazo.

Por ello, los profesionales de la información no debemos permanecer ajenos a estos movimientos tecnológicos, sino que nuestra actitud debe ser la de adoptar una posición decidida por la innovación. Sin ir más lejos, la comunidad bibliotecaria se ha convertido en uno de los pilares fundamentales del movimiento Link Open Data (datos abiertos enlazados), cuyo objetivo es compartir y hacer semánticamente accesibles piezas de información, datos y conocimiento en la web semántica. En el caso de las bibliotecas, con lo que se está contribuyendo es con los metadatos de los registros almacenados en los catálogos de las bibliotecas, lo que genera una masa crítica de recursos semánticamente accesibles a disposición de cualquiera que los necesite.

Por lo tanto, no podemos esperar a que un día nos comuniquen que la web semántica ya está plenamente implantada para empezar a considerarla. Debemos contribuir desde ya a la adopción e implantación del modelo no sólo para reivindicar nuestra figura profesional, sino para convertirnos en elementos clave en el nuevo escenario tecnológico que se avecina. No se trata tanto de estar preparados para cuando llegue el cambio, sino de actuar de manera proactiva para que el cambio llegue.

REFERENCIAS BIBLIOGRÁFICAS

- Berners-Lee, T. (2010). 5 Stars Open Data. *5 Star Data*. Disponible el 2 de julio de 2016 en <http://5stardata.info/>.
- Berners-Lee, T., J. Hendler y O. Lassila. (2001). The Semantic Web. *Scientific American*.
- García Marco, F. J. (2013). Schema.org: la catalogación revisitada. *Anuario ThinkEPI*, (1), 169–172.
- Heath, T., y C. Bizer. (2011). Linked Data: Evolving the Web into a Global Data Space. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 1(1), 1–136. doi:10.2200/S00334ED1V01Y201102WBE001.
- Library of Congress. (2015). Overview (Bibliographic Framework Initiative Technical Site - BIBFRAME.ORG). Disponible el 18 de mayo de 2015 en <http://bibframe.org/>.
- LODLAM. (2010). LODLAM Linked Open Data in Libraries, Archives & Museums. Disponible el 15 de noviembre de 2015 en <http://lodlam.net/>.
- Nowak, B. (2009). The Semantic Web. Not a Piece of Cake. Entrada de blog. Disponible el 27 de septiembre de 2015 en <http://bnode.org/blog/2009/07/08/the-semantic-web-not-a-piece-of-cake>.
- Picco, P. y V. Ortiz Repiso. (2012). RDA, el nuevo código de catalogación: cambios y desafíos para su aplicación. *Revista Española de Documentación Científica*, 35(1), 145–173. doi:10.3989/redc.2012.1.848
- Sardo, L. (2004). Functional Requirements for Authority Records. *Bollettino AIB*, 44(4), 455–470.
- Tillett, B. (2005). What is FRBR? A Conceptual Model for the Bibliographic Universe. *Australian Library Journal*, 54 (marzo 2015), 24–30. doi:Article.

- Universidad de Manheim. (2016). Silk: the Linked Data Integration Framework. Disponible el 3 de julio de 2016 en <http://silkframework.org>.
- W3C (2015). RDFa Core 1.1. Disponible el 18 de mayo de 2015 en <http://www.w3.org/TR/rdfa-syntax/>.
- . (2014a). RDF - Semantic Web Standards. Disponible el 18 de mayo de 2015 en <http://www.w3.org/RDF/>.
- . (2014b). RDF Schema 1.1. Disponible el 18 de mayo de 2015 en <http://www.w3.org/TR/rdf-schema/>.
- . (2013). RIF Overview. Disponible el 28 de septiembre de 2015 en <http://www.w3.org/TR/rif-overview/>.
- . (2012a). OWL 2 Web Ontology Language Document Overview. Disponible el 18 de mayo de 2015 en <http://www.w3.org/TR/owl2-overview/>.
- . (2012b). SKOS Simple Knowledge Organization System-home page. Disponible el 29 de septiembre de 2015 en <http://www.w3.org/2004/02/skos/>.
- . (2003). Extensible Markup Language (XML). Disponible el 17 de mayo de 2015 en <http://www.w3.org/XML/>.

Las ontologías en la representación y organización temática de la información bibliotecológica

ADRIANA SUÁREZ SÁNCHEZ

Posgrado de Bibliotecología y Estudios de la Información UNAM

INTRODUCCIÓN

La representación y organización de la información por temas ha sido una actividad esencial de la disciplina bibliotecológica. En consecuencia, a lo largo del tiempo los especialistas del campo desarrollaron procesos y herramientas que les posibilitaron presentar de forma ordenada el contenido albergado en los dominios y recursos de información. Desde tal perspectiva, la representación y organización temática de la información se fundamentó en la determinación del *aboutness* y su concreción a través de sistemas para la organización del conocimiento (listados de temas, clasificaciones, encabezamientos de materias, tesauros, etcétera).

Hasta 1980, la catalogación temática y los sistemas para la organización del conocimiento constituyeron el paradigma en relación al arreglo por temas. Sin embargo, durante las últimas dos décadas del pasado milenio surgieron numerosas innovaciones tecnológicas. En la Bibliotecología, por un lado, aparecieron entidades de información de naturaleza distinta a la impresa; por otro lado, como resultado de la presencia tecnológica, el contexto digital introdujo modificaciones en la manera como los usuarios recuperaban

Tendencias multidisciplinares...

información, se hizo más presente el uso de Internet, la preferencia de búsqueda por navegación (*browsing search*) sobre la búsqueda por interrogación (*query search*) y la recuperación semántica (Chu, 2010).

En este nuevo contexto tecnológico, la búsqueda por temas se ha consolidado como un punto de acceso sustancial, pero se encuentra en un momento renovación. Sobre el tema, los Requisitos Funcionales de los Registros Bibliográficos (FRBR) (IFLA, 1998) y los Requisitos Funcionales para Datos de Autoridad Temática (FRSAD) (IFLA, 2009) indican que los nuevos sistemas de acceso temático deberán:

- Contar con sistemas basados en entidades, atributos y relaciones.
- Ofrecer a los usuarios instrumentos de acceso por temas que permitan la navegación desde la visión conceptual, jerárquica y asociativa.
- Explorar las relaciones entre temas, esto es, explorar las relaciones con el objetivo de entender la estructura de un dominio de información y su terminología.
- Diferenciar entre clases e instancias.
- Establecer un entramado rico de atributos y relaciones que permita a los usuarios definir con precisión el tema de su interés.

Actualmente, los usuarios necesitan recuperar información con mayor precisión, pero además requieren hacerlo de manera vinculada, con posibilidades de navegación y generación de conocimiento. Ante tales exigencias, los sistemas para la organización del conocimiento tradicionales (encabezamientos de materia, tesauros) están en posibilidades de responder a varias de las exigencias. Sin embargo, en el ámbito de la representación y organización de la información, en el contexto digital ha surgido un nuevo sistema denominado *ontología*, cuyo potencial no ha sido debidamente estudiado desde la perspectiva bibliotecológica.

Considerando lo anterior, el presente capítulo tiene por objetivo determinar la viabilidad de las ontologías en la representación/organización temática de la información realizada por la Bibliotecología. Para cumplir con el objetivo, el apartado se divide en tres secciones: (1) definición de las ontologías, (2) la semánticidad de las ontologías y (3) las ontologías en la representación y organización temática.

DEFINICIÓN DE LAS ONTOLOGÍAS

El término *ontologías* ha sido usado con una amplia variedad de significados en distintos campos de conocimiento (Poli, 2002; Staab, 2009); sin embargo, para comprender y recuperar su sentido primario es necesario ir a su origen, enraizado en la Filosofía.

Para la Filosofía, la Ontología —con mayúscula y en singular— es “la ciencia del ser en cuanto ser” (Ferrater-Mora, 1944). En este marco disciplinar, tuvo dos connotaciones. Para Aristóteles, Husserl y Heidegger, era la ciencia de la existencia, del ser último e irreductible, de un *primo ens* en que todos los demás consisten, en este caso es realmente Metafísica, es decir, la ciencia de la realidad, mientras que para Hartmann tenía como misión la determinación de aquello en lo cual los entes consisten, es decir, una ciencia de las esencias y no de las existencias, lo que implicaría una teoría de los objetos (Ferrater-Mora, 1944).

Fuera de la Filosofía, las ciencias computacionales fueron el segundo nicho que empleó el término Ontología, aunque ya no registra la palabra como tal, sino ontología(s) —con minúscula y en singular o plural, según sea un instrumento o varios—. Tal cambio se debe a que en el nicho informático el concepto no se refiere a una disciplina sino a un “instrumento técnico que denota un artefacto diseñado con un propósito” (Gruber T., 2008).

Ahora bien, ¿cómo la Ontología filosófica se traslada a la(s) ontología(s) computacional(es)? Si hacemos una revisión de la historia y las necesidades de las Ciencias Computacionales, encontramos que a principios de la década de los ochenta la disciplina

buscaba métodos para modelar el “mundo” en los sistemas electrónicos, comunicar conocimiento consensuado y establecer sistemas para la transferencia de información entre agentes mecánicos y humanos, así como entre diversos agentes mecánicos.

Derivado de sus necesidades, emprendió un conjunto de estudios sobre organización del conocimiento y de la información. Autores como McGraw y Harbisson-Briggs (1991) analizaron y evaluaron distintos instrumentos: estructuras de referencias conceptuales, mapas cognitivos, taxonomías, diccionarios de conceptos, tesauros, análisis de clusters y otras técnicas que implementaron en contextos programados.

Durante este periodo de estudio y bajo la premisa de que el modelado, la comunicación y la transferencia de información implican la representación de “mundos” (Vickery, 1997), la Ingeniería del Conocimiento reconstruyó postulados para la representación y organización del conocimiento y una parte fundamental de sus estudios se dirigió a la Ontología y la Epistemología filosóficas —consideradas disciplinas fundamentales sobre el tema— de las cuales extrajeron muchos principios. Finalmente, tras consecutivos periodos de trabajo, definieron un nuevo instrumento representador y organizador al que denominaron *ontología* debido a que era “una especificación de una conceptualización” (Gruber, 1993) o “una teoría lógica que da cuenta explícita o parcial de una conceptualización” (Guarino, 1995), manipulada mediante conceptos y las relaciones existentes entre éstos.

La ontología computacional se acerca a la Ontología filosófica en tanto que es una teoría de los objetos, mientras que se aleja de ésta al ser una herramienta construida con fines pragmáticos y no de reflexión sobre las existencias de la realidad. Bajo esta noción de aplicabilidad, las ontologías son “[...] teorías de contenido acerca de los tipos de objetos, propiedades de los objetos y relaciones entre los objetos que son posibles de especificar en un dominio de conocimiento” (Chandrasekaran, Josephson y Benjamins, 1999).

Desde su surgimiento en las ciencias computacionales hasta la actualidad, su trascendencia y fortalecimiento en la representación y organización del conocimiento radica en que

[...] son un medio para modelar formalmente la estructura de un sistema, por ejemplo, las entidades relevantes y las relaciones que emergen de esta observación, la cual es usada para ciertos propósitos [...] la ingeniería ontológica analiza las entidades relevantes y las organiza en conceptos y relaciones, siendo representadas, respectivamente, por predicados unarios y binarios (Guarino, Oberle, & Staab, 2009, p. 2).

Las ontologías se fortalecieron en las ciencias computacionales durante la década de los noventa, aplicadas a Inteligencia Artificial, las bases de datos y la ingeniería del software. Al principio, fueron un asunto exclusivo de la informática; sin embargo, poco a poco llamaron la atención de otros campos disciplinares en los que la representación y organización de la información eran tareas prioritarias. En el campo bibliotecológico, el concepto se introdujo a mitad de la década de los noventa. Así, del espacio de los sistemas inteligentes, las ontologías se extrapolaron a la representación y organización temática de la información en ámbitos particulares de la bibliotecología asociados con dominios.

Hasta la mitad de la década de los noventa, la comunidad bibliotecológica se mantuvo recelosa ante las ontologías. No obstante, paulatinamente el tópico se revistió de interés hasta convertirse en un tema de artículos de revistas y encuentros de la especialidad (foros, congresos, simposios, seminarios). Así, para principios del siglo XX las ontologías se consolidaron como un tema de investigación y la producción literaria sobre el tema se incrementó.¹

El tópico captó la atención de los bibliotecólogos debido a las ventajas ordenadoras, señaladas por especialistas. Sobre el asunto,

1 Al realizar una búsqueda empleando las palabras clave “ontolog?” y “ontologies” en *Library and Information Science*, *Emerald* e *Information Science and Technology*, se encuentra que la literatura sobre ontologías a partir del año 2000 se quintuplicó en comparación con la literatura existente de 1995 a 2000. Francisco García Marco (2008) confirma esta tendencia cuando dice que a partir de 2001 hay un incremento de las publicaciones sobre ontologías en el ámbito computacional y documental.

Tendencias multidisciplinares...

Guerrero y Lozano (1999) señalaron que:

Las ontologías proporcionan un método más rico de organizar el conocimiento, permitiendo realizar deducciones mediante el uso de axiomas entre los conceptos formalizados. El conocimiento formalizado de las ontologías tiene como característica fundamental su disposición para ser reutilizado, por lo que inducirán un gran avance en el entendimiento de diversos sistemas de representación del conocimiento y llevará a fomentar el desarrollo de aplicaciones basadas en estos sistemas (1999, p. 1).

Aunque es importante que los bibliotecólogos conozcan las nociones filosóficas y computacionales de las ontologías, en la disciplina deben ser concebidas como un tipo de sistema para la organización de la información manifestado en relaciones de cuatro tipos: conceptuales, jerárquicas, definitorias y asociativas. Son aplicables a la representación de dominios de conocimiento/información desde la parte de la organización temática y clasificatoria, y hacen explícita la terminología, así como las relaciones entre los elementos formantes. Se diferencian de otro tipo de sistemas a partir de la comprensión del dominio tanto para los humanos como para las computadoras, lo que incide en la posibilidad de inferencias que una máquina puede realizar en beneficio de una recuperación de información más precisa.

Las ontologías coinciden con otros sistemas organizadores del conocimiento propios de la bibliotecología en tanto que son sistemas terminológicos. Mientras que se diferencian en el establecimiento de atributos, relaciones y axiomas que permiten instaurar vinculaciones y extensiones de vocabulario. De manera que, además de presentar los conceptos de un área temática los definen y asocian. Así:

Una ontología abarca la estructura clasificatoria usada por las taxonomías y los tesauros. Pero su principal característica es la presentación de propiedades para cada clase dentro de la estructura clasificatoria. Con una taxonomía completa y propiedades exhaustivas, una ontología funciona de dos modos: como un vocabulario

conceptual y como una plantilla de trabajo que permite el almacenamiento, la búsqueda y el razonamiento, basado en instancias y reglas (Zeng, 2008).

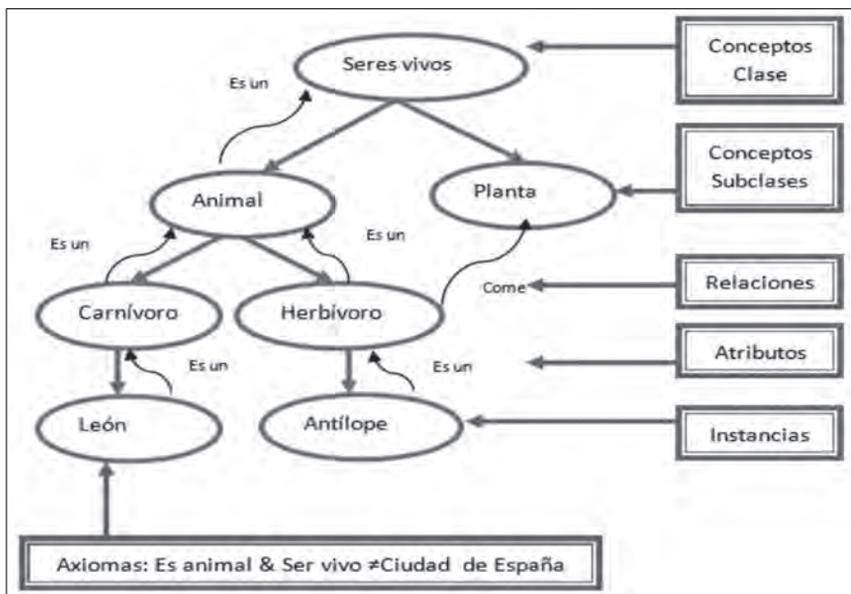
La riqueza ontológica radica en que ofrece una estructura compleja a través de la cual los usuarios del instrumento están en posibilidad de obtener más y mejor información sobre el área representada y encontrar la terminología de esta herramienta vinculada a una posible indización de recursos.

LA SEMANTICIDAD DE LAS ONTOLOGÍAS

Es común escuchar que las ontologías son altamente semánticas, pero ¿qué es la semanticidad? Por semanticidad debemos entender un instrumento representativo y comunicativo en el cual prevalezcan elementos atributivos y lazos asociativos entre términos y situaciones del dominio. En el marco de los sistemas para la organización del conocimiento, la semanticidad se ha observado como una propiedad elemental. Así, por ejemplo, en los encabezamientos de materia persisten asociaciones de sinonimia y reenvíos de uso. Lo mismo sucede en los tesauros donde encontramos vinculaciones de términos amplios (BT), términos específicos (NT) y términos relacionados (NT).

A semejanza de los sistemas antes mencionados, las ontologías poseen semanticidad, pero expandida, es decir, forman un entramado representador más complejo, posible gracias a su estructura. Obsérvese el gráfico siguiente:

Figura 1. Estructura de la ontología



La semántica ontológica radica en la estructura atributiva y relacional que supera a los sistemas para la organización del conocimiento que le anteceden, lo que posibilita mapas semánticos para la total comprensión de dominios de información.

ONTOLOGÍAS EN LA REPRESENTACIÓN Y ORGANIZACIÓN TEMÁTICA

Aunque las ontologías desde el punto de vista teórico parecen ofrecer un conjunto de ventajas frente a otros sistemas organizadores empleados en las bibliotecas, Hepp establece —oportunamente— que “la promesa de las ontologías no es buena para todos los problemas” (Hepp, 2007, p. 10).

Así pues, su función no es sustituir la tendencia de organización temática hasta ahora seguida entorno a la organización de

libros, revistas y algunos otros recursos, sino más bien insertarse como instrumentos para el contexto digital que posibiliten crear simbolizaciones y ordenamientos en las cuales el usuario pueda tener a su alcance el contenido de dominios, navegar en el instrumento e ir delimitando sus temas de interés a la vez que genera conocimiento respecto al campo.

La representación del conocimiento o la información es el primer objetivo de las ontologías. Una vez logrado su primer objetivo, si el instrumento se quiere vincular con la indización de un conjunto de recursos, se puede optar por esta posibilidad que deberá establecerse desde su planeación y diseño.

Es importante señalar que tras la elección de una ontología como herramienta representadora y organizadora, subyace un grupo de factores que para fines del capítulo llamaremos “principios generadores”. Éstos son aquellas necesidades que exigirán la creación de una ontología. La presencia de las ontologías no está motivada por un entusiasmo momentáneo o una idea de que los sistemas vinculados a las tecnologías serán la tendencia imperante en el futuro, más bien deberán derivarse de las necesidades específicas de un contexto.

Sobre el asunto, la teoría señala que las principales aplicaciones de las ontologías son:

- a) Necesidades representadoras con características específicas (dominios nuevos, dominios en los que sea necesarias explicitaciones conceptuales, dominios no abarcados adecuadamente por otros sistemas para la organización del conocimiento) bajo el esquema de ontologías profundas.
- b) Dominios específicos de conocimiento con necesidades particulares de organización y recuperación: terminología, jerarquía, descripción, etc., bajo el esquema de ontologías ligeras.
- c) Necesidades de organización-recuperación que se inserten dentro del modelo de la web semántica (Berners, 2001).

CONCLUSIONES

Las ontologías son un sistema para la organización del conocimiento nacido en las ciencias computacionales que se extrapoló a la disciplina debido a su riqueza para representar y organizar conocimiento e información. En vista de sus ventajas en términos de arreglo de la información, deben ser un tema de estudio desde su ámbito teórico pero, especialmente, deben probarse sus aplicaciones prácticas en ámbitos bibliotecológicos. Actualmente, contamos con gran cantidad de artículos sobre el tema y es momento de que una institución o algún grupo de investigación tome en sus manos la tarea de implementación de ontologías que afirmen o desmientan la teoría.

La disciplina bibliotecológica vive momentos de renovación. En consecuencia, los especialistas del campo deben transformarse en verdaderos profesionales de la información. Es decir, deberán ser capaces de catalogar libros o revistas impresas y, al mismo tiempo, tener habilidades para construir una estructura clasificadora de recursos digitales, ordenar dominios particulares de conocimiento o implementar una herramienta vinculada al contexto web que además tenga la capacidad de vincular semánticamente cada una de sus entidades formantes.

Finalmente, sólo resta mencionar que la investigación en el ámbito de la organización del conocimiento y la información es un campo de trabajo altamente rico y productivo en el que se debe invertir trabajo cotidiano y esfuerzos de investigación en beneficio de las unidades de información.

REFERENCIAS BIBLIOGRÁFICAS

- Berners-Lee, T., J. Hendler y O. Lassila. (2001). The Semantic Web. *Scientific American*, 284 (5), 34-43.
- Chandrasekaran, B., J. R. Josephson y V. R. Benjamins (1999). What Are Ontologies, and Why Do We Need Them? *IEEE Intelligent Systems*, (enero-febrero), 20-26.

- Chu, H. (2010). *Information Representation and Retrieval in the Digital Age*. Medford: American Society for Information Science and Technology by Information Today.
- Ferrater-Mora, J. (1944). *Diccionario de filosofía*. México: Atlante.
- García-Marco, Francisco Javier. 2008. El nacimiento y el despegue de la investigación moderna sobre ontologías. *Anuario ThinkEPI*, pp.122-125.
- Gruber, T. (2008). Ontology. En *Encyclopedia of Database Systems*, Ling Liu y M. Tamer Özsu (eds.). Heidelberg: Springer Verlag. Disponible el 17 de febrero de 2016 en <http://tomgruber.org/writing/ontology-definition-2007.htm>.
- . (1995). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human and Computer Studies*, 43, 907-928.
- Guarino, N. (1995). Formal Ontology, Conceptual Analysis and Knowledge Representation. *International Journal of Human-Computer Studies*, 43 (5-6), 624-640.
- Guarino, N., D. Oberle y S. Staab. (2009). What is an Ontology? En S. Staab y R. Studer (eds.), *Handbook of Ontologies*. Heidelberg: Springer-Verlag.
- Guerrero, V. y A. Lozano Tello. (1999). La representación y organización del conocimiento en sus distintas perspectivas: su influencia en la recuperación de la información. *Actas del IV Congreso ISKO-España*, pp, 25-31. Granada.
- Hepp, M. (2007). *Ontology Management: Semantic web Services and Business Applications*. Nueva York: Springer.

Tendencias multidisciplinares...

IFLA. (2009). *Requisitos funcionales para datos de autoridad de materia*. Disponible el 10 de octubre de 2016 en <http://www.ifla.org/files/assets/cataloguing/frsad/frsad-final-report-es.pdf.pdf>.

———. (1998). *Requisitos funcionales de los registros bibliográficos: informe final*. La Haya, Holanda: Ministerio de Cultura.

McGraw, K. y K. Harbisson-Briggs. (1991). *Knowledge Acquisition*. Englewood Cliffs, Nueva Jersey: Prentice-Hall.

Poli, R. (2002). *Ontological Methodology*. *Human-Computer Studies*, 56, 639-664.

Staab, Steffen y Studer, eds. (2009). *Handbook on Ontologies*. Nueva York: Springer.

Vickery, B. (1997). Ontologies. *Journal of Information Science*, 23(4), 277-286.

Zeng, M. (2008). Knowledge Organization Systems (KOS). *Knowledge Organization*, 35(2-3), 160-182.

USUARIOS, CONSUMIDORES
Y PROSUMIDORES DE
INFORMACIÓN

El uso de metadatos para la gestión del patrimonio digital mundial en la Sociedad de la Información

LUIS FERNANDO RAMOS SIMÓN

SILVIA COBO SERRANO

Universidad Complutense de Madrid

INTRODUCCIÓN

Los metadatos son elementos de información estructurados, embebidos en la fuente del objeto digital al que hacen referencia para otorgarle un significado, un contexto y una organización (Muñoz de Solano, 2014, p. 154). Dicho de una forma más sencilla, son datos que describen las propiedades de otros datos para facilitar su gestión y uso al vincular la información descriptiva con el objeto digital. En resumen, para un bibliotecario-documentalista, como dice Méndez Rodríguez (2002, p. 35), “los metadatos son un tipo de datos que usamos para describir el contenido, la estructura, la representación y el contexto de algún conjunto de datos específicos”.

Los metadatos comenzaron a utilizarse en 1968 en el ámbito de la informática. Más tarde, en los años setenta, se hicieron populares en Europa y Estados Unidos, donde su uso se vinculó a documentos militares. Como es bien sabido, “metadata” es una palabra de origen griego que significa “datos sobre los datos” y hoy en día se emplean en todo tipo de sectores. La información sobre metadatos define los requisitos necesarios sobre los datos en cuanto a su estructura, contenido y precisión para incorporarlos a una base

de datos. Estas descripciones constituyen un modelo en el que también se incorporan protecciones para evitar el mal uso de los datos en el *dataset*. En algunos sectores, se emplean metadatos estandarizados, como sucede en el ámbito de la información geográfica (GIS). Estos estándares pueden, por un lado, estar reconocidos por organismos internacionales de estandarización y ser un estándar dominante por su aplicación generalizada en un grupo de organizaciones o, por otro lado, solamente ser usados por algunas empresas del sector privado. Los elementos de los metadatos pueden ser separados en tres tipos: diccionario de definición de metadatos, directorio de metadatos y, por último, contenido de metadatos, también llamado metadatos de metadatos (Aalders, 2005).

Se puede decir que todos los campos científicos organizan sus servicios de información electrónica incorporando metadatos. Sin duda, los metadatos son tributarios de la larga tradición bibliotecaria de los catálogos, adaptando ese amplio saber y experiencia a las nuevas formas digitales de organización de la información. De hecho, los metadatos descriptivos más conocidos proceden del mundo bibliotecario (por ejemplo, Dublin Core, desde 1995). Los metadatos proporcionan de una forma sencilla la información más relevante sobre el contenido en donde se representan o apuntan, de tal modo que indican a las máquinas y las personas sus características y aspectos sobre lo que son, no son y lo que se puede encontrar en ellos. De ahí, la importancia de los buenos metadatos para recuperar de la información, así como para facilitar su descubrimiento y hacer visible el objeto y la colección (Dangerfield *et al.*, 2015, p. 19).

Tradicionalmente, los estudiosos de esta materia establecen varias clasificaciones de los metadatos (descriptivos, administrativos, técnicos, de conservación, funcionales, explícitos...), si bien los metadatos descriptivos y administrativos son los más comunes. Los primeros (descriptivos) se emplean para la representación e identificación de los recursos para organizar los objetos digitales (catálogos, índices...), mientras que los segundos (administrativos) se orientan hacia la gestión de los recursos en cuanto a su ubicación, uso, versiones, etcétera.

Por tanto, se pretende llamar la atención sobre la creciente importancia de los metadatos en el ámbito del patrimonio digital y su funcionalidad en la gestión de los recursos culturales.

LOS METADATOS EN EL PATRIMONIO DIGITAL

Al igual que los datos, los metadatos han alcanzado en los últimos años una notable importancia económica, como se ha mencionado. Aunque existen metadatos en todo tipo de productos y sectores, las bibliotecas y las editoriales tienen mucha experiencia en su gestión.

Algunos datos de esta realidad, referentes al sector académico, manifiestan que hay unos cincuenta millones de libros y revistas de los que podrían ser extraídas referencias bibliográficas. A pesar de lo que parece, el sector está muy atomizado y las fuentes de datos muy dispersas, por lo que hay varios intentos para tratar de crear repositorios abiertos que permitan una fácil gestión de tales referencias (Open Citation Corpus, por ejemplo).

En Wikipedia, aparecen referencias de más de 125 bases de datos que gestionan metadatos bibliográficos. Las dos principales fuentes, ambas accesibles por suscripción, son Thomson Reuters Web of Science y Scopus (Elsevier). La tercera, Google Scholar, es de libre acceso y pertenece a un operador comercial. Microsoft también intentó operar Microsoft Academic Search en 2009, pero el servicio desapareció definitivamente en 2016.

En el ámbito del acceso abierto, los principales servicios que dan acceso a referencias son las siguientes:

- Open Access Subset of PubMed Central, que reúne 468.805 artículos. En su conjunto, el servicio da acceso a las referencias de unos 2,1 millones de artículos.
- Preprints, en el repositorio Arxiv, que se estiman en más de 1,1 millones.
- CiteSeerX con 1,2 millones de artículos, principalmente del área de Informática.
- CitEc (Citation in Economics) contiene 545.641 documentos.

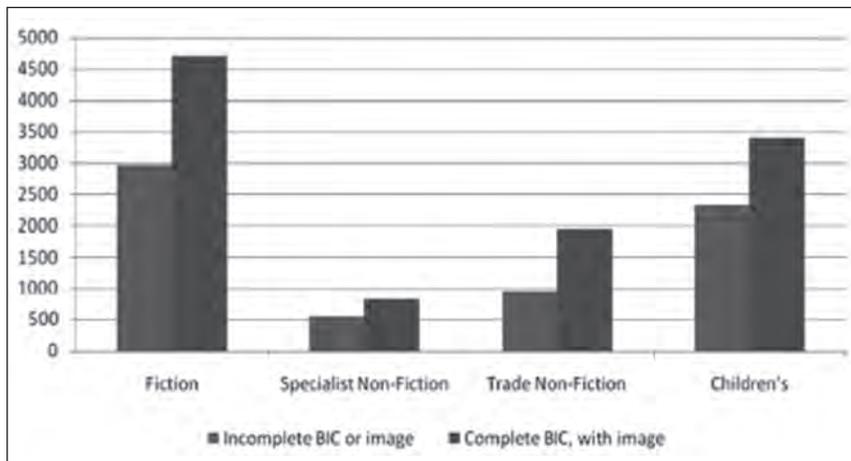
Tendencias multidisciplinares...

En conjunto, Shotton (2013) calcula que las referencias en acceso abierto están en torno al 4 por ciento. La fuerte presencia de las empresas comerciales y la resistencia para facilitar la puesta de datos en común dan una idea del interés económico por estos metadatos.

Para poner de relieve la importancia comercial de los metadatos, son más contundentes aún los resultados del informe elaborado para la firma Nielsen por Breedt y Walter (2012), en el que se pone de manifiesto la estrecha vinculación entre las ventas comerciales y los metadatos en el sector del libro.

La tendencia de este informe apunta que la relación entre la calidad de los metadatos y las ventas de un libro es cada vez mayor. En concreto, el informe comparó la venta de libros en el mercado inglés en 2011 según el tipo de metadatos distribuidos en el proceso de comercialización y venta: En un caso, los datos básicos del libro (autor, título, editorial, etc.); en otro, aquellos libros en los que se mejoraron los metadatos (imagen de portada, descripciones larga y corta, biografía del autor y reseña de la obra).

Figura 1. Comparación de las ventas de libros en relación a los metadatos según la temática.



Fuente: Breedt, André, y David Walter (2012, p.7).

Las principales conclusiones del análisis son:

- Las ventas con buenos metadatos y la imagen del libro son siempre muy superiores a los que tienen los datos incompletos o que carecen de imagen.
- La diferencia es más notoria cuando las ventas son en línea y no en la librería. Asimismo, cuanto mayor es el número de metadatos mejorados, se produce un mayor aumento de las ventas en línea.
- Las diferencias en las ventas son mucho mayores en los libros de ficción. En los libros especializados, las diferencias son menores. Las ventas son un 140 por ciento mayores cuando hay varios metadatos mejorados.
- Se observan diferencias claras cuando se analizan los libros con más metadatos, en donde el incremento de los mismos conlleva a un incremento en las ventas.
- El metadato más influyente es la imagen de portada. Los títulos con imagen tuvieron ventas superiores en un 268 por ciento que aquellos sin imagen.
- Se venden casi siete veces más los libros que tienen metadatos estándar, imagen y metadatos mejorados que los que no completan los metadatos estándar, no tienen imagen y no usan metadatos mejorados.

Por otro lado, resulta necesario definir el concepto de patrimonio cultural en el ámbito digital. En este sentido, entendemos el nuevo patrimonio documental digitalizado del que son titulares las bibliotecas, los archivos, los museos y todas las tecas, con mención especial por su importancia a las filmotecas y fototecas o colecciones fotográficas. Este conjunto documental es un patrimonio creciente, cuya presencia en Internet ya es visible, aunque sólo se haya digitalizado y puesto en línea una pequeña parte del total de documentos que conservan estas instituciones.

La carta sobre la preservación del patrimonio digital (Unesco, 2003), en su artículo primero, considera patrimonio digital a los “recursos de carácter cultural, educativo, científico o administrativo

e información técnica, jurídica, médica y de otras clases, que se generan directamente en formato digital o se convierten a éste a partir de material analógico ya existente”.

El objetivo señalado por la carta en la conservación de este patrimonio es el de hacerlo accesible al público, así como elaborar estrategias encaminadas a su preservación, si bien distingue entre los elementos en dominio público —cuyo acceso no debería estar sujeto a requisitos poco razonables— y el patrimonio privado para el que propugna un equilibrio entre los derechos de sus titulares y el interés del público por tener acceso a dichos elementos. En el aspecto institucional, la carta otorga un papel determinante en esa función a las bibliotecas, archivos y museos (filmotecas y tecas en general) para llevar a cabo esas tareas en colaboración con el sector privado (creadores, editores, productores, distribuidores, fabricantes de equipos y software...).

Los metadatos —que, como se ha mostrado, son claves en el sector cultural comercial— también son esenciales para construir el patrimonio cultural digital. Así se pone de manifiesto en algunos informes, en los que se destaca el papel de los metadatos como uno de los activos más valiosos para la reutilización dentro del sector cultural, junto con los archivos, las obras en dominio público inéditas y las fotografías (Davies, 2009b).

Para lograr un patrimonio digital de alcance mundial, es preciso que los metadatos sean libres, como lo han entendido European (más de 53 millones de objetos digitales) y la Digital Public Library of America (DPLA) (13,3 millones de objetos en mayo de 2016), probablemente las dos iniciativas más importantes a nivel mundial en este ámbito.

La discusión sobre el libre intercambio de metadatos está justificada por el alto coste que conllevan la recopilación y organización de los metadatos en las grandes instituciones culturales. Los pocos datos que se conocen ponen de relieve que los ingresos por este concepto se cuentan entre los más importantes obtenidos por las instituciones culturales en países como Reino Unido o Alemania (Davies, 2009a).

FUNCIONALIDAD DE LOS METADATOS EN LA GESTIÓN DE LOS RECURSOS CULTURALES

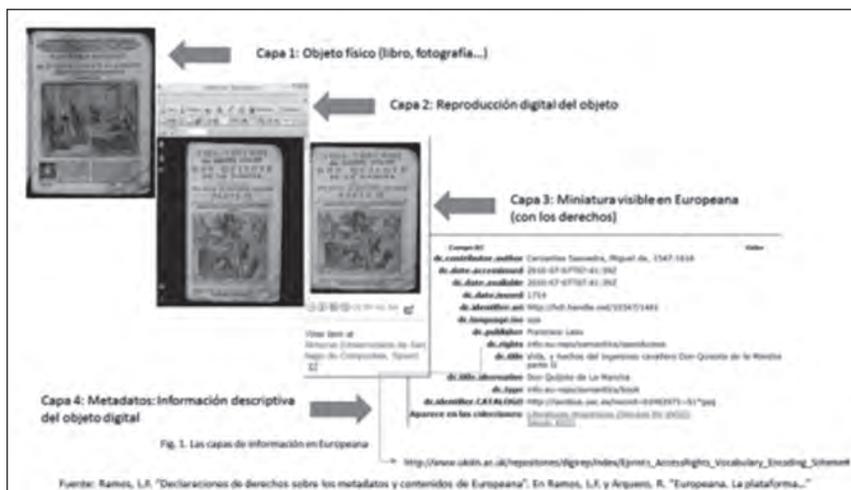
Es importante distinguir entre el acceso abierto a los metadatos que gestionan un recurso digital del acceso y el contenido mismo. Ambos aspectos se indican en los metadatos, por eso a veces se crea confusión. Por ejemplo, cuando de una base de datos que contiene metadatos se dice que es de libre acceso o abierta, parece indicarse que tanto el acceso a la base de datos como al contenido son libres, cuando en absoluto es así.

Podemos observar la importancia de los metadatos en la construcción del patrimonio digital al analizar cómo cumplen su papel en Europeana y en la DPLA. Estas plataformas en esencia sólo mantienen una base de datos compuesta por la información catalográfica y una imagen digital en miniatura de los objetos de la colección, en forma de capas. A partir de este esquema, la búsqueda del usuario conduce al enlace con el objeto digital.

La complejidad de estos procesos técnicos de gestión de los metadatos es notable. Desde la creación de Europeana y la implantación del modelo de datos EDM (European Data Model), el enriquecimiento de los metadatos ha continuado siendo una de las principales vías para mejorar la calidad de los datos, el contexto de los objetos digitales y la información multilingüe. Estas actividades en la actualidad se llevan a cabo a través de los distintos grupos de trabajo para mejorar los distintos procesos de enriquecimiento de los datos y hacer más transparentes los intercambios de información (Europeana, 2013 y 2015; Dangerfield *et al.*, 2015).

El símil de las capas, utilizado por Keller (2011), nos permite mostrar de una forma sencilla los distintos niveles de abstracción que van desde el objeto material (por ejemplo, un cuadro de pintura), su imagen digital y después la representación miniaturizada, acompañada de los datos descriptivos del objeto (por ejemplo, título, autor, museo, titular de los derechos, etc.). Europeana organiza este esquema en cuatro capas, tal como se muestra en la figura 2.

Figura 2. Capas de los objetos digitales en las instituciones culturales.



- 1ª Capa de objetos físicos: es el objeto, como un libro, una pintura, escultura u otro elemento del patrimonio cultural o científico.
- 2ª Capa de reproducciones digitales: corresponde al objeto anterior digitalizado por cualquier procedimiento, ya sea escaneado, fotografía o cualquier otra forma de archivo digital. Si el objeto tiene origen digital, por ejemplo, una imagen o un libro digital, la primera y la segunda capa coinciden.
- 3ª Capa de miniaturas (*previews*, vistas previas o miniaturas): es una visualización en baja resolución que permite una rápida identificación de las características generales del objeto, ya sea una pequeña imagen digital, una secuencia de imagen o video, o una muestra de cualquier cualidad de la imagen digitalizada. Como hemos visto para el caso de los libros comerciales, es un elemento atractivo para captar el interés de los usuarios hacia el contenido y actúa a modo de metadato complementario. Es un archivo de imagen, generalmente en formato jpg o la portada de un archivo pdf de pequeño tamaño, menos de 200 píxeles. Los proveedores

de los objetos digitales deben proporcionar un *link* a una imagen de mayor tamaño y calidad que la miniatura misma. Cuando no se puede usar una miniatura adecuada al objeto representado, entonces se debe sustituir por una imagen genérica que representa uno de los formatos comunes: texto, imagen, vídeo, sonido y 3D.

- 4^a Capa de metadatos: contiene la información sobre los objetos mediante una descripción catalográfica del objeto físico y digital de las capas anteriores.

Esta capa de información de metadatos es de libre acceso y uso para todos los usuarios, lo que significa que el proveedor de los metadatos renuncia a los derechos sobre los mismos mediante la iniciativa de licencia CC0. Esta licencia, impulsada por la organización Creative Commons, tanto en estas plataformas de recursos culturales como en otros entornos de datos abiertos, ha favorecido a nivel internacional la apertura, la interoperabilidad y los estándares comunes, lo que está cambiando la forma de ver el sector cultural y el uso del contenido digital en línea. De hecho, la DPLA —abierta en abril de 2013— está basada en el European Data Model (Modelo de Datos de Europea) y opera con el mismo esquema de derechos en cuanto a los metadatos, “lo que supone que los dos masivos conjuntos de datos de ambos lados del Atlántico sean interoperables” (Cousins y Daley, 2013, p. 35). En efecto, a los más de cincuenta millones de objetos digitales, hay que sumar los más de siete millones que ya ofrece el portal estadounidense. De hecho, es posible la búsqueda conjunta sobre ambas plataformas. La búsqueda está disponible en: <http://www.digibis.com/dpla-europeana>.

Entre los objetivos básicos de Europea está, por un lado, que los metadatos financiados con fondos públicos estén ampliamente disponibles para todos y, por otro, que las instituciones proveedoras de contenido no se introduzcan nuevos derechos después de la digitalización de objetos que están en dominio público, en línea con las recomendaciones del Comité de sabios (Comisión Europea, 2011).

CONCLUSIONES

Tras el estudio de la temática, las principales conclusiones que se pueden extraer son las siguientes:

- Los metadatos identifican, orientan y condicionan los contenidos.
- Los metadatos son decisivos para asegurar la colaboración a largo plazo entre las instituciones al facilitar el acceso, la recuperación y organización de los contenidos.
- En el ámbito del patrimonio digital, los metadatos tienen una notable importancia económica en los sectores comerciales, de ahí su atomización y la dificultad de ponerlos en acceso abierto. Por ejemplo, en el sector del libro está demostrada la eficacia de los metadatos mejorados para aumentar las ventas.
- El papel clave de los metadatos en el sector cultural-comercial demuestra su importancia para construir el patrimonio cultural digital, de ahí la importancia de favorecer la reutilización libre de estos metadatos en la economía digital, a pesar del alto coste que supone su recuperación y organización en las grandes instituciones culturales.
- Las grandes plataformas de acceso en línea al patrimonio cultural, como Europeana y DPLA, disponen una capa de metadatos que contiene una descripción catalográfica del objeto físico y digital de las capas anteriores. Estos datos son de libre acceso para todos a través de la licencia libre CC0. Esta modalidad, impulsada por Creative Commons, favorece la apertura de datos y unos estándares comunes.
- Como conclusión general, los metadatos deben ser libres y contener orientaciones sencillas sobre los derechos de los contenidos.

REFERENCIAS BIBLIOGRÁFICAS

- Aalders, H. (2005). An Introduction to Metadata for Geographic Information. En H. Moellering, H. Aalders y A. Crane. *World Spatial Metadata Standards*. Elsevier.
- Breedt, André, y D. Walter (2012). *White Paper: The Link Between Metadata and Sales*. Disponible el 24 de julio de 2016 en: http://www.nielsenbookdata.co.uk/uploads/3971_Nielsen_Metadata_white_paper_A4%281%29.pdf.
- Comisión Europea (2011). *The New Renaissance: Report of the 'Comité des Sages': Reflection Group on Bringing Europe's Cultural Heritage Online*. Disponible el 24 de julio de 2016 en: www.ace-film.eu/wpcontent/uploads/2011/02/CdS_Final_report.pdf.
- Cousins, J. y B. Daley (2013). Europeana: el patrimonio cultural de Europa disponible en línea para todo el mundo, *Boletín de Anabad*, 63, 3.
- Dangerfield, M. C. et al. (2015). *Report and Recommendations from the Task Force on Metadata Quality*. Disponible el 24 de julio de 2016 en: http://pro.europeana.eu/files/Europeana_Professional/Publications/Metadata%20Quality%20Report.pdf.
- Davies, Rob (2009a). *PSI in the Cultural Sector*. European PSI Platform. Topic Report No: 4
- . (2009b). *Economic and Social Impact of the Public Domain: EU Cultural Institutions and the PSI Directive*. Informe para la Comisión Europea. Rightscom/Comisión Europea.
- Europeana (2015). *Report on Enrichment and Evaluation*. Disponible el 24 de julio de 2016 en: http://pro.europeana.eu/files/Europeana_Professional/EuropeanaTech/EuropeanaTech_taskforces/Enrichment_Evaluation/FinalReport_EnrichmentEvaluation_102015.pdf.

Tendencias multidisciplinares...

- Europeana (2013). *Europeana Data Model Primer*. Disponible el 24 de julio de 2016 en: http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf.
- Keller, P. (2011). *The European Licensing Framework, Next Steps*. Disponible el 24 de julio de 2016 en: <http://www.bnl.public.lu/fr/actualites/communiques/2011/11/europeana/keller--2011-11-28.pdf>.
- Méndez Rodríguez, E. (2002). *Metadatos y recuperación de información. Estándares, problemas y aplicabilidad en bibliotecas digitales*. Gijón: Trea.
- Muñoz de Solano, B. (2014). Metadatos en Europeana. En L.F. Ramos Simón y R. Arquero Avilés. *Europeana. La plataforma del patrimonio cultural europeo*, p. 153-176.
- Shotton, D. (2013). Open citations. *Nature*, 502, 295:7.
- UNESCO (2003). *Carta sobre la preservación del patrimonio digital*. Disponible el 24 de julio de 2016 en http://portal.unesco.org/es/ev.php-URL_ID=17721&URL_DO=DO_TO PIC&URL_SECTION=201.html.

Interoperabilidad global en las comunidades bibliotecaria, archivística y museográfica

ARIEL ALEJANDRO RODRÍGUEZ GARCÍA

Universidad Nacional Autónoma de México

INTRODUCCIÓN

En lo referente al desarrollo de los metadatos, las comunidades bibliotecaria, archivística y museográfica están viviendo tiempos de cambio. No hace mucho se decía que los datos bibliográficos se habían preparado para ser leídos e interpretados sólo por los bibliotecarios y usuarios, pero con la automatización en los procesos, esos datos han pasado la frontera para que en la actualidad se sustente que los datos además deban ser interpretados por una computadora o robot.

La tecnología de la World Wide Web (W3) y, más específicamente, el crecimiento de los estándares para el uso de los datos vinculados (*linked data*), mostraron que con los metadatos había mayor posibilidad de interconectar los recursos de información, puesto que la web provee plataformas para vincular éstos, pese a su procedencia.

De ahí que surgieran preguntas tales como, ¿los datos bibliográficos pueden ponerse directamente en la web semántica? ¿Qué tipo de reglas o lineamientos permitirán la explicación del uso y reuso de los datos? ¿Cuáles serán las iniciativas de metadatos que

permitirán introducir modelos para entender las partes componentes de un sistema datos?

Los últimos tres lustros, las preguntas anteriores y otras más que han salido a la luz de las investigaciones sobre los metadatos ya no parten del cuestionamiento ¿qué se entiende por metadato? Ahora inician con interrogantes respecto a cómo se debe entender la participación de estas comunidades en las iniciativas sobre los Vocabularios Abiertos Vinculados (Linked Open Vocabularies por su término en inglés), así como los Datos Abiertos Vinculados (Linked Open Data).

Si estamos de acuerdo con este preámbulo, nos gustaría decir que el propósito de este texto es explicar, *grosso modo*, la dirección general que está siguiendo la línea de investigación sobre la interoperabilidad global y los datos vinculados en las comunidades bibliotecaria, archivística y museográfica o comunidades culturales.

LA INTEROPERABILIDAD, UN ASUNTO PARA REFLEXIONAR EN LAS COMUNIDADES CULTURALES

El primer acercamiento que se tiene respecto a la interoperabilidad, nos refiere Caplan (2003), es que generalmente hablamos acerca de la interoperabilidad en la búsqueda, o la habilidad que se tiene para representar un conjunto diverso de registros de metadatos para obtener un resultado significativo. Es decir, los diferentes esquemas de metadatos interactúan entre sí con la finalidad de facilitar el cruce de datos.

El cruce de datos o *crosswalk* ha servido como base para desarrollar las especificaciones de conversión de registros físicos de un esquema a otro mediante el intercambio de los registros, lo que ha contribuido al campo bibliotecario, de acuerdo con Caplan (2003), a crear catálogos de unión o también a la cosecha de metadatos. De ahí que, por ejemplo, la Biblioteca del Congreso constituyera el *crosswalk* de MARC21 con Dublin Core, el FGDC Content Standards for Spatial Metadata, el Global Information Locator Service (GILS), entre otros más.

Con el desarrollo de esquemas de metadatos y el aumento en el cruce de datos entre las comunidades culturales, la interoperabilidad está siendo un asunto no sólo de abstracción sino de estudio, debido a que se está presenciando la forma en que diversos sistemas interactúan unos con otros sin importar el tipo de hardware o software.

Lo más difícil que puede apreciarse sobre la interoperabilidad es resolver lo concerniente a las diferencias semánticas. Esto es, cada esquema de metadatos conceptúa y nombra de manera distinta sus elementos descriptivos, lo que produce que las reglas de operación en cada esquema no concuerden, y por eso requieren ser mapeadas para interpretarse semánticamente en cada esquema.

En suma, la perspectiva de la interoperabilidad vista desde el actual modelo conceptual sugerido por la IFLA a través de su modelo conceptual de los Requerimientos Funcionales, tiene la manera de unir y armonizar diferentes reglas de catalogación, lo que hace que se promueva la interacción más real y palpable entre estándares de metadatos y normas de descripción. Además, disemina la forma de hacer visibles las categorías inherentes a lo que se está reconociendo actualmente como la metadescripción de los recursos de información.

LAS DIFERENCIAS SEMÁNTICAS COMO BARRERA DE LA INTEROPERABILIDAD

Existen barreras que causan dificultades para que la interoperabilidad se lleve a cabo íntegramente, entre las cuales se encuentran las diferencias semánticas, de práctica, de representación, de vocabulario, de problemas especiales entre la descripción de las unidades o ítems y las colecciones, de las versiones múltiples y las diferencias de lenguaje múltiple. No ahondaremos en cada una de ellas porque no es motivo de este trabajo.

Desde nuestro punto de vista, las diferencias que consideramos que importan para comprender el desarrollo de los metadatos en

las comunidades culturales son aquellas enunciadas como diferencias semánticas debido a que son las que, por regla, de acuerdo con Caplan (2003), no necesariamente deben tener el mismo significado ya que a cada elemento es posible asignarle un grupo de reglas que define el esquema de metadatos.

Al respecto, Mitchell (2013) refiere que en las comunidades bibliotecaria, archivística y museográfica (*LAM communities* en inglés) se han desarrollado nuevos marcos normativos para la catalogación como son las RDA (Resources Description and Access), las DACS y las CCO (Cataloging Cultural Objects) que, consecuentemente, impactarían en la creación de nuevos sistemas de información.

Mitchell señala que en la comunidad bibliotecaria, la discusión se centra en la iniciativa del BIBFRAME, esencialmente en la polémica subyacente sobre la implementación de las especificaciones de RDA. Algo similar pasa en la comunidad archivística con el estándar Encoded Archival Description (EAD por sus siglas en inglés) y el DACS (por sus siglas en inglés), que es la Descripción de Archivos: un Estándar de Contenido y el Encoded Archival Context (EAC por sus siglas en inglés) que es un nuevo sistema de información que sirve para definir los metadatos de archivos y las necesidades de administración de los objetos. Finalmente, en la comunidad museográfica se ha trabajado en la actualización del modelo de referencias conceptual (CIDOC-CRM por sus siglas en inglés), así como las especificaciones del CCO.

Es evidente que los cambios referidos por los nuevos marcos normativos han puesto a las comunidades culturales a prestar atención de manera distinta en lo que está sucediendo en la comunidad de metadatos. Esta última, típicamente se enfoca a explorar cómo los sistemas de información y las estructuras que lo soportan necesitan de una visión más amplia del cruce de información entre comunidades, así como de acceso y agregación de datos en los repositorios.

En definitiva, el marco de normativo en las comunidades culturales está puntuando que cualquier iniciativa que surja deberá enfocarse en definir las técnicas, los estándares, los sistemas y los servicios que permitan unir los cambios tanto en las fuentes de

información como en las necesidades de patrones útiles para los usuarios de dichas comunidades.

Considerando que la tendencia en el uso de los metadatos es la anterior, lo que nos parece valioso indicar es que la historia sobre los cambios en los marcos normativos es muy reciente, por ello es que las comunidades culturales se encuentran en un constante estudio sobre lo que podría acontecer en un futuro inmediato respecto a la implementación de las normas descriptivas y cambios en los sistemas de información.

En la comunidad bibliotecaria, señala Mitchell (2013), hablar de BIBFRAME y RDA es remitirse al hecho de que MARC como estándar armonizó en gran medida el desarrollo de las bibliotecas al permitir señalar detalles para la interconexión que pudieran crear, compartir y validar tanto los registros bibliográficos como los de autoridad. De ahí la polémica entre permanecer a la zaga o avanzar hacia la interoperabilidad global.

En cuanto a la comunidad museográfica, los protocolos que han seguido en la última década han puesto en práctica modelos de catalogación y métodos de diseminación que han superado sus discrepancias en el uso de las CCO y el modelo de referencia CIDOC.

Por lo que se refiere a la comunidad archivística, sus marcos normativos como el EAD y el DACS vienen acompañados de estándares sustentados en la tesis de dar cabida a nuevos tipos de recursos, así como a la persistente tarea de posicionarse en el ambiente digital.

ESQUEMA GENERAL PARA DISCUSIÓN SOBRE INTEROPERABILIDAD GLOBAL

A propósito de los nuevos marcos normativos en las comunidades culturales y el interés por hacer interoperables estas iniciativas, desde 2013 se comenzaron a perfilar las estrategias para que las normativas confluyan en una dirección. Es decir, se proyecta que el diseño y la aplicación de una interoperabilidad global se ejecute con base en lo establecido por los datos abiertos vinculados

(LOD, Linked Open Data) y el Vocabulario Abierto Vinculado (LOV, Linked Open Vocabularies) con la finalidad de crear un bloque de metadatos comunes.

Al respecto, Mitchell (2013) refiere que hay cinco bloques de estructuras de metadatos que permiten definir el bloque del LOD, como el modelado de datos, las reglas de contenido, los esquemas y vocabularios de metadatos, la serialización de los datos y el intercambio de los datos. Por supuesto que el esquema general del bloque LOD para las comunidades culturales plantea una forma de explorar los aspectos de los metadatos asociados al BIBFRAME, CIDOC-CR y el DACS con la finalidad de considerar las diferencias y similitudes entre sus normativas, así como sus servicios. De ser cierto lo anterior, surge la pregunta, ¿en qué proceso se encuentra la adopción de los Datos Abiertos Vinculados (LOD por sus siglas en inglés) en las comunidades culturales? Sobre esto, Godby, Wang y Mixter (2015) señalan que hace catorce años, hablar de la Web Semántica era como una novela de ciencia ficción, especialmente con el detalle como el que ahora se hace, ya que a lo largo de estos años las colecciones de documentos se han transformado y han cambiado la estructura que consideraba la lectura humana un sustento, por lo que ahora se tiene como un sistema estructurado de datos que están asociados a los objetos del mundo real.

Con certeza, podemos argumentar que este cambio no sólo ha modificado el método para leer, sino para elaborar los documentos y la representación de su contenido, también ha influido en la percepción y estructuración de los marcos normativos, los cuales aprovechan al máximo la tecnología para que la descripción y representación de los recursos de información sea más apegada al ambiente digital.

Desde el punto de vista de Coyle (2012), la web semántica y los datos vinculados no solamente son asociaciones de datos creadas por las tecnologías de la web, sino que los datos vinculados van más allá. Esto es, definen estructuras básicas que hacen que los datos por sí mismos generen paquetes de datos para operar en el ambiente de Internet.

Por consiguiente, los datos vinculados no son simples estándares o formatos como los ha citado Berner-Lee, sino que su énfasis es crear las relaciones para que nuestros datos se vinculen en el ambiente web. En la misma línea de discusión, podemos señalar que el estado de adopción de los datos vinculados en las comunidades culturales aún está en un proceso de maduración.

Si no hay inconveniente, retomaríamos el estudio realizado por Mitchell (2016), donde refiere que la adopción de los datos vinculados considera las aportaciones técnicas en contextos específicos en las comunidades culturales, como sería el caso de la evolución de los estándares.

A propósito de la aceptación de los principios de los datos vinculados, las comunidades culturales, según Mitchell (2016), han encausado sus proyectos hacia la identificación de actividades que permitan posicionar a las comunidades en el contexto de la web y el cada vez mayor uso del Internet de las cosas.

LOS DATOS VINCULADOS EN LAS COMUNIDADES CULTURALES Y SU PROYECCIÓN HACIA LA INTEROPERABILIDAD GLOBAL

Sobre la interoperabilidad global y los datos vinculados en las comunidades culturales, estamos de acuerdo con Zengenene (2013), quien refiere que la innovación tecnológica ha cambiado la naturaleza de la información en su generación, transferencia, manera de compartir y consumo. Este cambio ha forzado a las instituciones culturales a reconsiderar su rol dentro de la nueva sociedad de información. Por esta razón, los datos vinculados como una práctica estandarizada en dichas comunidades se ve solamente como una manera de publicación de datos en la web que deja entrever que las máquinas sean las únicas capaces de otorgar el significado al recurso web y sus relaciones.

Sin embargo, señala Zengenene (2013), los beneficios por adoptar las tecnologías de los datos vinculados son diversos y con altos potenciales para transformar, por ejemplo, los servicios bibliotecarios,

así como los trabajos bibliotecarios, los vendedores de programas, los usuarios y la biblioteca como institución.

A propósito de los beneficios, Tharani (2015) menciona en la investigación que realizó sobre el caso de estudio del BIBFRAME, que el propósito es reemplazar a MARC, pero su futuro en la catalogación es incierto ya que en el paradigma bajo el cual se soportan los datos vinculados, no se tiene plena conciencia respecto a la manera en que serán publicados y compartidos los datos en la web.

Teniendo en cuenta el resultado anterior, nos atreveríamos a mencionar desde la perspectiva filosófica de los datos vinculados, que la base en la que se sustenta BIBFRAME es un modelo y conjunto de especificaciones diseñadas para el control bibliográfico debido a que cada una de las clases contienen atributos demasiado jerárquicos que sirven para la descripción y vinculación de las instancias de éstas clases.

Otros investigadores como Oomen y Aroyo (2011) refieren que la masificación del patrimonio cultural está en gestación y la mayoría de los estudios que surjan pueden ser empíricos en la medida que se descubra el modelo del ciclo de vida del contenido digital y su relación entre los diferentes tipos de actividades básicas para la masificación del patrimonio cultural. Esto es, determinar en cuál de los tres modelos principales se sustenta la investigación, ya sea en los proyectos de contribución, proyectos de colaboración o proyectos de co-creadores.

Por su parte, Bermes (2011) indica que hay una situación compleja entre la interoperabilidad y los datos culturales porque el modelado de datos bibliográficos se sustenta en dos conceptos básicos: la descripción de los documentos y la descripción de las entidades de autoridad que están contenidas en el mismo documento, mientras que el modelo de datos en los archivos enfatiza el concepto de jerarquía y contexto, de ahí que el EAD represente las posibilidades de representar en orden categórico la organización natural de los fondos localizados en el archivo. Finalmente, el modelo de datos en los museos se caracteriza en la descripción única de los objetos, donde estos son descritos tomando en cuenta su contexto de creación como lo hacen las bibliotecas y los archivos, pero se diferencia

en relación a los diferentes eventos que afectan el ciclo de vida del objeto en el museo, por eso es que el modelo de referencia CIDOC CRM centra su idea en la preservación, la restauración, los cambios de posesión y su manera de exhibición.

En suma, las investigaciones enunciadas nos sirven de marco para señalar que la interoperabilidad global y los datos vinculados en las comunidades culturales se encuentren en fase de creación de sus fundamentos, por ello es que proyectos como BIBFRAME, Europeana y ArchivalSpace se observan como un nuevo estilo de interoperabilidad, construido sobre los principios web, que ofrecerán a los usuarios nuevos servicios con adaptación al descubrimiento afortunado o serendipia.

REFERENCIAS BIBLIOGRÁFICAS

- Bermes, E. (2011) Convergence and Interoperability: a Linked Data Perspective. *World Library and Information Congress 77th. IFLA 2011 General Conference and Assambly*. Disponible el 29 de abril de 2016 en <http://conference.ifla.org/ifla77>.
- Caplan, P. (2003) *Metadata Fundamentals for all Librarians*. Chicago: ALA.
- Coyle, K. (2012) Semantic Web and Linked Data. *Library technology reports*, 4 (mayo/junio).
- Godby, C. J., S. Wang y J. K. Mixter. (2015) *Library Linked Data in the Cloud: OCLC's Experiments with new Models of Resource Description*. Chicago: Morgan & Claypool Publisher.
- Jonsson, G. (2002) "Las bases para un registro en la mayoría de las reglas de catalogación y la relación a FRBR". Traducción de Mar Hernández. 68 conferencia IFLA Council and General Conference, Glasgow.

Tendencias multidisciplinarias...

- Mitchell, E. T. (2016) The Current State of Linked data in Libraries, Archives, and Museums. *Library Technology Report*, 52 (2) (enero).
- . (2013) Metadata Development in Libraries and Other Cultural Heritage Institutions. *Library Technology Report*, 5 (julio).
- Oomen, J. y L. Aroyo. (2011) Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges. Disponible el 30 de abril de 2016 en http://www.iisi.de/fileadmin/IISI/upload/2011/p138_oomen.pdf.
- Svenonius, E. (2000). *The Intellectual Foundation of Information Organization*. Cambridge, Mass.: MIT Press.
- Tharani, K. (2015) Linked Data in Libraries: a Case Study of Harvesting and Sharing Bibliographic Metadata with BIBFRAME. *Information Technology and Libraries*, 34 (1) (marzo).
- Zengenene, D. (2013) Global Interoperability and Linked Data in Libraries. *New Library World*, 113, ½: 84-87.

Tecnologia, Metadados e Interoperabilidade: aspectos na modelagem de ambientes informativos digitais

FABIANO FERREIRA DE CASTRO

*Programa de Pós-Graduação em Ciência da Informação
(UFSCar), Brasil.*

INTRODUÇÃO

A preocupação da comunidade biblioteconômica e da Ciência da Informação advinda com as novas regras para a descrição bibliográfica aplicadas ao ambiente digital levam os profissionais a repensarem o seu fazer profissional, na tentativa de caminharem em consonância com as novas transformações trazidas no bojo da recontextualização tecnológica das bibliotecas.

Nas últimas décadas observou-se que mais e mais informações têm sido publicadas, armazenadas e disponibilizadas na *web*. No entanto, somente uma parte dessas informações e em ambientes específicos, tais como as bibliotecas digitais, por exemplo, possuem um padrão de metadados para a descrição bibliográfica e que possibilite a padronização dos recursos informativos para atender aos requisitos da interoperabilidade.

As tecnologias atuais disponíveis baseadas em pesquisas na temática *web* semântica desenvolvem-se com o intuito de tentar

agregar semântica à descrição bibliográfica para um tratamento mais efetivo dos conteúdos digitais. A web semântica oferece algumas soluções para problemas como a padronização das formas de representação da informação criando uma nova visão para o armazenamento e o processamento dos dados. Algumas dessas soluções poderiam ser implementadas para potencializar os resultados das buscas no âmbito das bibliotecas digitais.

No desenvolvimento de um arcabouço teórico e metodológico sobre as ferramentas tecnológicas disponíveis para a construção de formas de representação de recursos informacionais, no âmbito digital, o presente trabalho caracteriza-se por ser uma pesquisa de análise exploratória e descritiva do tema, na tentativa de identificar o embasamento conceitual subjacente às ferramentas para a padronização da descrição dos aspectos de forma e do conteúdo dos recursos bibliográficos, no âmbito digital, na abordagem da Ciência da Informação.

Como pontos de reflexão, destaca-se em que medida os atuais ambientes informacionais digitais precisam ser repensados, a partir da modelagem conceitual e estrutural dos dados, na camada intangível, definida na representação e na descrição dos dados bibliográficos, com alcance à interoperabilidade.

APONTAMENTOS E IMPACTOS TECNOLÓGICOS NO CENÁRIO DA CATALOGAÇÃO DESCRITIVA

Percebe-se que no cenário atual da comunidade da Biblioteconomia e da Ciência da Informação, há uma preocupação e um reconhecimento crescente de que haverá a necessidade de um sucessor do formato bibliográfico MARC 21 (Machine Readable Cataloging), devido às novas transformações no domínio bibliográfico, permeada pelo uso intensivo das TIC (Tecnologias da Informação e Comunicação). Segundo Coyle (2011) tais discussões tendem a se concentrar principalmente nas questões estruturais. Dessa forma, questiona-se se o novo formato para a descrição e a representação dos dados será XML (eXtensible Markup Language), ele fará uso de

RDF (Resource Description Framework) e padrões de dados ligados¹

O que essas questões não se referem é a tarefa muito mais complexa de traduzir a semântica dos dados bibliográficos para a construção de um novo padrão. De acordo com Thomale (2010), basta apenas uma pequena investigação dos dados codificados no formato MARC 21, para revelar que as *tags* e os subcampos em si, são insuficientes para definir os elementos dos dados reais realizados pelos registros de um catálogo bibliográfico. “O primeiro passo na transformação de MARC 21 para um outro formato é identificar quais são os elementos contidos num registro bibliográfico MARC 21, o que não parece tão simples” (Thomale, 2010, p. 3, tradução nossa).

O entusiasmo recém-desenvolvido para RDF se tornar a base para os dados bibliográficos de uma biblioteca tem ocorrido uma série de esforços e aplicações que convertam MARC 21 para RDF, mas nenhuma oficialmente ou recomendada pelos organismos internacionais da Catalogação Descritiva.

Dentre as iniciativas oficialmente aceitas da transformação dos dados em RDF encontram-se na Library of Congress (LC) padrões para a descrição de recursos, tais como o MODS (Metadata Object Description Standard). Outras propostas para a transformação dos dados de bibliotecas para RDF podem ser vistas em: ISBD (International Standard Bibliographic Description) em RDF, FRBR (Functional Requirements of Bibliographic Records) em RDF e RDA (Resource Description and Access) em RDF.

Cada um desses esforços toma uma biblioteca padrão e usa RDF como sua tecnologia subjacente, criando um esquema de metadados completo que define cada elemento do padrão em RDF. O resultado é que se tem uma série de “silos” RDF, e cada elemento

1 O termo dados ligados (*Linked Data*) refere-se ao conjunto das melhores práticas para a publicação e a interligação de dados estruturados na web. Essas melhores práticas foram introduzidas por Tim Berners-Lee em sua nota *Web Architecture* e tornaram-se conhecidas como princípios de dados ligados (Heath e Bizer, 2011, tradução nossa).

de dados definido como se pertencessem exclusivamente a esse padrão.

Existem quatro declarações diferentes, por exemplo, no elemento “lugar de publicação” em ISBD, RDA, FRBR e MODS, cada um com seu próprio URI (Uniform Resource Identifier) e também há diferenças entre elas (por exemplo, RDA separa lugar de publicação, produção etc., enquanto ISBD não o faz), certamente deverão ter uma estrutura (vocabulário) comum para solucionar essas questões. (Coyle, 2012).

Uma possível solução seria se as diferentes instâncias do “lugar de publicação” pudessem ser tratadas como tendo um significado comum, tal que elementos de FRBR pudessem ser ligados a um elemento ISBD, mas o mesmo não ocorre.

A razão pela qual não ocorre se justifica, porque cada uma dessas restringe os elementos de forma única (individual) que define seu relacionamento com um contexto de dados particular (o que geralmente pensamos das estruturas de um registro). Os elementos não são independentes do contexto, e isso significa que cada um somente pode ser usado dentro desse contexto particular. Essa é a antítese do conceito de dados ligados, onde conjuntos de dados de fontes diversas compartilham elementos de metadados. E esta reutiliza os elementos que cria o “link” em dados ligados (vinculados). Para conseguir tal processo, os elementos de metadados precisam ser irrestritos por um contexto particular (Coyle, *ibid.*).

A ligação pode também ser alcançada através dos relacionamentos verticais, similares aos termos (mais amplos e mais restritos) num tesauro. Essa opção é menos direta, mas torna possível a “mistura” de conjuntos de dados que têm diferentes níveis de granularidade. No caso do “lugar de publicação” da ISBD, poderiam ser definidos mais amplamente para os três elementos de RDA que tratam esse separadamente. Coyle (*ibid.*) afirma que infelizmente isso não é possível devido à forma que ISBD e RDA foram definidos em RDF.

Coyle (*ibid.*) reforça que o resultado é que agora tem-se uma série de “silos” RDF, expressões dos dados em RDF que faltam a capacidade de cruzamento dos dados ligados, porque eles são obrigados

a especificar estruturas de dados e pouco se ganha em termos de dados ligados sob a óptica bibliográfica. Não somente os esquemas RDF são incompatíveis com outros, ninguém será ligado a dados bibliográficos de comunidades fora de bibliotecas que publicarem seus dados na web. O que significa a não ligação dos dados à Amazon, à Wikipedia, para citações em documentos.

Coyle (*ibid.*) diz que devido ao estágio inicial em que estamos no desenvolvimento de dados ligados para o ambiente das bibliotecas, têm-se duas opções em relação ao uso do RDF.

- Definir “super-elementos” que estejam acima dos registros de formatos e que não estão vinculados pelas restrições dos registros definidos em RDF. Neste caso, haveria um “lugar de publicação” geral que seria um “super” elemento correspondente a todos os lugares de publicação nos vários registros e seriam subordinados a um conceito geral de “lugar” utilizado amplamente. Para implementar a ligação, cada elemento de registro seria extrapolado para os seus “super-elementos”.
- Definir primeiro os elementos de dados fora de qualquer formato de registro específico, e então usá-los em esquemas de registros. Neste caso haveria apenas uma instância do “lugar de publicação” e que seria utilizado nos vários registros bibliográficos, sempre que um elemento é necessário. O intercâmbio desses registros seria possível com a ligação dos dados de seus elementos componentes, e iria interagir com outros dados bibliográficos na *Web* utilizando os elementos RDF definidos e seus relacionamentos.

Coyle (*ibid.*) conclui ainda que precisa-se criar dados, não somente registros, e que necessita-se criar primeiramente os dados, em seguida, registros com os mesmos para aplicações de acordo com a necessidade de cada ambiente informacional. Esses registros irão operar internamente nos sistemas de bibliotecas, enquanto os dados têm o potencial para fazer conexões no espaço de dados ligados. Um esforço maior tem que acontecer em descobrir e definir

os elementos de nossos dados e olhar para os diversos dados que deseja unir, no vasto universo de informações.

As bibliotecas têm seu foco nos registros bibliográficos que compõem os acervos institucionais, geralmente um documento complexo que atua como um substituto do catálogo, tal como um livro ou uma gravação de música. RDF, nesse contexto, não menciona nada sobre os registros, apenas diz que existem dados que representam coisas (recursos) e os relacionamentos entre essas coisas. O que é muitas vezes confundido é que qualquer coisa pode ser um recurso em RDF, assim, o livro, o autor, a página, a palavra na página, qualquer uma, ou todas essas poderiam ser as coisas (ou recursos) em seu universo. (Coyle, *ibid.*).

As discussões acerca do futuro das bibliotecas digitais e sua configuração para a adoção de tecnologias da web semântica manifestam-se da necessidade que os ambientes e sistemas de informação criem uma estrutura de dados que aproveitem as potencialidades do RDF, permitam os relacionamentos de forma explicitada e promovam a interoperabilidade.

Vale destacar que os relacionamentos bibliográficos sempre existiram no interior dos catálogos bibliográficos entre os dados de um registro bibliográfico, por meio das regras e dos esquemas de descrição oriundos da catalogação, porém não explicitados aos sujeitos institucionais para o uso e (re) uso das informações nos ambientes de manipulação de dados bibliográficos e catalográficos.

DA HETEROGENEIDADE DOS DADOS AOS METADADOS: CONSTRUCTOS ACERCA DA INTEROPERABILIDADE

A Internet está promovendo o acesso a um número amplo, diversificado e crescente de sistemas de informação digital. Nesse contexto, as bibliotecas digitais concebidas originalmente como sistemas locais de recuperação da informação para um grupo especializado de usuários, graças à universalização da *web* tem permitido o acesso aos recursos informacionais de forma globalizada.

A construção de bibliotecas digitais implica a integração de sistemas complexos que incluem coleções de documentos com estruturas diferentes, distintas mídias e conteúdos heterogêneos. Atrelado a isso, tem-se que considerar uma variedade de componentes de *hardware* e *software*, que almejam a interoperabilidade através de diferentes estruturas de dados, algoritmos para o processamento, e ainda múltiplos sujeitos, comunidades e instituições que possuem objetivos, políticas e culturas diferenciadas.

Segundo Méndez Rodríguez (2002) a noção fundamental e o conceito-chave para o funcionamento pleno da recuperação da informação em ambientes informacionais digitais estão baseados no uso efetivo dos metadados, que estabelecerá a interoperabilidade. Essa interoperabilidade requer conhecer todo o sistema de informação sob vários pontos de vista: a sintaxe (a combinação das palavras ou gramáticas para comunicar semântica e estrutura); a estrutura (que pode ser lida por humanos e pode validar-se através de máquinas); o vocabulário (que seria a correspondência entre os símbolos e os conceitos); e a semântica (o significado dos elementos ou do vocabulário).

Muitas discussões das investigações em temas que envolvem a recuperação da informação, sobretudo, em bibliotecas digitais distribuídas em rede recaem sobre a infraestrutura, em estabelecer ferramentas comuns, tecnologias e padrões compartilhados que possam facilitar a base para diversos aspectos da interoperabilidade.

Um objetivo comum desses esforços é permitir que diferentes comunidades, com diferentes tipos de informação e usando diferentes tecnologias, consigam um nível geral de compartilhamento de informação e, por meio de processos de agregação apoiados por tecnologia da informação, criem novos e mais poderosos tipos de serviços de informação. (Sayão e Marcondes, 2008). De acordo com Méndez Rodríguez (2002, p. 312, tradução nossa),

A interoperabilidade supõe complementar uma série de estratégias humanas ligadas à descrição consistente e flexível da informação mediante metadados, com um conjunto de ferramentas, *softwares* e protocolos que permitam os distintos sistemas de informação

(bibliotecas digitais) estabelecer mecanismos de conversão automática entre eles, e ainda, reconciliar suas diferenças e alcançar, ao menos, um acesso universal integrado aos recursos informacionais.

A interoperabilidade, entretanto, no campo científico das áreas de Ciência da Informação e da Biblioteconomia não é algo novo, principalmente com o fenômeno da denominada explosão da informação, em meados do século XX, onde as unidades de informação, em especial as bibliotecas, sempre trabalharam numa filosofia de cooperação, colaboração e compartilhamento de recursos informacionais, criando uma rede estruturada e ordenada de informações para o acesso e a recuperação de documentos. Tais características da interoperabilidade podem ser visualizadas intrinsecamente na construção dos primeiros catálogos bibliográficos.

Na área de tecnologias da informação a interoperabilidade é vista como a capacidade de comunicação entre programas de computadores; entretanto, no campo das bibliotecas digitais o conceito de interoperabilidade é mais abstruso e estratificado e com muitas visões interdisciplinares subjacentes (Sayão e Marcondes, 2008).

A interoperabilidade pode ser considerada como o processo contínuo de assegurar que sistemas, procedimentos e a cultura de uma organização sejam gerenciados de tal forma que possibilitem a maximização das oportunidades para intercâmbio e reuso de informação. (Guy, 2005; Riley, 2010).

Vale destacar que a interoperabilidade pode ser realizada em diferentes níveis, e possui muitas facetas, conforme pode ser observado em Guy (2005), Sayão e Marcondes (2008):

- Interoperabilidade técnica – as considerações sobre os aspectos técnicos incluem assegurar envolvimento de um conjunto de organizações no contínuo desenvolvimento de padrões de comunicação, transporte, armazenamento e representação de informações, tais como são o Z39.50 e o *Search Retrieval Web Service* (SRW). Inclui também os esforços cooperativos para assegurar que padrões individuais evoluam em benefício da comunidade envolvida e para facilitar, onde

for possível, convergência desses padrões, de forma que seja possível que os sistemas possam ter como base mais de um conjunto de padrões.

- Interoperabilidade semântica – está relacionada com o significado ou semântica das informações originadas de diferentes recursos e é solucionada pela adoção de ferramentas comuns ou/e mapeáveis de representação da informação, como esquemas de metadados, classificações, tesouros e mais recentemente, ontologias; um exemplo de questão endereçada por essa faceta da interoperabilidade pode ser o seguinte: o que significa “autor” para um recurso informacional? Será a mesma coisa que “criador” para um outro recurso? No nível semântico Marino (2001) acrescenta que existem dois subníveis: Epistemológico (relacionado ao significado dos elementos descritores do formato e das relações nele existentes) e Ontológico (relacionado ao uso de ontologias, vocabulários controlados e padrões de metadados para o estabelecimento dos significados dos dados representados). Como exemplo desse nível de interoperabilidade pode-se destacar os metadados que têm o mesmo significado entre elementos de descrição do formato MARC 21.
- Interoperabilidade política/humana – independente das questões relacionadas à maneira pela qual a informação é descrita e disseminada, a decisão de tornar os recursos informacionais mais amplamente disponíveis e interoperáveis tem implicações para a organização, para as equipes envolvidas e para os usuários em termos comportamentais, de recursos e de treinamento. A ênfase dada por parte de alguns setores governamentais aos problemas de democratização do acesso, da exclusão digital e da federação de fontes de informação voltadas para a educação a distância, tem impacto nas políticas públicas para a área, e estão enquadrados nesse item.
- Interoperabilidade intercomunitária – enfoca a necessidade, cada vez mais urgente, impulsionada pela crescente interdisciplinaridade, principalmente nas áreas de pesquisa, de

acesso a informações provenientes de um espectro amplo de fontes distribuídas por organizações, áreas de conhecimento e comunidades de natureza distintas. Geralmente exige o estabelecimento de fóruns para discussão e consenso em torno de práticas e procedimentos comuns.

- Interoperabilidade legal – considera as exigências e as implicações legais de tornar livremente disponíveis itens de informação.
- Interoperabilidade internacional – quando se atua em escala internacional é necessário contornar a diversidade de padrões e normas, os problemas de comunicação, as barreiras linguísticas, as diferenças no estilo de comunicação e na falta de uma fundamentação comum.

Essa investigação tem como interesse os requisitos de interoperabilidade nas esferas semântica, estrutural e sintática, onde defende-se que a estruturação e a modelagem de catálogos bibliográficos (banco de dados) ocorrem nessas camadas, para que a interoperabilidade entre ambientes informacionais digitais heterogêneos ocorra de maneira efetiva e se estabeleçam os relacionamentos bibliográficos semânticos numa mesma interface.

Méndez Rodríguez (2002) elucida que a interoperabilidade sintática (sintaxe) diz respeito à forma como se codificam os metadados digitalmente. Um exemplo dessa natureza é o RDF, que codifica os tipos de propriedades e os valores desses elementos. A estrutura descreve os mecanismos para agrupar os elementos de metadados, assim como as regras sobre como devem ser codificados os valores para cada elemento.

Metadados podem ser considerados instrumentos cada vez mais necessários no ambiente da *Web* atualmente, permitindo em larga escala a gestão de recursos informacionais distribuídos. Nos últimos anos tem-se assistido a um crescimento na interação de comunidades de metadados, que são cada vez mais impulsionadas por uma necessidade de colaboração e intercâmbio entre domínios heterogêneos.

No entanto, os padrões de metadados isoladamente não têm sido capazes de atender as necessidades de interoperabilidade entre as comunidades, instituições e organizações independentes responsáveis pelo estabelecimento da normalização e padronização de metadados.

Como podemos definir conceitualmente interoperabilidade no contexto dos metadados? O IEEE (2000) diz que a interoperabilidade corresponde à habilidade de dois ou mais sistemas ou componentes trocar informações e utilizá-las neste processo. Nilsson (2010) diz que a interoperabilidade de metadados consiste na capacidade de dois ou mais sistemas, ambientes informacionais ou componentes trocarem dados descritivos sobre recursos informacionais, e interpretá-los de uma maneira consistente, não ocorrendo falhas no intercâmbio para sua compreensão por máquinas.

De acordo com Sayão e Marcondes (2008, p. 136),

Uma organização verdadeiramente interoperável é capaz de maximizar o valor e o potencial de reuso da informação que está sob o seu controle. É também capaz de intercambiar efetivamente estas informações com outras organizações igualmente interoperáveis, permitindo que novos conhecimentos possam ser gerados a partir da identificação de relacionamentos entre conjuntos de dados previamente não relacionados. Na perspectiva do usuário, as interfaces devem apresentar para o usuário uma visão unificada em termos semânticos de diferentes recursos informacionais heterogêneos, ou seja: como nomeá-los, como referenciá-los, como utilizá-los em buscas, como acessá-los, como apresentá-los para o usuário.

De acordo com Castro (2012), pode-se dizer que a interoperabilidade é a capacidade de dois ou mais sistemas, ambientes informacionais digitais e padrões de metadados heterogêneos, intercambiarem seus dados, a partir da codificação das regras e esquemas de descrição utilizadas pelo catalogador, na confecção padronizada e metodológica dos metadados, e na estruturação conceitual do ambiente informacional proporcionados pelas ontologias. Nesse sentido, Alves (2005, p. 110) diz que as ontologias funcionam como

“qualificadores dos metadados e não haverá estabelecimento da semântica sem a representação por metadados”.

Os metadados são introduzidos nos sistemas de gerenciamento de informações, com o intuito de apoiar determinadas operações administrativas, incluindo a busca, apresentando resumos ou configurando interfaces. Na sua essência, os metadados criam um nível de compreensão, permitindo que os sistemas gerenciem recursos sem ter que mergulhar em seu desenvolvimento físico ou digital internos (Nilsson, 2010).

Os metadados podem consistir de todo tipo de informação sobre um item, por exemplo, seu título, descrições textuais e classificações de assunto, para o acesso e as relações contextuais entre o item a ser descrito e outras coisas. Nesse momento é oportuno resgatar os conceitos intitulados pelos Requisitos Funcionais para Registros Bibliográficos (FRBR), onde um item pode ser entendido como um único exemplar de uma manifestação (IFLA, 1998), ou ainda, segundo as regras de catalogação do AACR2 (2005), um documento ou um conjunto de documentos sob qualquer forma física, publicado, distribuído ou tratado como uma entidade autônoma, constituindo a base de uma descrição bibliográfica única.

O núcleo de proposição dos metadados é que seu uso efetivo permite aos sistemas, às aplicações e aos usuários gerenciar e acessar itens sem nenhuma necessidade de interação direta com o item em si (Lytras e Sicilia, 2007). Por essa razão, a administração e o intercâmbio de metadados é a atividade central em muitos sistemas que gerenciam os objetos digitais e não digitais, tais como os sistemas de gerenciamento de conteúdo, repositórios de objetos de aprendizagem e bibliotecas (digitais ou convencionais).

Nilsson (2010) acrescenta que especificações e padrões de metadados acrescentam valores adicionais ao reduzir o limiar para o desenvolvimento de sistemas que realizam intercâmbio, reusam e combinam metadados de diferentes fontes. Um padrão comum garante uma melhor documentação e um melhor acesso às ferramentas reutilizáveis. Este é o valor central para a interoperabilidade de metadados. Percebendo o potencial inerente ao uso de metadados interoperáveis informado requer uma coordenação em grande

escala entre os atores relevantes em um campo de prática e aplicação. Especificações de metadados tendem a ser concebidos para uma determinada comunidade particular, com mais ou menos itens bem definidos e descritos e cenários comuns de uso.

Neste cenário, pode-se destacar no domínio bibliográfico, metadados na forma de catalogação tem sido a questão desde as primeiras bibliotecas. Como os dados de uma biblioteca geralmente estão abertos e disponíveis ao mundo, os principais problemas na interoperabilidade de metadados começam a surgir. O desenvolvimento de um novo código de catalogação Resource Description and Access (RDA) aparece nesse contexto, como uma forma de auxiliar a interoperabilidade de metadados das bibliotecas, destacando a situação complexa com uma multiplicidade de padrões de metadados utilizados em bibliotecas do mundo todo, como por exemplo, o MARC 21.

Tanto as bibliotecas como as ferramentas tecnológicas que permeiam seu ambiente informacional atingem o campo dos metadados orientados a web, onde o Resource Description Framework (RDF) tem feito progresso ao longo das últimas décadas, juntamente com uma série de especificações que sustentam a web semântica, tal como a Web Ontology Language (OWL) (World Wide Web Consortium, 2009).

Quando os metadados projetados de acordo com diferentes especificações e de diferentes domínios se encontram, ou seja, quando comunidades se envolvem para aumentar sua interação, dificuldades consideráveis no gerenciamento dos metadados tendem a aparecer, pois muitas vezes os dados se apresentam incompatíveis. O resultado é que os benefícios da interoperabilidade de metadados dentro de um padrão são perdidos quando são combinados, aumentam os custos de desenvolvimento e falhas na comunicação entre os sistemas, aplicando-se em muitos casos soluções não reutilizáveis.

Godby, Smith e Childress (2003) argumentam baseados em experimentos com o cruzamento de metadados (*crosswalks*), que as traduções completas de um padrão para outro são possíveis numa comunidade de prática específica, enquanto somente traduções

parciais são possíveis entre elas. Os autores exemplificam os metadados de uma biblioteca que podem ser combinados com outras especificações de bibliotecas diferentes, desde que usando um padrão comum para a descrição dos recursos bibliográficos.

Nesse momento pode-se corroborar com as autoras Santos e Corrêa (2009) ao mencionarem que nos ambientes de manipulação de dados bibliográficos, em especial, os ambientes informacionais digitais, quem define a interoperabilidade é o padrão de metadados a ser adotado para representar e descrever os recursos informacionais.

Subjacentes aos formatos e aos modelos abstratos é a noção sutil de semântica. Com a ascensão do RDF e a iniciativa da web semântica pelo W3C, a descrição semântica de metadados tem recebido especial atenção, no desenvolvimento e na construção de ambientes informacionais digitais interoperáveis.

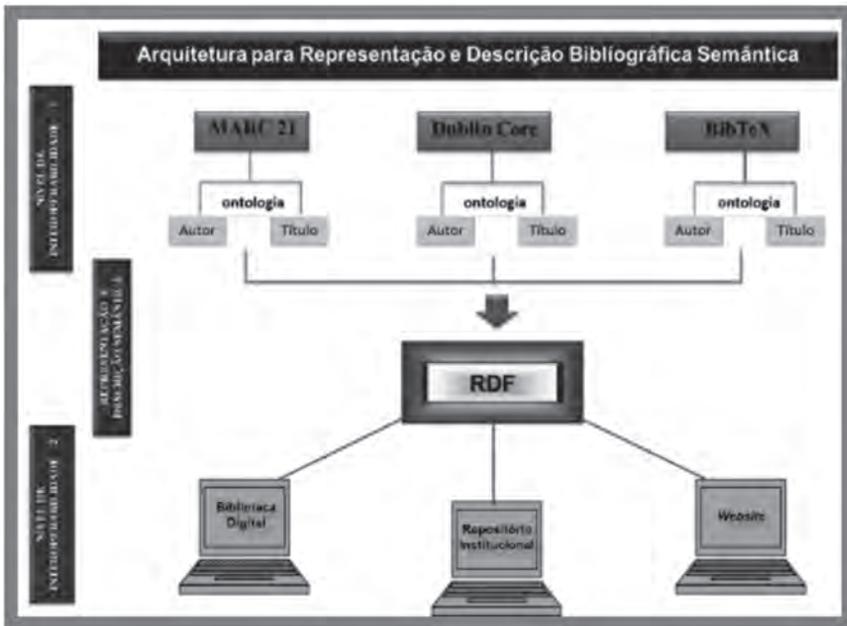
PROPOSTA DE MODELAGEM ESTRUTURAL DE DADOS NOS DOMÍNIOS BIBLIOGRÁFICO E WEB

De acordo com os apontamentos e reflexões acerca das tecnologias digitais, e com respaldo na literatura científica, verifica-se que a palavra de ordem na constituição de ambientes informacionais digitais, sobretudo, no âmbito das bibliotecas digitais é a interoperabilidade. Para tanto, para que a interoperabilidade aconteça de maneira eficiente, faz-se necessário um olhar mais acurado nas camadas estruturantes e delimitadoras dos catálogos bibliográficos digitais, ou seja, na representação e na descrição dos dados, a fim de potencializar as formas de busca e recuperação das informações.

Castro (2012), com base na literatura científica apresentada sobre a representação, a descrição, a modelagem e a interoperabilidade de recursos bibliográficos e pautados nas principais definições, arquiteturas e tecnologias, e como resultados dessa investigação, apresenta uma proposta de um modelo/esquema para a representação e a descrição bibliográfica semântica de um recurso informacional, contempladas num único ambiente, com vistas a

possibilitar a interoperabilidade semântica entre padrões de metadados e ambientes informacionais heterogêneos, de acordo com a *Figura 1*.

Figura 1. Modelo para a Representação e a Descrição Bibliográfica semântica e níveis de interoperabilidade.



Fonte: Castro (2012, p. 178).

A arquitetura para a Representação e a Descrição Bibliográfica Semântica vislumbra a possibilidade da construção de representação e de descrição bibliográfica semântica e a interoperabilidade semântica entre, não somente padrões de metadados, mas também entre ambientes e sistemas informacionais digitais heterogêneos.

Na parte superior da arquitetura, têm-se alguns padrões de metadados (MARC21, Dublin Core e BibTeX), descrevendo alguns

metadados em comum, tais como, autor e título (considerados elementos de calção de um recurso bibliográfico). Na fase da definição dos metadados precisa-se considerar para estruturar os elementos as ontologias, ou seja, o delineamento conceitual de cada atributo e os modelos conceituais de dados, tais como os FRBR e FRAD.

No padrão de metadados e/ou formato de intercâmbio MARC 21, por exemplo, as ontologias são determinadas através das regras e esquemas de descrição bibliográfica (AACR2) e também, do conceito fornecido pelo próprio MARC 21.

A definição dos elementos de metadados pela ontologia de descrição bibliográfica é considerada uma das principais partes da interoperabilidade por poder fornecer informações significativas para que sejam compreensíveis por agentes não humanos (computador/máquina).

Percebem-se conceitos semânticos na ontologia de descrição bibliográfica AACR2 e MARC 21, pautados na lógica de descrição estabelecidos no processo da Catalogação Descritiva, uma vez que a base de conhecimento do formato MARC 21 é o AACR2. No RDA, por exemplo, não é explicitado o conceito Autor, o mesmo é chamado por *Creator* (Criador), linguagem que está em consonância com o padrão de metadados Dublin Core, desenhado para a localização de recursos na web.

Vale destacar nesse momento que a arquitetura supracitada apresenta ontologias para apenas dois elementos da descrição bibliográfica na estrutura MARC 21 (autor e título), uma vez que é necessário o desenvolvimento e a criação de ontologias para determinar a semântica de todos os metadados do padrão MARC 21, pois quanto mais complexa e completa for a estruturação, haverá uma compreensão mais intuitiva e efetiva pela máquina promovendo a interoperabilidade.

Ainda na fase de confecção dos metadados, a atribuição das ontologias para a descrição bibliográfica não é condição suficiente para determinar a interoperabilidade. O catalogador além de conhecer a estrutura MARC 21 deverá também dominar e refletir o uso das regras de catalogação (AACR2) para saber aplic-las no preenchimento do conteúdo dos metadados (valores dos elementos).

Para a descrição bibliográfica de um recurso bibliográfico (livro, por exemplo) na estrutura MARC 21, o catalogador deverá seguir as regras do capítulo 2 do código de catalogação (AACR2) que norteiam a construção padronizada e unívoca de como deverão ser representados e apresentados os dados aos usuários.

É através do uso adequado das regras e dos esquemas que as representações terão maior consistência na camada intangível do sistema propiciando formas de apresentação dos recursos bibliográficos de forma mais efetiva.

A meta no nível de interoperabilidade 1, de acordo com a figura 1, é a possibilidade de integração e comunicação entre padrões de metadados distintos, para tanto só será possível a partir da determinação de ontologias de descrição bibliográfica explicitadas nas regras e códigos de catalogação e em formatos de metadados potencializando os relacionamentos bibliográficos semânticos e estabelecendo a interoperabilidade.

A camada RDF terá o objetivo de fazer o armazenamento e a representação dos dados, localizada na parte central da arquitetura, com as descrições bibliográficas semânticas. Independente da estrutura original do padrão permitirá que os dados em MARC 21, Dublin Core, BibTeX ou qualquer outro padrão, seja convertido em RDF e podendo ser reaproveitados na construção web semântica.

Já na parte inferior da arquitetura é possível verificar a presença de ambientes informacionais heterogêneos, que podem ser interpretados tais como, Biblioteca Digital, Repositório Institucional, Website etc., definidos no modelo, sendo que a representação e a descrição dos recursos informacionais, podem estar ou não contemplados num mesmo padrão. O nível de interoperabilidade 2 objetiva a possibilidade de interoperabilidade semântica entre múltiplos ambientes e sistemas informacionais digitais e o armazenamento das descrições bibliográficas semânticas numa única plataforma/interface.

Ressalta-se, que quanto mais completa e complexa for uma descrição, melhores possibilidades nos níveis de interoperabilidade semântica haverá, com a possibilidade de uso, de preservação e de (re)

uso das informações de forma mais efetiva pelo usuário final, multidimensionando as formas de acesso aos recursos informacionais.

Portanto, acredita-se que tal modelo de arquitetura proposto nessa investigação, pode possibilitar o compartilhamento entre padrões de metadados e ambientes e sistemas informacionais distintos, trabalhando numa filosofia de colaboração entre os recursos informacionais disponíveis e as tecnologias que estão abarcadas na sua construção, no estabelecimento da interoperabilidade, na otimização dos relacionamentos bibliográficos e ampliados para a construção padronizada de recursos na *web*.

À GUIA DE CONSIDERAÇÕES E REFLEXÕES

A utilização de ontologias, de metadados e dos fundamentos teóricos e metodológicos da Catalogação Descritiva uma vez articuladas, podem redesenhar novas opções aos ambientes informacionais, seja na modelagem do banco de dados (catálogos), ou ainda, na forma de representação dos recursos informacionais, garantindo a possibilidade de interoperabilidade semântica e no auxílio às formas de apresentação dos recursos, multidimensionando as formas de busca aos usuários, de acesso e de recuperação de informações pertinentes, significativas e relevantes, bem como o uso, a preservação e o (re) uso numa única interface. (Castro, 2016, tradução nossa).

Quanto ao “futuro” do MARC 21, sabe-se que há iniciativas internacionais de construção de conversores, ferramentas e instrumentos tecnológicos que podem potencializar a transformação dos dados bibliográficos, numa estrutura RDF, mas ainda nenhuma oficializada pelos organismos internacionais de catalogação. Essa investigação não tem a pretensão de desenvolver o que poderia ser chamado de MARC RDF, mas refletir sobre um caminho inicial com orientações teóricas e metodológicas, para a determinação e a construção de ambientes informacionais melhor estruturados, com descrições bibliográficas semânticas explicitadas, propiciando a interoperabilidade em ambientes informacionais digitais.

As discussões acerca do futuro das bibliotecas digitais e sua configuração para a adoção de tecnologias da *Web Semântica* manifestam-se da necessidade que os ambientes e sistemas de informação criem uma estrutura de dados, que aproveitem as potencialidades do RDF, permitam os relacionamentos de forma explicitada e promovam a interoperabilidade.

Para Castro (2012), o formato de metadados MARC 21 fornece o protocolo pelo qual os computadores realizam o intercâmbio, o uso e a interpretação da informação bibliográfica, ligando elementos de dados para formar a base da maioria dos catálogos de bibliotecas utilizada hoje. A *Web Semântica* em muitos aspectos será baseada na utilização desse tipo de ligação de dados, mas talvez em um catálogo ou banco de dados muito maior e globalizado.

O modelo para a representação e a descrição bibliográfica semântica e os níveis de interoperabilidade desenvolvido nesse estudo, propicia a modelagem estrutural dos ambientes informacionais digitais atualmente, a partir da heterogeneidade de esquemas de metadados, tornando-se, dessa forma, uma estrutura única a ser adotada por vários catalogadores, ampliando seu escopo para a construção padronizada de recursos na *web*. Além disso, a adoção desse modelo configura o estabelecimento efetivo da interoperabilidade, principalmente com os impactos tecnológicos na Catalogação Descritiva e o desenvolvimento de ambientes informacionais digitais interoperáveis.

REFERÊNCIAS BIBLIOGRÁFICAS

- AACR2. (2005). *Anglo American Cataloging Rules*. Preparado sob a direção de The Joint Steering Committee for Revision of AACR. Trad. Federação Brasileira de Associações de Bibliotecários, Cientistas da Informação e Instituições (FEBAB). São Paulo: FEBAB, 2005.
- Alves, R. C. V. (2005). Web Semântica: uma análise focada no uso de metadados. 2005. 180 f. Dissertação (Mestrado em Ciência da Informação). Faculdade de Filosofia Ciências, Universidade Estadual Paulista, Marília.
- Castro, F. F. de. (2016). Functional requirements for Bibliographic Description in Digital Environments. *TransInformação*, 28: 223-231.
- . (2012). Elementos de interoperabilidade na catalogação descritiva: configurações contemporâneas para a modelagem de ambientes informacionais digitais. Tese (Doutorado em Ciência da Informação). Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.
- Coyle, K. (2012). *Bibliographic Framework: RDF and linked data*. Berkeley, United States. Disponível em: <http://kcoyle.blogspot.com/2012/01/bibliographic-framework-rdf-and-linked.html>. Acesso em: 20 jan. 2016.
- . (2011). MARC 21 as Data: a Start. *Code4Lib Journal*, 14.
- Godby, C. J., D. Smith e E. Childress. (2003). Two Paths to Interoperable Metadata. *Proceedings of DC-2003: Supporting Communities of Discourse and Practice. Metadata Research & Application*. Seattle, Sep./Oct. Disponível em: http://www.siderean.com/dc2003/103_paper-22.pdf. Acesso em: 20 jan. 2016.

- Guy, M. (2005). Interoperability Focus: Looking at Interoperability. Ukoln. Disponível em: <http://www.ukoln.ac.uk/interop-focus/about/leaflet.html>. Acesso em: 20 mar. 2016.
- Heath, T. e C. Bizer. (2011). *Linked Data: Evolving the Web into a Global Data Space. Morgan & Claypool, 1* (1). (Synthesis Lectures on the Semantic Web: Theory and Technology). Disponível em: < <http://linkeddatabook.com/editions/1.0/#linkedData>>. Acesso em: 15 set. 2016.
- IEEE. (2000). Standards Information Network. *The Authoritative Dictionary of IEEE Standards Terms*. New York: IEEE.
- IFLA Study Group on the Functional Requirements for Bibliographic Records. (1998) *Functional Requirements for Bibliographic Records: final report*. UBCIM Publications - New Series, vol. 19. München: K. G. Saur. Disponível em: <<http://www.ifla.org/VII/s13/frbr/frbr.pdf>>. Acesso em: 18 jan. 2016.
- Lytras, M. D. e M.A. Sicilia. (2007). Where is the Value of Metadata? *International Journal of Metadata, Semantics and Ontologies*, 2 (4): 235-241.
- Méndez Rodríguez, E. (2002). *Metadados y recuperación de información: estándares, problemas y aplicabilidad en bibliotecas digitales*. Gijón: Trea, 2002.
- Nilsson, M. (2010). From Interoperability to Harmonization in Metadata Dstandardization: Designing an Evolvable Framework for Metadata Harmonization. Thesis (Doctoral in Computer Science and Communication). KTH School of Computer Science and Communication, Stockholm.
- Riley, J. (2010). *RDF for librarians*. DLP Brown Bag Series, p. 1-38, sep. 2010. Disponível em: <http://www.dlib.indiana.edu/education/brownbags/fall2010/rdf/rdf.pdf>. Acesso em: 15 out. 2016.

Tendências multidisciplinares...

- Santos, P. L. e M. R. Corrêa. (2009). *Catálogo: trajetória para um código internacional*. Niterói: Intertexto.
- Sayão, L. F. e C. H. Marcondes. (2008). O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. *TransInformação*, 20, (2): 133-148.
- Thomale, J. (2010). Interpreting MARC: Where's the bibliographic data? *Code4Lib Journal*, 11, 2010.
- World Wide Web Consortium (W3C). (2009). *OWL 2 Web Ontology Language, Structural Specification and Functional-Style Syntax*, W3C Recommendation. Disponível em: <<http://www.w3.org/TR/owl-syntax/>>. Acesso em: 20 jan. 2016.

**GOBIERNO ABIERTO
Y CIUDADANÍA**

La participación ciudadana en el contexto del gobierno abierto

ALEJANDRO RAMOS CHÁVEZ
Universidad Nacional Autónoma de México

INTRODUCCIÓN

En los últimos años, han existido tendencias por parte de muchos países que han tenido por objetivo generar procesos más transparentes, plurales y participativos en las formas de llevar a cabo los procesos gubernamentales. En este sentido, iniciativas como la Alianza Internacional para el Gobierno Abierto (AGA) o la Alianza para el Gobierno Abierto México, apuntan a alcanzar mejores procesos de rendición de cuentas y generar canales más plurales de información y comunicación, así como de participación de los ciudadanos y los gobiernos.

La calidad de la información que el gobierno pone a disposición de la ciudadanía resulta un punto fundamental en este análisis, ya sea que esta información esté fundamentada en la transparencia de las acciones gubernamentales, o bien en mecanismos que permitan la interacción permanente entre la sociedad y su gobierno. En este punto, el correcto desarrollo de procesos y de generación de marcadores y metadatos para obtener la información de forma sencilla y oportuna es vital para el correcto desempeño del gobierno abierto.

DE LAS FORMAS TRADICIONALES DE GOBIERNO AL GOBIERNO ABIERTO

Existen muchas formas en las que los gobiernos atienden los asuntos públicos y se relacionan con sus ciudadanos; en este sentido, dependiendo del grado de apertura que adopte el gobierno, pero también de la acción colectiva llevada a cabo por la ciudadanía, se abren o cierran las posibilidades de incidencia más o menos activa por parte de los ciudadanos en la forma en la que son tratados los asuntos públicos y son puestas en marcha las políticas. Algunas de estas formas son, en diferentes grados, las adoptadas por la mayoría de los gobiernos en la actualidad para llevar a cabo las acciones gubernamentales. De igual manera, sus características pueden ser encontradas en las formas de gobierno del caso específico de México.

De forma puntual, nos referimos a la gobernabilidad (Alcántara-Sáez, 1995; Altman, 2001; Camou, 1995; Pérez Linan, 2003) y a la gobernanza (Aguilar-Villanueva, 2007; Brugue, Goma y Subirats, 2005; Kaufmann, Kraay y Zoido-Lobaton, 2000; Kenis y Schneider, 1991; Mayntz, 2000, 2002) que, dependiendo de las acciones y los matices adoptados por los gobiernos, pueden dar como resultado otra diversidad de formas en las que se insertaría, por ejemplo, el de la señalada por muchos autores como gobernabilidad democrática (Altman, 2001; Camou, 1995; Gentes, 2009; Prats, 2001). Cabe señalar que la forma que adopte el gobierno puede visualizarse de forma piramidal de acuerdo al grado de apertura en el involucramiento de los ciudadanos en el quehacer gubernamental.

En este sentido, la gobernabilidad desde una perspectiva conceptual ha sido definida como una forma de gobierno en la que tanto los procesos como el ejercicio de gobierno están direccionados a alcanzar tres objetivos:

- 1) Generar elementos para disminuir las incertidumbres de los mercados.
- 2) Reducir los riesgos en las formas en las que son tomadas en consideración las decisiones gubernamentales.

- 3) Alcanzar formas de gobierno que permitan obtener aceptabilidad y estabilidad por parte de la población con relación a la gestión pública, lo que podría facilitar la permanencia del partido político en el gobierno al finalizar el periodo gubernamental (Torres y Ramos, 2008).

Uno de los primeros usos que se le dio al concepto de gobernabilidad apuntaba a la necesidad de alcanzar un equilibrio entre la carga de exigencias sociales y las acciones gubernamentales para cumplirlas; es decir, a la capacidad del gobierno para atender las problemáticas de la sociedad al incluirlas, en una primera instancia, en las agendas gubernamentales. De igual forma, el término surgió en el periodo de crisis en los que se vieron inmersos muchos países en la década de los setenta del siglo pasado. Sin embargo, a últimos tiempos, la gobernabilidad ha sido más utilizada por su contraparte negativa, es decir, la ingobernabilidad, para dar cuenta de la sobre carga financiera, política y de demandas sociales que el gobierno no puede atender de forma eficiente (Aguilar-Villanueva, 2007; Bartolomé, 2006; Camou, 1995).

Al realizar un ejercicio de esquematización de estas formas de gobierno, se podría ver en el caso de la política tradicional bajo la forma de gobernabilidad, un proceso en donde existe una agenda política; después se llevaría a cabo la decisión de las formas y los elementos que abordará la política, para posteriormente hacer una formulación, su implementación y evaluación. Hablando en términos de la teoría de sistemas, en este tipo de gestión de políticas existiría una caja negra en donde entran demandas y se llevan a cabo los procesos para elaborar políticas donde sólo tienen competencia los agentes gubernamentales; finalmente, habría una salida de políticas y, en el mejor de los casos, de solución de problemáticas sociales (véase *Figura 1*).

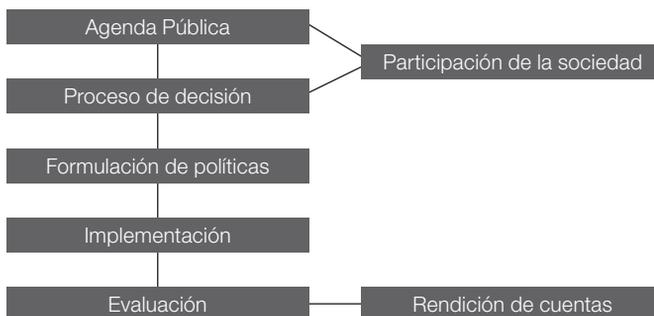
Figura 1. Esquematación del modelo de gobernabilidad.



Fuente: Elaboración propia.

Para el caso de la política en su forma de gobernabilidad democrática, se habla de políticas públicas, pues existe una participación de la sociedad en el establecimiento de la “agenda pública” y en los procesos de decisión de la urgencia o prioridad en el establecimiento de políticas para atender los asuntos públicos. Además, se incorpora en la fase de evaluación un proceso de rendición de cuentas del gobierno a la sociedad (véase *Figura 2*).

Figura 2. Esquematación del modelo de gobernabilidad democrática.



Fuente: Elaboración propia.

Finalmente, en el esquema de política bajo el modelo de gobernanza, existiría un proceso en el que los actores coordinados, llámense éstos agentes gubernamentales, ciudadanía, iniciativa privada, organizaciones sociales, organizaciones internacionales e instituciones de educación superior, sólo por mencionar algunos, se encuentran relacionados permanentemente por los procesos de agenda pública, la formulación de políticas, la toma de decisiones, y la implementación y evaluación de las políticas, por lo que existiría una transparencia amplia en cada una de los procesos de la gestión de las políticas (véase *Figura 3*).

Figura 3. Esquematzación del modelo de gobernanza.



Fuente: elaboración propia.

Una vez tomadas en consideración estas perspectivas teóricas, en el siguiente capítulo se abordarán el análisis del gobierno abierto, así como el de los pilares conceptuales que lo sustentan, para finalmente realizar un análisis de los resultados que México ha tenido en cuanto a la adopción de los lineamientos del propio gobierno abierto.

PILARES CONCEPTUALES DEL GOBIERNO ABIERTO

Una de las primeras necesidades de los gobiernos con la incorporación del gobierno abierto fue alejarse de las problemáticas en las

Tendencias multidisciplinarias...

que se encontraba la llamada gobernabilidad, o ingobernabilidad, antes señaladas. En este sentido, su utilización se popularizó tanto en el análisis político como en el discurso gubernamental a partir del año 2009 por iniciativa del gobierno del presidente de Estados Unidos de América, Barack Obama. Sin embargo, existen registros que dan cuenta de que su utilización ya tiene mucho tiempo, pues inclusive Parks, en el año de 1957, lo utilizaba para hacer referencia a una necesidad de gobierno encaminada a la incorporación de un derecho que proporcionara a los individuos mayores niveles de conocimiento del funcionamiento y el quehacer gubernamental.

Sin embargo, a partir de la iniciativa de Obama, se extendieron la idea y los principios del gobierno abierto mediante el establecimiento de directrices gubernamentales que varios países se comprometieron a adoptar en la iniciativa llamada “Alianza para el Gobierno Abierto” (AGA). En la actualidad, suman 65 países que comparten la visión del gobierno abierto.

Es oportuno señalar que a la iniciativa original de Estados Unidos de América se sumaron ocho gobiernos en el año 2011, entre los que se encontraba México. Cada uno de los gobiernos incorporados a la AGA elaboró un plan de acción específico para su nación tomando en consideración las problemáticas y los retos particulares de las naciones. Sin embargo, todos los planes de acción se circunscribieron en las líneas o retos generales enmarcados en la idea del gobierno abierto.

Estas líneas generales han sido relacionadas por algunos autores (Blomgren y Foxworthy, 2014) con conceptos como transparencia, rendición de cuentas y participación ciudadana en términos de una gestión gubernamental más horizontal, pues se piensa que sólo con una participación amplia de la ciudadanía en la elaboración de políticas públicas, y no sólo en la elección de representantes, se podría hablar de una verdadera democracia. Estos elementos constituyeron la base de la propuesta del gobierno abierto, los cuales se establecieron en cuatro dimensiones:

- 1) Transparencia
- 2) Participación ciudadana

- 3) Rendición de cuentas
- 4) Innovación y tecnología

La transparencia y la rendición de cuentas toman mayor protagonismo en las discusiones de la profundización democrática de los países, donde destaca el elemento de la información, que resulta esencial en el desarrollo ciudadano, pues ahora se percibe como un derecho no sólo el conocer las decisiones públicas tanto de gobernantes como de funcionarios públicos, sino también el saber por qué tomaron una decisión con relación a otras posibilidades. Cabe destacar que para que la información realmente nos lleve a la transparencia, ésta debe ser suficiente, oportuna, verídica, y ser de fácil acceso para los ciudadanos.

El segundo pilar es la participación ciudadana, en su entendimiento en términos de la gobernanza, es decir, como una coordinación de diversos actores que tienen interés y pueden participar en la elaboración de políticas. Este pilar se aleja de la forma clásica de tomar decisiones públicas verticales o unilaterales de los gobiernos fuertemente jerárquicos.

Finalmente, se encuentra el pilar relacionado con el uso de tecnologías en los procesos de rendición de cuentas y transparencia, así como en su uso para abrir canales y formas novedosas para acercar a la ciudadanía a su gobierno con el objetivo de alcanzar mayores niveles de involucramiento y participación en la forma de gobernar. Tomando en cuenta estos elementos, en el siguiente apartado de este documento se analizan los resultados que ha alcanzado México a partir de su involucramiento con las ideas del gobierno abierto y su incorporación a la AGA.

RESULTADOS DE MÉXICO EN EL GOBIERNO ABIERTO

Para abordar los resultados de México en el gobierno abierto, es necesario señalar que se han tenido tres planes de acción. El primero fue en los años 2011 y 2012, el segundo entre 2013 y 2015, y el actual los años 2016 a 2018. En estos años se ha buscado de forma

Tendencias multidisciplinarias...

consistente alcanzar resultados en veintiséis temas específicos, los cuales se enmarcan en cinco ejes fundamentales para el desarrollo del país:

- 1) Gobierno centrado en la ciudadanía
- 2) Presupuesto abierto y participativo
- 3) Datos abiertos para el desarrollo
- 4) Empoderamiento y participación ciudadana
- 5) Gobernanza de los recursos naturales

Los resultados alcanzados por cada uno de los ejes pueden visualizarse en la *Tabla 1*.

Tabla 1. Resultados por eje de acción del gobierno abierto en México.

| Eje de acción | Resultados |
|--|---|
| Gobierno centrado en la ciudadanía | <ol style="list-style-type: none">1) Tu gobierno en un solo punto gov.mx2) Regulación clara y transparente3) Normas accesibles4) Registro de detenidos5) Base de datos de personas desaparecidas6) Padrón único y abierto de beneficiarios |
| Presupuesto abierto y participativo | <ol style="list-style-type: none">7) Compras abiertas8) Compras inteligentes9) Gasto abierto10) Impuestos abiertos11) Fondo emprendedor abierto y participativo12) Infraestructura para todos13) Sigue la obra pública |
| Datos abiertos para el desarrollo | <ol style="list-style-type: none">14) Política nacional de datos abiertos- datos.gov.mx15) Consejo para datos abiertos16) Datos carreteros abiertos17) Datos de educación superior18) Datos para un México en paz |
| Empoderamiento y participación ciudadana | <ol style="list-style-type: none">19) Democratización de las becas y apoyos gubernamentales para la educación20) Retos públicos |

| | |
|--------------------------------------|--|
| Gobernanza de los recursos naturales | 21) Fonden transparente – reconstrucción.mx 22) Petróleo en beneficio de todos 23) Minería para todos 24) Cuidado del medio ambiente basado en la evidencia 25) Protección participativa del medio ambiente 26) Adhesión de México a la Iniciativa de Transparencia de las Industrias Extractivas |
|--------------------------------------|--|

Fuente: Elaboración propia con datos de <http://gobabiertomx.org/resultados/>.

Se realizó un ejercicio de revisión a fondo del discurso gubernamental de cada uno de los veintiséis resultados proporcionados por el gobierno al finalizar la segunda etapa del país en la AGA. Los datos arrojados fueron dados en el mes de noviembre del año 2015, por lo que de cierta forma resultan novedosos, y muestran que en diferente medida han sido tomados en consideración los fundamentos conceptuales antes mencionados.

Se revisó a qué se hace referencia en cada uno de los resultados con palabras clave, y se tomaron en consideración cuatro dimensiones: en primer lugar, la información/ rendición de cuentas/ y la transparencia; en segundo, la mejora o eficiencia de servicios públicos; en tercero, la participación ciudadana, y finalmente la tecnología/ las TIC/ e Internet.

En veinticinco de los resultados, se hace mención a la información, la rendición de cuentas o la transparencia, por lo que nos podemos dar cuenta de la importancia que le da el gobierno mexicano a estos temas. Sólo en el resultado 15, “Consejo para datos abiertos”, no se hace una mención directa al tema de la transparencia, la información o la rendición de cuentas.

Por otro lado, en dieciocho resultados se habla de la mejora o eficiencia de servicios públicos, por lo que también se puede apreciar que, para el caso específico de México, resulta necesario hacer más eficientes los servicios públicos que históricamente han sido criticados por la población. Por mencionar un ejemplo de lo anterior, está el resultado 10, “Impuestos abiertos”, en el que se intenta hacer más accesible, claro y fácil para cualquier tipo de contribuyente, el pago de impuestos, así como conocer de mejor forma los derechos y las obligaciones en esta misma materia.

Tendencias multidisciplinares...

Algo que sin lugar a dudas es sorprendente es que solamente en seis de los veintiséis resultados se habla de la participación ciudadana, por lo que se entiende que, por el momento, y a casi cinco años de la incorporación del país a la AGA, no resulta fundamental para el gobierno, de manera transversal, el tema de la participación ciudadana.

Finalmente, en veintitrés resultados se hace alusión a las TIC y a Internet; sin embargo, en todas estas menciones se hace alusión a sitios, plataformas o herramientas web para brindar los servicios y la información. Sin embargo, en ninguna se aborda el problema de una mayor amplitud y la necesidad de redoblar esfuerzos por conectar a más personas en el uso de las tecnologías. En este sentido, se considera fundamental tomar en consideración, en las iniciativas del gobierno abierto, la disparidad en el acceso a las tecnologías, conocida por algunos autores como *brecha digital*, pues sólo con un acceso más amplio e incluyente se podría hablar de la verdadera democratización de estos medios, así como de la democracia que se pueda obtener con el uso de estas tecnologías.

REFLEXIONES FINALES

Existen avances muy significativos tanto a nivel internacional como nacional en materia de gobierno abierto, en donde se da cuenta de que la idea de gobierno abierto nace como una alternativa a la forma de gobierno que había permeado la acción pública de muchas naciones hasta el año 2009, cuando se relacionaba con la adopción de prácticas de la iniciativa privada, el adelgazamiento del Estado y, en general, con las visiones gerenciales del New Public Management.

A la par de estos logros, se torna fundamental apostar a la inclusión de más procesos que permitan tanto una mayor participación ciudadana, como una utilización de las tecnologías más extendida que incluya sobre todo, a aquellos sectores de la población que han permanecido al margen de estos beneficios, pues sólo de esta forma se podría hacer realidad el discurso de Obama

al iniciar los trabajos de la AGA, relacionados con la necesidad de alcanzar un nivel de apertura en el gobierno sin precedentes y un sistema de transparencia, participación pública y colaboración que refuercen la democracia, aseguren la confianza pública y promuevan la eficacia y eficiencia gubernamentales.

REFERENCIAS BIBLIOGRÁFICAS

- Aguilar-Villanueva, L. F. (2007). El aporte de la Política Pública y de la Nueva Gestión Pública a la gobernanza. *Revista del CLAD, Reforma y Democracia*, 39: 1–9.
- Alcántara-Sáez, M. (1995). *Gobernabilidad, crisis y cambio*. México: Fondo de Cultura Económica.
- Altman, D. (2001). Crisis de gobernabilidad democrática: orígenes y mapa de lectura. *Revista Instituciones y Desarrollo*, 8 y 9: 1–26.
- Bartolomé, A. I. (2006). Consecuencias de la sobrecarga del Estado y la globalización en la concepción del Estado de bienestar. Hacia un régimen de prestaciones sociales condicionadas. *Revista de Sociología*, (81): 119–129.
- Blomgren, L. y S. Foxworthy. (2014). Collaborative Governance and Collaborating Online: The Open Government Initiative in the United States. En E. Bohne, J. D. Graham, J. Raadschelders y P. Lehrke (eds.), *Public Administration and the Modern State*. Nueva York: Palgrave Macmillan.
- Brugue, Q., R. Goma y J. Subirats. (2005). Gobernar ciudades y territorios en la sociedad de las redes. *Revista del CLAD Reforma y democracia*, (32): 5–18.
- Camou, A. (1995). *Gobernabilidad y democracia*. México: Instituto Federal Electoral.

Tendencias multidisciplinares...

- Gentes, I. (2009). Modernización del Estado y gestión pública participativa en Chile: alcances, obstáculos y perspectivas. *América Latina Hoy*, 42 (0).
- Kaufmann, D., Kraay, A., & Zoido-Lobaton, P. (2000). *Aggregating governance indicators*. Nueva York: Policy Research Working Paper.
- Kenis, P. N., & Schneider, V. (1991). Policy Networks and Policy Analysis: Scrutinizing a New Analytical Toolbox. En B. Marin & R. Mayntz (Eds.), *Policy Networks: Empirical Evidence and Theoretical Considerations*. Boulder, Colorado: Westview Press.
- Mayntz, R. (2002). National States and Global Governance. En *II Inter-American Congress of CLAD on State and Public Administration Reform Lisbon, Portugal* (pp. 1–8). Lisboa: CLAD.
- . (2000). Nuevos desafíos de la teoría de Governance. *Instituciones y desarrollo*, (7), 1–10.
- Parks, W. (1957). Open Government Principle: Applying the Right to Know Under the Constitution. *George Washington Law Review*, 26, 1.
- Pérez Linan, A. (2003). Pugna de poderes y crisis de gobernabilidad: Hacia un nuevo presidencialismo? *Latin American Research Review*, 38(3): 149–164. Disponible en: <http://doi.org/10.1353/lar.2003.0040>
- Prats, J. (2001). Gobernabilidad democrática para el desarrollo humano Marco Conceptual y Analítico. *Revista Instituciones y Desarrollo*, (10): 1–24.
- Torres, G. y A. Ramos. (2008). Gobernanza y territorios. Notas para la implementación de políticas para el desarrollo. *Revista Mexicana de Ciencias Políticas y Sociales*, 1 (203): 75–95.

Centro de Información Político-Electoral

PABLO PRIETO MONTES DE OCA

MOISÉS RUBÉN RAMÍREZ ZÁRATE

*Unidad Técnica de Transparencia y Protección de Datos Personales
Instituto Nacional Electoral*

TENDENCIAS EN GOBIERNO ABIERTO

El término *Open Government* se originó a finales de la década de 1970 en Gran Bretaña (Ramírez-Alujas y Güemes, 2012). De acuerdo con Richard Chapman y Michael Hunt, “[E]n su plano original, se refiere a diversas cuestiones relacionadas con el secreto de Gobierno e iniciativas para abrir las ventanas del sector público hacia el escrutinio ciudadano en aras de reducir la opacidad” (ibíd., p. 101).

Transparentar información pública es una tarea compleja, implica el rediseño de procesos al interior de las oficinas gubernamentales, la asignación de nuevas responsabilidades y la adopción de una nueva cultura con relación al ciclo de vida de la información. El escrutinio ciudadano ejerce presión sobre las burocracias y demanda que éstas generen, organicen, almacenen y difundan la información gubernamental de manera ordenada, sistemática y eficiente.

En este sentido, las Tecnologías de la Información y la Comunicación (TICs), en específico los sistemas de cómputo orientados a la gestión de información, proveen herramientas que facilitan tanto la administración de la información generada, como su difusión y recuperación.

Estas tecnologías no deben asumirse como soluciones totales. La aplicación de las mismas no implica que los procesos de gestión de información sean más efectivos, ni mucho menos que un gobierno sea más transparente. Simplemente son herramientas para administrar la información. Su implementación permite no sólo poner a disposición del público la información, sino que ésta sea localizable, relevante, comprensible, que esté actualizada y sea gratuita o, al menos, de bajo costo (incluso para el propio gobierno) (Darbshire, 2009). No obstante, estos beneficios dependen fundamentalmente del adecuado tratamiento descriptivo de la información, es decir, del uso extenso de metadatos.

El Instituto Nacional Electoral (INE), a través de la Unidad Técnica de Transparencia y Protección de Datos Personales, ha emprendido un proyecto para implementar un repositorio de información institucional denominado “Centro de Información Político Electoral” (CIPE). A continuación, se exponen las etapas de trabajo y las herramientas informáticas implementadas para el desarrollo del proyecto.

DIAGNÓSTICO DEL ESTADO DE LA INFORMACIÓN DEL INSTITUTO NACIONAL ELECTORAL

Durante 2014, el INE efectuó un diagnóstico que tenía como objetivo analizar el volumen, la accesibilidad y el grado de preservación de la información institucional, así como la infraestructura utilizada para su almacenamiento y difusión. También, se analizaron los macroprocesos de generación y administración de la información con el propósito de proponer un nuevo modelo de gestión en la materia.

Entre otros resultados, se observó, por ejemplo, que además del portal web (www.ine.mx) existen microsítios que agrupan información sobre temas en particular, plataformas para la consulta de bases de datos, y el catálogo web de la Biblioteca Central. Cada plataforma fue evaluada para determinar su capacidad y calidad de almacenamiento, preservación, organización y recuperación de información.

Destacó que a pesar de que el portal web da un acceso rápido y relativamente sencillo a los documentos institucionales, debido al volumen de información que aloja esta plataforma, no es la idónea para organizar y preservar la información, pues es susceptible de cambios repentinos o de abandono de la información.

Por otro lado, el catálogo de la Biblioteca Central, mientras brinda mecanismos excelentes para preservar y organizar la información, está orientado a usuarios especializados, lo que dificulta la localización de información para usuarios que no conocen la herramienta de búsqueda.

A partir de estos hallazgos, se decidió desarrollar una plataforma web para la gestión y consulta de información digital de fácil uso y mantenimiento que contara con las siguientes características:

- Un motor para realizar búsquedas en texto completo y un mecanismo de descubrimiento mediante filtros.
- La capacidad para utilizar estándares de metadatos que garantizaran la interoperabilidad con sistemas externos.

Para cumplir con lo anterior, se decidió utilizar un software de código abierto, pues éste brinda las siguientes ventajas:

1. El costo de implementación es más bajo que las soluciones comerciales.
2. Se puede conocer su funcionamiento interno, lo que permite adecuarlo a necesidades específicas.
3. Los equipos de TI institucionales se pueden familiarizar con el código fuente, lo que facilita su mantenimiento.
4. Existen grandes comunidades de usuarios, documentación y ejemplos disponibles en la web, lo que acelera su implementación.

Para el desarrollo de estas tareas, se constituyó un equipo multidisciplinario que permitió integrar distintos conocimientos y habilidades al proyecto. En la *Ilustración 1*, se detallan los perfiles involucrados, así como sus aportaciones.

Tendencias multidisciplinarias...

Ilustración 1. Equipo Multidisciplinario



Fuente: Elaboración propia

Para clasificar semánticamente los documentos, se elaboró una lista de términos controlados a partir del análisis del Código Federal de Instituciones y Procedimientos Electorales y de tesauros político-electorales elaborados por el Instituto en años anteriores. Este listado consta de 25 temas principales, 237 subtemas y 615 temas específicos con los que se puede clasificar desde lo general hasta lo particular cualquier documento generado por el Instituto.

DISEÑO Y APLICACIÓN DE ESTÁNDARES DESCRIPTIVOS

A partir de los resultados del diagnóstico, se identificaron las áreas que generaban la mayor cantidad de información. Éstas se consideraron de alta prioridad para el proyecto:

- Consejo General
- Junta General Ejecutiva
- Dirección Ejecutiva de Capacitación Electoral y Educación Cívica
- Dirección Ejecutiva de Organización Electoral
- Dirección Ejecutiva del Registro Federal de Electores

Como primer paso, se estudiaron los siguientes documentos para determinar el grado de utilización de metadatos al interior del Instituto:

1. Reglamento Interior del Instituto Nacional Electoral
2. Reglamento en Materia de Transparencia y Acceso a la Información Pública
3. Lineamientos para la Publicación y Gestión del Portal de Internet e Intranet
4. Guía para la Identificación de Información Socialmente Útil

Se observó que el uso de metadatos era escaso, discrecional y no estandarizado. Además, se identificó que la información se encontraba dispersa, por lo que antes de describirla era necesario localizarla.

Una vez definidas las áreas iniciales, se estableció un alcance para el proyecto. Se determinó que se sistematizaría la información publicada por las áreas seleccionadas entre los años 2010 y 2015.

ASPECTOS TÉCNICOS DE LA IMPLEMENTACIÓN

Escapa del alcance del presente documento ahondar en los detalles de los componentes de software utilizados. No obstante, cabe aclarar algunos aspectos sobre los sistemas de catalogación (formularios web):

| Componente | Software | Versión |
|-----------------------------|--------------------------|---------|
| Sistema operativo | Linux Fedora Workstation | 23 |
| Sistema de bases de datos | PostgreSQL | 9.4 |
| Servidor HTTP | Nginx | 1.10 |
| Lenguaje de desarrollo | Python | 3.4 |
| Web Framework | Django | 9.0 |
| Administrador de contenidos | Mezzanine | 4.0 |

Tendencias multidisciplinarias...

En cuanto al repositorio de información institucional:

| Componente | Software | Versión |
|---------------------------|--------------|---------|
| Sistema operativo | Linux CentOS | 6.5 |
| Sistema de bases de datos | PostgreSQL | 9.4 |
| Repositorio digital | DSpace | 5.0 |
| Lenguaje de desarrollo | Java | |

INVENTARIADO

Como primer acercamiento al universo documental de las direcciones ejecutivas, se generó un inventario a partir de las Guías Simples de Archivo. Éstas “sirve[n] para identificar los documentos de archivo por medio de su descripción archivística, en la que [...] se identifica el contexto y el contenido de los documentos [...] así como los productores” (Archivo General de la Nación, 2006, p. 3). Esto facilitó conocer los asuntos tratados por cada unidad, los medios de almacenamiento y el promedio de producción documental. De cada documento, se obtuvo y registró la siguiente información:

1. Órgano responsable (Dirección Ejecutiva).
2. Área generadora (Dirección, subdirección o departamento responsable).
3. Título del documento.
4. Descripción breve.
5. Año de inicio y fin del asunto.
6. Volumen (medio de almacenamiento y tamaño)¹.

En el caso de los documentos del Consejo General (CG) y la Junta General Ejecutiva (JGE), el inventario se realizó con base en los reportes de Datos Estadísticos del Consejo General y Junta General. Éstos detallan la fecha y el tipo de sesiones, así como el número

1 Este dato permitió conocer la información que ya se encontraba digitalizada.

documentos emitidos por cada órgano colegiado. Los datos recuperados de estos reportes fueron:

1. Clave: Descriptor alfanumérico único que consta de las siglas del área generadora, el número consecutivo del documento y año correspondiente (por ejemplo, CG01/2015).
2. Autor: Instituto Federal Electoral o el Instituto Nacional Electoral.
3. Área generadora: Unidad responsable del contenido del documento. En este caso, Consejo General o Junta General Ejecutiva.

Por otra parte, se aprovechó la información de los registros bibliográficos existentes en el catálogo de la Biblioteca Central.² Estos datos se cruzaron con la información obtenida en los reportes estadísticos para obtener los siguientes metadatos:

1. Título
2. Tipo de documento³
3. Fecha de sesión⁴

Finalmente, se establecieron dos niveles para la descripción de documentos:

1. General: Metadatos comunes y obligatorios para todos los documentos institucionales.
2. Específico: Metadatos no obligatorios que dependen del área de procedencia y, en especial, del tipo de documento del que se trate. Estos permiten aumentar la precisión de la recuperación de información.

2 Almacenados en un Sistema de Gestión de Bibliotecas Aleph Versión 22.

3 Resolución, acuerdo, dictamen, plan, entre otros.

4 Los acuerdos y resoluciones se discuten y aprueban en sesiones del órgano colegiado correspondiente.

Tendencias multidisciplinarias...

En las siguientes tablas se muestran los metadatos correspondientes a cada nivel descriptivo.

Tabla 1. Metadatos generales

| Metadato | Descripción |
|----------------------|---|
| Id | Identificador único del registro. |
| Autor | Persona o institución propietaria del documento. |
| Área / colaborador | Área administrativa o persona que generó la información. |
| Título | Nombre asignado a cada documento. |
| Tipo de documento | Clase o naturaleza del documento: acuerdo, resolución, programa, entre otros. |
| Tipo de archivo | Formato genérico: texto, audio, video, imagen, etc. |
| Fecha de publicación | Por lo menos, año de publicación. |
| Resumen | Descripción breve del contenido. Texto libre. Máximo cien palabras. |
| URL / URI | Enlace web. |
| Materia | Descripción temática restringida a un listado de términos aceptados. Obligatoriamente se asigna, por lo menos, un tema. |

Fuente: elaboración propia.

Tabla 2. Metadatos específicos

| Metadato | Descripción |
|---------------------------------|--|
| Clave | Identificador alfanumérico del documento. |
| Votación | Párrafo que indica la forma en la que fue aprobado el documento. |
| Fecha de sesión | Día, mes y año de la sesión en la que se aprobó el documento |
| Tipo de sesión | Indica si la sesión correspondiente fue ordinaria, extraordinaria o especial. |
| Aplicación de sanción | Booleano que indica si el documento incluye la aplicación de una amonestación a un partido, institución o persona. |
| Documentos relacionados | Listado de los documentos mencionados como antecedentes. |
| Partidos Políticos involucrados | Indica los partidos políticos mencionados en el título de los documentos. |

Fuente: elaboración propia.

RASTREO Y RECOPIACIÓN DE INFORMACIÓN

Cada documento se rastreó en los sistemas de información del Instituto⁵ y en Internet mediante motores de búsqueda distintos⁶ para reducir sesgos en la recuperación de información. Como criterios de búsqueda, se utilizaron los títulos o las claves de los documentos. Cada documento se buscó en por lo menos dos buscadores. En caso de no localizarlo, se marcaría como “no localizado” para después solicitarlo al área generadora.

Para administrar esta tarea, se diseñaron formularios web accesibles a todos los miembros del equipo e interconectados a una base de datos central, lo que facilitó la distribución de actividades entre el personal. Véase *Ilustración 2*.

Ilustración 2. Formularios para la gestión de metadatos

The image shows a screenshot of a web form titled "Metadatos Generales". The form is used for entering document metadata. It includes a search bar at the top left with the text "Buscar en Google". Below this, there are several input fields and a large text area. The "Clave documento" field contains the value "88/CJ/15/11/2015". The "Título" field contains a long text describing an agreement of the General Council of the National Electoral Institute. Below the title, there are several icons representing different entities: "Instituto Nacional Electoral", "Partido Acción Nacional", "Partido Acción Social", "Partido Revolucionario Democrático", "Partido Nueva Alianza", "Partido Verde Ecologista de México", and "Partido del Trabajo". At the bottom of the form, there are dropdown menus for "Tipo documento" (set to "Acuerdo"), "Institución" (set to "Instituto Nacional Electoral"), "Órgano" (set to "Consejo General"), and "Fecha de emisión" (set to "2015-11-26 - Ordena: Consejo General").

- 5 El portal web del INE y los repositorios de información especializada como NormaINE: <http://norma.ine.mx/es/web/normateca/inicio>. Último acceso en mayo de 2016.
- 6 Google Search, DuckDuckGo y Bing.

Tendencias multidisciplinares...

Los formularios se complementaron con funciones de verificación de los datos ingresados, lo que evitó que se guardaran registros incompletos o con errores de entrada (poner una URL en el campo de título o viceversa).

En la *Ilustración 3*, se muestran las capas de software utilizadas para la gestión del proyecto, así como los beneficios de cada una.

Ilustración 3. Sistemas interconectados



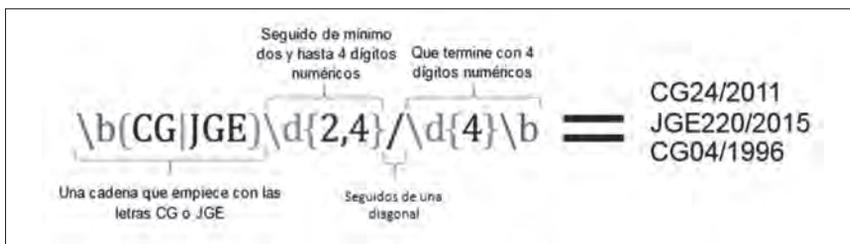
CLASIFICACIÓN ASISTIDA POR COMPUTADORA

Para acelerar la descripción de cada documento se implementaron técnicas de minería de texto y extracción de información desarrolladas con dos herramientas: *Apache Tika-Server* (<https://tika.apache.org/>); un extractor de texto simple, y regex (<https://docs.python.org/3.4/library/re.html>)⁷, un lenguaje de programación especializado en la definición y recuperación de patrones de texto.

El proceso consiste en extraer el texto simple de los archivos a catalogar para después analizarlos mediante regex a fin de detectar patrones de texto correspondientes a un título, párrafos de votación o claves.

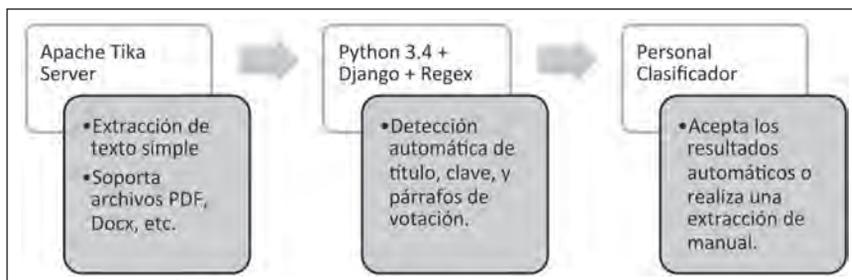
7 Implementación del módulo de regex de Python 3.4.

El siguiente regex se utilizó para detectar las claves de documentos mencionados en los antecedentes de los acuerdos y resoluciones:



El resultado de esta operación es evaluado por el personal clasificador, quien decide si es necesario realizar una extracción manual o aceptar la información extraída automáticamente. Véase *Ilustración 4*.

Ilustración 4. Extracción semiautomática de metadatos



Por otro lado, para extraer información semántica de los documentos, se crearon regex por cada tema y subtema de la lista controlada de términos. Esto conformó un conjunto reglas de reconocimiento de frases relacionadas con temas específicos.

El texto simple es procesado por cada regla y las frases resultantes se contabilizan para calcular la frecuencia de coincidencias del texto con las reglas preestablecidas. El formulario muestra los tres temas con mayor frecuencia y las frases relacionadas con cada

Tendencias multidisciplinares...

uno (véase *Ilustración 5*). Finalmente, el catalogador determina, con base en la cantidad de frases coincidentes y a partir de una lectura rápida del documento, si los temas detectados son correctos. De no ser así, se realiza la clasificación manual.

Ilustración 5. Extracción automática de información semántica

| Clasificación Temática | | |
|---|------------|-------------|
| Buscar Temas | | |
| Tema : Subtema: Microtema | Frecuencia | Reglas |
| Autoridad para la investigación de los delitos electorales-Tipificación de los delitos electorales-Difusión de propaganda gubernamental personalizada | 14 | 189,187,183 |
| Autoridad para la investigación de los delitos electorales-Tipificación de los delitos electorales-Proselitismo o presión a electores | 12 | 218,175 |
| Proceso Electoral-Campañas y Precampañas-Regulación de propaganda electoral | 2 | 48,53 |

| Regla | Frase |
|-------|--|
| 189 | 2 del Reglamento del Instituto Federal Electoral en Materia de Propaganda Institucional y Politico Electoral de Servidores Publicos, con fundamento en lo dispuesto en los articulos 363, parrafo |
| 189 | g del Reglamento del Instituto Federal Electoral en Materia de Propaganda Institucional y Politico Electoral de los Servidores Publicos |
| 189 | la competencia electoral entre los partidos politicos y que dicha propaganda incluya nombres, imagenes, voces o simbolos que impliquen promocion personalizada de cualquier servidor publico, puede motivar el control y vigilancia del Instituto Federal Electoral, |
| 189 | REQUISITOS PARA SU INICIO Y EMPLAZAMIENTO TRATANDOSE DE PROPAGANDA POLITICA O ELECTORAL QUE IMPLIQUE LA PROMOCION DE UN SERVIDOR PUBLICO |
| 218 | factura atinente para la elaboracion de las pintas; e Precisara si para la implementacion del mecanismo en cuestion, utilizo recursos publicos de alguna dependencia o entidad de la Administracion Publica, del partido politico al cual pertenece o, en su caso, de caracter privado; f En caso de que los recursos utilizados sean de origen publico, indicara las circunstancias de modo, tiempo y lugar en que se autorizo o utilizo el egreso respectivo; y g Si para la difusion de la propaganda en cuestion conto con el apoyo de algun funcionario, servidor publico, institucion publica o gubernamental o partido politico, sirviendose precisar nombres, domicilio, terminos y circunstancias de tiempo, modo y lugar en que hubiera recibido el mencionado apoyo; y 5 Girar atento oficio al Vocal Ejecutivo de la 02 Junta Distrital Ejecutiva en el Estado de Mexico para que en apoyo de esta autoridad, se constituyera en los lugares que a continuacion se senalan |
| 218 | politica o electoral que difundan los poderes publicos, los organos autonomos y cualquier otro ente de los tres ordenes de gobierno, bajo cualquier modalidad de medio de comunicacion, pagada con recursos publicos, que pueda influir en la equidad de la competencia electoral entre los partidos politicos y que dicha propaganda incluya nombres, imagenes, voces o simbolos que impliquen promocion personalizada de cualquier servidor publico, puede motivar el control y vigilancia del Instituto Federal Electoral, en atencion al ambito de sus atribuciones y a la especialidad de la materia |

Por último, el registro se almacena en la base datos. Una vez que los metadatos quedan organizados de forma estándar, éstos pueden ser migrados a otros sistemas informáticos, en este caso, a un repositorio digital.

PROTOTIPO DE REPOSITORIO DIGITAL

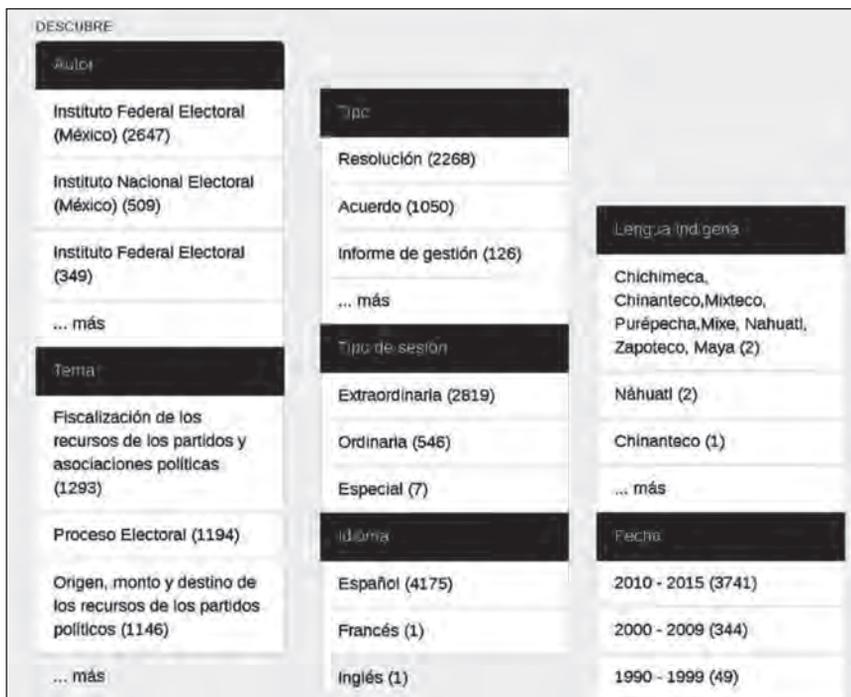
Los documentos catalogados se cargaron en un prototipo de repositorio que permitió demostrar las ventajas de la implementación de estándares descriptivos. El sistema brinda opciones de búsqueda y recuperación de información enriquecidas, con cajas de búsqueda simples y avanzadas que no sólo buscan en los metadatos, sino en el texto completo del documento. También cuenta con capacidad de explorar los documentos mediante filtros y delimitadores que permiten agrupar o discriminar información.

Ilustración 6. Pantalla principal del CIPE



Como se observa en la *Ilustración 6*, el sistema permite organizar la información en colecciones y comunidades. En el caso del CIPE, estas últimas corresponden a las unidades administrativas del Instituto.

Ilustración 7. Filtros disponibles



En la *Ilustración 7*, se muestra cómo el amplio uso de metadatos permite que los documentos del repositorio se puedan filtrar de acuerdo con diversos criterios al utilizar el sistema de descubrimiento integrado en DSpace.

REFLEXIONES FINALES

El desarrollo del Centro de Información Político Electoral se ha basado en dos principios fundamentales. Primero, la aplicación de estándares descriptivos extensos que fortalezcan la organización de la información y que faciliten su recuperación, y segundo, pero no menos importante, la implementación de herramientas informáticas

para administrar, agilizar y hacer más sencillos los procesos de catalogación documental.

Como se mencionó inicialmente, transparentar información pública es una tarea compleja. No obstante, como se ha demostrado, el rediseño de los procesos de gestión de información no es un objetivo inalcanzable. El uso creativo de herramientas de cómputo para la extracción de texto de documentos digitales y la detección de patrones textuales puede apoyar considerablemente a la descripción de documentos. La aplicación de estas técnicas permitió avanzar rápidamente: en un año se lograron clasificar alrededor de cuatro mil documentos correspondientes a cinco años de historia institucional.

Por otro lado, el diseño y la aplicación de estándares descriptivos extensos que privilegien la descripción del contenido sobre otros metadatos, fortalecen la recuperación de información tanto para usuarios generales como especializados.

El uso extenso de metadatos resulta en una capacidad de búsqueda flexible, amplia y precisa: se pueden recuperar, por ejemplo, los documentos del Consejo General que hayan sido aprobados con votos particulares en sesiones extraordinarias del año 2015, en los que además esté implicado algún partido político en particular y que traten sobre fiscalización electoral. Este ejercicio difícilmente es reproducible en buscadores convencionales o con documentos carentes de metadatos.

El conjunto de experiencias obtenidas y los beneficios alcanzados puede representar una referencia de utilidad para el desarrollo de otros proyectos que pretendan fortalecer la gestión de documentación institucional pero, sobre todo, que busquen ofrecer a la ciudadanía información pública de manera transparente, a bajo costo, y de fácil recuperación.

El uso sistemático de metadatos es una condición necesaria para garantizar el derecho de acceso a la información pública. En otras palabras, un documento no descrito es un documento perdido. El diseño y la aplicación de estándares descriptivos debe consolidarse como un pilar de la gestión documental en las instituciones gubernamentales.

REFERENCIAS BIBLIOGRÁFICAS

- Archivo General de la Nación. (2006). *Instructivo para la elaboración de la Guía simple de archivos*. Disponible el 3 de Mayo de 2016 en http://www.agn.gob.mx/menu_principal/archivistica/pdf/instructivo_guia%20simple_abril06.pdf.
- Câmpeanu, C., K. Salomaa y S. Yu. (s.f.). A Formal Study of Practical Regular Expressions. *International Journal of Foundations of Computer Science*.
- Darbishire, H. (2009). *Proactive Transparency: The future of the right to information?* World Bank Institute. Disponible el el 27 de mayo de 2016 en http://siteresources.worldbank.org/WBI/Resources/213798-1259011531325/6598384-268250334206/Darbishire_Proactive_Transparency.pdf.
- Django Documentation*. (2016). The web framework for perfectionists with deadlines. Disponible el 27 de abril de 2016 en <https://docs.djangoproject.com/en/1.9/>
- Instituto Federal Electoral. (2016). Datos Estadísticos de la Junta General Ejecutiva 1990 - 2014. Disponible el 12 de enero de 2016 en http://www.ine.mx/archivos3/portal/historico/recursos/IFE-v2/DS/DS-JGE/DS-DatosEstadisticosSesionesJGE/Estadistico_JGE_90-14.pdf.
- . (2016b). Datos Estadísticos del Consejo General 1990-2014. Disponible el 9 de febrero de 2016 en http://www.ine.mx/archivos3/portal/historico/recursos/IFE-v2/DS/DS-CG/DS-SesionesCG/CG-DatosEstadisticos-docs/Estadistico_CG_1990-2014.pdf.
- . (2016c). Lineamiento para la Publicación y Gestión del Portal de Internet e Intranet. Disponible el 27 de mayo de 2016 en http://www.ife.org.mx/archivos2/Normateca/UTSID/05_Gestoria_Web/01_Lineamientos/2011_JGE103-2011_Lineamientos_Publicacion_Gestion.pdf

- . (2016d). Reglamento en Materia de Transparencia y Acceso a la Información Pública. NormaINE. Disponible el 25 de mayo de 2016 en http://norma.ine.mx/documents/27912/276852/2014_Reglamento_Transparencia.pdf.
- . (2016e). *Reglamento Interior del Instituto Nacional Electoral*. NormaINE. Disponible el 13 de febrero de 2016 en http://norma.ine.mx/documents/27912/276876/Reglamento_Interior_INE.pdf.
- Python Software Foundation. (23 de Mayo de 2016). *Regular expression operations*. Obtenido de Python 3.4.5 documentation: <https://docs.python.org/3.4/library/re.html>
- Ramírez-Alujas, A. y M. Güemes. (2012). Gobierno abierto, reforma del Estado y modernización de la gestión pública: alcances, obstáculos y perspectivas en clave latinoamericana. En Á. Ramírez-Alujas, A. Hofmann, y J. Bojórquez Pereznieto, *La Promesa del Gobierno Abierto* (193-224). México: ITTAIP, InfoDF.

**CULTURA DIGITAL
Y MEDIOS SOCIALES**

Consumidores de información y descripción de documentos: las etiquetas de los usuarios en los catálogos de bibliotecas

ANDRÉS FERNÁNDEZ RAMOS
Universidad Nacional Autónoma de México

INTRODUCCIÓN

La constante evolución de las tecnologías ha propiciado importantes cambios en Internet, que ha evolucionado desde una primera generación basada en el papel pasivo de los usuarios, que eran fundamentalmente consumidores de información, a una segunda generación, denominada la web 2.0, en la que los usuarios participan de forma activa en la creación, organización, difusión y utilización de la información. El motor de dicha evolución ha sido el desarrollo de multitud de aplicaciones, como las redes sociales, las wikis o los blogs, que están centradas en el usuario y permiten un alto grado de interactividad con la información y con otros usuarios.

Las bibliotecas no han sido ajenas a esta evolución y han ido incorporando poco a poco elementos 2.0 a sus sitios web, como la presencia en las redes sociales, la creación de blogs y wikis, los servicios de referencia virtual a través del chat, el uso de plataformas como Youtube para publicitarse o para formar a sus usuarios, o los OPACs 2.0. Todas estas aplicaciones y la filosofía de la web 2.0 están permitiendo a las bibliotecas acercarse más al usuario, conocerle mejor y lograr que se involucre en el desarrollo de la

biblioteca, a la vez que otorga un papel activo y central al sistema bibliotecario y la prestación de servicios (Alonso Arévalo *et al.*, 2014). Tanto es así, que se está hablando de un nuevo modelo de biblioteca denominado “biblioteca 2.0”, cuya esencia es el cambio centrado en el usuario donde éste participa activamente al momento de configurar la prestación de los servicios bibliotecarios (Casey y Savastinuk, 2006).

Aunque no existe un consenso amplio sobre el concepto de biblioteca 2.0, puesto que algunas definiciones inciden más en la interacción, otras en los usuarios y otras en las tecnologías (Holmberg *et al.*, 2009), lo cierto es que esos tres elementos resultan centrales al conceptualizar esta nueva visión de la biblioteca. Cabe destacar, por tanto, que la tecnología en sí misma no es la característica fundamental de la biblioteca 2.0, sino que es un medio que permite al usuario interactuar con la biblioteca y con los otros usuarios de una forma más activa y dinámica.

De entre los muchos servicios bibliotecarios que han incorporado elementos de la web 2.0 para mejorar la experiencia del usuario, destaca el catálogo en línea u OPAC. Estas herramientas, creadas por las bibliotecas para facilitar la búsqueda de documentos en sus colecciones a través de Internet, han evolucionado desde sistemas unidireccionales, en los que el usuario únicamente podía hacer consultas, hasta sofisticadas plataformas con funcionalidades que permiten al usuario disponer de más información y posibilidades de usarla y gestionarla según sus necesidades y preferencias. Algunas de las posibilidades que ofrecen este tipo de catálogos a los usuarios, que han sido denominados *catálogos sociales* o *catálogos 2.0*, serían: disponer de un espacio virtual y personalizado en el cual organizar los documentos en carpetas; la gestión de referencias bibliográficas; la renovación del préstamo; recibir notificaciones de nuevas adquisiciones; la valoración de los documentos; compartir y comentar registros en redes sociales; elaborar reseñas y comentarios, y etiquetar los documentos en función de sus necesidades y preferencias.

De esta última funcionalidad, la posibilidad de etiquetar los documentos de la biblioteca para describir su contenido se hablará a detalle

a lo largo del presente trabajo. Se abordarán las características de dicho sistema de descripción de documentos en comparación con los lenguajes controlados, el uso y la utilidad de las etiquetas de los usuarios en las descripciones de los documentos, tanto en los catálogos como en otros sistemas de información, sus ventajas e inconvenientes, así como su potencial en los diferentes tipos de bibliotecas.

LA DESCRIPCIÓN DE CONTENIDOS MEDIANTE ETIQUETAS

Las *tags* (traducidas al español como “etiquetas”) son palabras o signos que los usuarios de un sistema de información utilizan libremente y en su propio lenguaje para describir el contenido de un documento en función de sus propios intereses (Gerolimos, 2013; Kim, Decker y Breslin, 2010; Rolla, 2009). La práctica del etiquetado de documentos se popularizó a principios del siglo XXI con la aparición de las redes sociales y de diferentes sitios web en los que los usuarios podían colgar y/o gestionar documentos digitales, como por ejemplo Flickr, Del.icio.us o CiteULike. La finalidad del etiquetado era la de permitir a los usuarios organizar la información (fotos, canciones, sitios web, referencias, etc.) en función de su contenido y de acuerdo con sus propias necesidades. Debemos tener en cuenta que esta era la única forma de organizar dicha información, ya que estos documentos no contaban con ninguna otra descripción que la que los usuarios pudieran aportar. Hoy en día, son muchas las aplicaciones que permiten el etiquetado de documentos, puesto que es la forma más sencilla y práctica de organizar y permitir la recuperación de la información en la web, especialmente la de carácter personal que los usuarios continuamente suben a la red (Gerolimos, 2013).

Aunque la utilidad más notoria de esta práctica es que los usuarios puedan organizar la información en su propio espacio virtual, también sirve cuando se aprovechan y gestionan todas esas etiquetas de manera colectiva para buscar información dentro de las aplicaciones o sitios webs donde se han generado (Macgregor y McCulloch, 2006).

Tendencias multidisciplinares...

El conjunto de las tags o etiquetas que los usuarios han utilizado para describir contenidos en un sistema de información da lugar a las *folksonomías* (Rolla, 2009), que son sistemas de clasificación de información no jerárquicos basados en el lenguaje natural de los usuarios, donde las relaciones entre los términos no están predeterminadas. La forma más habitual de búsqueda mediante las etiquetas de una folksonomía es la navegación (*browsing* en inglés) a través de nubes o listados de etiquetas, ya que su falta de especificidad hace que su consulta en una caja de búsqueda sea poco efectiva.

Esta forma de describir contenidos está ganando popularidad, ya que muchos sitios web la utilizan, y ha abierto un importante debate acerca de su utilidad en la recuperación de información, especialmente en comparación con los lenguajes controlados utilizados tradicionalmente en los catálogos y las bases de datos bibliográficas (tesauros, clasificaciones...). Sus ventajas e inconveniencias, así como las diferencias con respecto a los lenguajes controlados, han sido ampliamente analizadas en la literatura científica (Macgregor y McCulloch, 2006; Noruzi, 2006; Porter, 2013; Rolla, 2009; Spiteri, 2006; Steele, 2009) y podrían sintetizarse en la siguiente tabla:

Tabla 1. Diferencias entre las folksonomías y los lenguajes controlados.

| Folksonomías | Lenguajes controlados |
|--|--|
| Falta de normalización, ambigüedad | Normalización del vocabulario |
| Estructura pobre y lineal | Estructura rica en relaciones jerárquicas, asociativas |
| Falta de exhaustividad y precisión | Elevada exhaustividad y precisión |
| Vocabulario del usuario | Vocabulario del bibliotecario |
| Facilidad para hacer búsquedas | Necesita cierto aprendizaje |
| Facilidad y rapidez para asignar etiquetas | Necesita cierto aprendizaje y consume bastante tiempo |
| Actualización / dinamismo | Lentitud en su actualización |
| Subjetividad / aspectos emotivos | Objetividad |

Como puede observarse, las diferencias entre ambos sistemas son evidentes y están determinadas en gran medida por el contexto en el que surgieron. Así, los lenguajes controlados surgieron en las bibliotecas con la finalidad de normalizar el vocabulario que se iba a utilizar para describir el contenido de sus fondos documentales y facilitar su recuperación. La estructura de estos lenguajes es compleja, ya que organizan conceptualmente el conocimiento sobre una o varias disciplinas, y son los profesionales de la información los encargados de la asignación de términos a los documentos. Por el contrario, las folksonomías surgen de la necesidad de describir contenidos en la web por parte de los usuarios para organizar su propia información, por lo que es lógico que su uso sea sencillo y dinámico.

En los últimos años, se aprecia una marcada tendencia al uso de las etiquetas o tags en otros ámbitos, al margen del de la gestión de la información personal, y con otro tipo de documentos que, por su naturaleza, ya han sido descritos o van a serlo en las bibliotecas y centros de documentación. Así, además de en Facebook, Twitter, Instagram o Flickr, las tags se están utilizando para describir el contenido de publicaciones científicas en redes sociales académicas como ResearchGate o Academia.edu, y en gestores de referencias bibliográficas, como Mendeley o CiteULike. También se están utilizando para etiquetar libros en sitios web especializados como LibraryThing, e incluso empiezan a usarse en los catálogos de las bibliotecas.

LAS ETIQUETAS DE LOS USUARIOS EN LOS CATÁLOGOS DE BIBLIOTECAS

El informe del grupo de trabajo de la Biblioteca del Congreso sobre el futuro del control bibliográfico elaborado en 2008 ya recomendaba en su sección 4.1.2 la incorporación de las etiquetas de los usuarios a los catálogos de las bibliotecas, puesto que lo consideraba una posibilidad más de enriquecer el catálogo y proporcionar otro punto de acceso a la información (Library of Congress Working Group on

the Future of Bibliographic Control, 2008). Aunque el debate sobre su utilidad ya estaba abierto y en la literatura ya se podían encontrar trabajos sobre el tema, este informe fue una llamada de atención sobre su importancia. Hay que tener en cuenta que, a diferencia de lo que ocurre en otros sistemas de información, como Flickr o ResearchGate, en los catálogos los documentos ya cuentan con una descripción de su contenido. Entonces, ¿por qué se utilizan o se plantea la utilización de las etiquetas de los usuarios en los catálogos de las bibliotecas? La justificación de su uso ha sido fundamentada en base a diversos motivos, entre los que destacan los siguientes:

- El uso de etiquetas por parte de los usuarios serviría para paliar el desfase existente entre el vocabulario del usuario y el de los lenguajes documentales con que cuenta el catálogo, lo que permitiría a los usuarios buscar y recuperar información con base en su propio vocabulario. Se ha señalado en diversos trabajos de investigación que muchos usuarios no utilizan los lenguajes documentales porque les resultan muy complicados de utilizar, ya que no comprenden su estructura y funcionamiento y además contienen un vocabulario muy técnico que no se corresponde con el que los usuarios utilizan (Fast y Campbell, 2005; Lu y Kipp, 2014; Macías González y Pérez Casas, 2008). Un ejemplo ilustrativo de esta circunstancia es el estudio de Antell y Huang (2008), en el que observa que los estudiantes de la Universidad de Oklahoma utilizan catorce veces más las palabras clave que los encabezamientos de materia al hacer búsquedas en el OPAC.
- La posibilidad de que los usuarios etiqueten los documentos les permite interactuar con el catálogo, personalizar su espacio virtual y organizar la información de acuerdo a sus necesidades (Gerolimos, 2013). Muchos usuarios están familiarizados con las *tags* y las utilizan masivamente en otros espacios web, por lo que cada vez son más demandadas para poder organizar y recuperar la información.

- Las etiquetas de los usuarios aportan descripciones del contenido a las obras de ficción (novelas, poesía, películas, etc.), que generalmente en los catálogos no figuran o aparecen de una forma excesivamente general (Mendes, Quinonez-Skinner y Skaggs, 2009).

Aunque todos estos motivos resultan bastante convincentes, el uso de las etiquetas de los usuarios en los catálogos aún es bastante escaso y no existe unanimidad en la comunidad científica y profesional acerca de sus beneficios reales, tanto en términos de efectividad en la recuperación de la información, como en términos de uso por parte de los usuarios.

La literatura científica ha abordado profusamente el uso y la utilidad de las tags en sistemas de información, pero ha sido principalmente en el ámbito de las redes sociales, gestores de referencias sociales y en el de los sitios web que se permite a los usuarios etiquetar documentos digitales en la web. En el ámbito de los catálogos de bibliotecas, hay menos estudios, pero ya se puede apreciar cierta masa de conocimiento sobre el tema. Con una síntesis de los resultados de dichas investigaciones, se pueden señalar algunas ideas sobre el estado de la cuestión:

- Poco uso. Aún son pocos los catálogos que permiten a los usuarios etiquetar contenidos y utilizar las etiquetas de otros usuarios para buscar información, pero además son pocos los usuarios que aprovechan esta opción en los catálogos en los que sí es posible (Lee y Yang, 2012; Noorhidawati, Hanum y Zohoorian-Fooladi, 2013; Spiteri y Tarulli, 2012).
- Suple carencias en la descripción de ciertos documentos. Se ha visto especialmente en las obras de ficción, que carecen de descripción de materia, en documentos que tratan temas novedosos que no han sido recogidos en las listas de encabezamientos de materia, y en materias cuya descripción ha quedado obsoleta en los lenguajes controlados (Adler, 2009; Mendes Quinonez-Skinner y Skaggs, 2009; Voorbij, 2012).

Tendencias multidisciplinarias...

- Complementan las descripciones del catálogo pero no las sustituyen. Las carencias del lenguaje natural hacen que las búsquedas mediante etiquetas no sean exhaustivas ni precisas, pero puede ser un punto de acceso secundario que aporta otro tipo de información relevante para los usuarios (Macgregor y McCulloch, 2006; Pecoskie, Spitery y Traulli, 2014; Rolla, 2009).
- Actitud favorable de los bibliotecarios. Aunque los bibliotecarios son conscientes de las limitaciones de las etiquetas, la actitud es positiva hacia las iniciativas que enriquezcan el catálogo y mejoren la experiencia de los usuarios (Clements y Liew, 2016; Westcott, Chapell y Lebel, 2009).

CONCLUSIONES

Los catálogos de las bibliotecas han sido durante muchos años el recurso principal que utilizaban los usuarios para localizar documentos de su interés. Hoy en día, existen muchas otras herramientas y formas de buscar información, y los usuarios han cambiado su forma de relacionarse con la información, lo que ha redundado en que el uso de los catálogos esté decreciendo notablemente. Para que los catálogos puedan competir con otros sistemas de información y seguir siendo una herramienta fundamental para los usuarios, deben adaptarse a sus necesidades y preferencias, conocer lo que ofrece la competencia, e incorporar aquellos elementos y funcionalidades que mejoren su aceptación por parte de los usuarios.

La incorporación de las *tags* a los catálogos es una de las muchas formas que tiene la biblioteca de acercarse a sus usuarios y ofrecerle un servicio que usa habitualmente en otros sitios web y con el que ya está familiarizado. De esta forma, habrá mayor interacción del usuario con la biblioteca digital y podrá sentirse más involucrado en su desarrollo, ya que su papel será mucho más activo y además tendrá la capacidad de organizar y recuperar información de acuerdo a sus necesidades y preferencias. En todo caso,

como ya señalaba Margaix-Arnal (2008), será necesario que la biblioteca se implique en gestionar sus relaciones con los usuarios y adopte una actitud que los invite a la participación, puesto que el éxito de los catálogos sociales depende en gran medida de la implicación de los usuarios al etiquetar contenidos.

El uso de las *tags* no supone renunciar a la descripción bibliográfica que se hace en las bibliotecas, sino complementarla con información adicional. Aunque las *tags* carecen de las bondades de los lenguajes controlados y, por tanto, no serán tan eficientes como éstos en la búsqueda de información, también tienen su utilidad, ya que proporcionan otro tipo de punto de acceso a los contenidos que puede ser de ayuda a los usuarios.

REFERENCIAS BIBLIOGRÁFICAS

- Adler, M. (2009). Transcending Library Catalogs: A Comparative Study of Controlled Terms in Library of Congress Subject Headings and User-Generated Tags in LibraryThing for Transgender Books. *Journal of Web Librarianship*, 3(4): 309-331.
- Alonso Arévalo, J., J.A. Cordón García, R. Gómez Díaz, R y B. García-Delgado Giménez. (2014). Uso y aplicación de herramientas 2.0 en los servicios, producción, organización y difusión de la información en la biblioteca universitaria. *Investigación Bibliotecológica*, 28(64): 51-74.
- Antell, K., y J. Huang. (2008). Subject searching success: Transaction logs, patron perceptions, and implications for library instruction. *Reference & User Services Quarterly*, 48(1): 68-76.
- Casey, M. E., y L.C. Savastinuk. (2006). Library 2.0: Service for the Next-Generation Library. *Library Journal*, 131(14): 40-42.

- Clements, L., y C. L. Liew. (2016). Talking about Tags: An Exploratory Study of Librarians' Perception and Use of Social Tagging in a Public Library. *The Electronic Library*, 34(2): 289-301.
- Fast, K. V., y Campbell, D. G. (2005). "I still like Google": University Student Perceptions of Searching OPACs and the Web. *Proceedings of the American Society for Information Science and Technology*, 41(1): 138-146.
- Gerolimos, M. (2013). Tagging for Libraries: A Review of the Effectiveness of Tagging Systems for Library Catalogs. *Journal of Library Metadata*, 13(1): 36-58.
- Holmberg, K., I. Huvila, M. Kronqvist-Berg y G. Widén-Wulff. (2009). What is Library 2.0? *Journal of Documentation*, 65(4): 668-681.
- Kim, H.L., S. Decker y J. G. Breslin.(2010). Representing and sharing folksonomies with semantics. *Journal of Information Science*, 36(1): 57-72.
- Lee, Y. Y., y S. Q. Yang. (2012). Folksonomies as Subject Access - A Survey of Tagging in Library Online Catalogs and Discovery Layers. In *IFLA Annual Conference* (pp. 1-12).
- Library of Congress Working Group on the Future of Bibliographic Control. (2008). *On the Record: Report of the Library of Congress Working Group on the Future of Bibliographic Control*. Disponible en: <https://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>
- Lu, K., y M. E. I. Kipp. (2014). Understanding the Retrieval Effectiveness of Collaborative Tags and Author Keywords in Different Retrieval Environments: An Experimental Study on Medical Collections. *Journal of the Association for Information Science and Technology*, 65(3): 483-500.

- Macgregor, G., y E. McCulloch. (2006). Collaborative Tagging as a Knowledge Organisation and Resource Discovery Tool. *Library Review*, 55(5): 291-300.
- Macías González, J., y P. Pérez Casas. (2008). No es oro todo lo que reluce: una aproximación al nuevo concepto de catálogo social. *Mi biblioteca*, (13): 88-98.
- Margaix-Arnal, D. (2008). El Opac 2.0: Puerta De Acceso a Los Contenidos De La Biblioteca. In *IV Congreso Nacional de Bibliotecas Públicas*. Disponible en <http://eprints.rclis.org/3802/>.
- Mendes, L. H., J. Quinonez-Skinner y D. Skaggs. (2009). Subjecting the catalog to tagging. *Library Hi Tech*, 27(1): 30-41.
- Noorhidawati, A., N. F. Hanum y N. Zohoorian-Fooladi. (2013). Social Tagging in a Scholarly Digital Library Environment: Users' Perspectives. *Information Research*, 18(3), C43.
- Noruzi, A. (2006). Folksonomies: (Un) Controlled Vocabulary? *Knowledge Organization*, 33(4): 199- 203.
- Pecoskie, J., L. F. Spiteri y L. Tarulli. (2014). OPACs, Users, and Readers' Advisory: Exploring the Implications of User-Generated Content for Readers' Advisory in Canadian Public Libraries. *Cataloging & Classification Quarterly*, 52(4): 1-23.
- Porter, J. (2013). Folksonomies in the Library: their Impact on User Experience, and their Implications for the Work of Librarians. *The Australian Library Journal*, 60(3): 248-255.
- Rolla, P. J. (2009). User Tags versus Subject Headings Can User-Supplied Data Improve Subject Access to Library Collections? *Library Resources & Technical Services*, 53(3): 174-184.

Tendencias multidisciplinarias...

- Spiteri, L. F. (2006). The Use of Folksonomies in Public Library Catalogues. *The Serials Librarian*, 51(2): 75-89.
- Spiteri, L. F. y L. Tarulli. (2012). Social Discovery Systems in Public Libraries: If We Build Them, Will They Come? *Library Trends*, 61(1): 132-147.
- Steele, T. (2009). The New Cooperative Cataloging. *Library Hi Tech*, 27(1): 68-77.
- Voorbij, H. (2012). The Value of Library Thing Tags for Academic Libraries. *Online Information Review*, 36: 196-217.
- Westcott, J., A. Chappell y C. Lebel. (2009). LibraryThing for Libraries at Claremont. *Library Hi Tech*, 27(1): 78-81.

Camino para establecer la relación entre usuarios y metadatos

PATRICIA HERNÁNDEZ SALAZAR
Universidad Nacional Autónoma de México

INTRODUCCIÓN

Abordar la relación entre usuarios y metadatos resulta especialmente interesante porque existen eventos académicos en los que se tocan dos temas ejes: la organización de la información y los usuarios. Es de notar que, en términos generales, en estos eventos se abordan en forma separada. De hecho, son dos congresos integrados en uno. No hay participaciones que pretendan ser intradisciplinarias, vincular la organización a los requerimientos reales de los usuarios, se entiende la organización *per se* y se explica sólo para los entendidos.

La división temática es tan determinante, que al final de estos eventos habría que preguntarnos si lo que hacemos los especialistas de la información para organizarla le hace sentido al sujeto final, y si el usuario percibe o se entera y en consecuencia lo agradece. La respuesta a estas dos interrogantes requirió plantear tres objetivos: reflexionar sobre el vínculo existente entre los usuarios, consumidores o prosumidores de información con las plataformas y sistemas de información; determinar si la percepción de los usuarios ejerce influencia sobre las actualizaciones o cambios en los sistemas, y establecer los métodos y las técnicas que permitan

identificar si los sujetos consideran los metadatos al consumir o producir información. Para cubrirlos, se presentan tres apartados principales: conciencia de los metadatos, consumidores de información y prosumidores de información.

CONCIENCIA DE LOS METADATOS

Hablar de usuarios de la información implica considerar los procesos cognitivos que el usuario realiza al buscar, recuperar, usar y crear información. La primera aproximación a estos procesos se establece mediante la percepción de la información, con la consiguiente conciencia, ésta se entiende como: “[...] percatación o reconocimiento de algo, sea de algo exterior, como un objeto, una cualidad, una situación, etc., o de algo interior, como las modificaciones experimentadas por el propio yo” (Ferrater Mora 2001, p. 620).

La primera parte queda muy clara, una persona reconoce algo exterior. ¿Por qué importa la segunda? “Porque se es consciente en tanto que aparecen como modificaciones del yo psicológico” (ibíd.), la conciencia de algo externo mediante la modificación interna lleva al reconocimiento y asimilación de un conocimiento; el individuo/usuario es consciente del objeto y se da la relación sujeto-objeto; antes de esto, el objeto no es, no existe para el sujeto.

A partir de lo anterior, ¿el sujeto/usuario es consciente de los metadatos? ¿Los reconoce? ¿Los asimila como objetos de conocimiento? ¿Se modifica internamente por ellos? La contestación se dará a partir de delimitar el concepto de metadatos. La National Information Standards Organization (NISO) los define como “[...] información estructurada que describe, explica, localiza, o de otra manera hace más fácil recuperar, usar o manejar un recurso de información” (Understanding metadata 2004, p. 1).

Su principal función es “facilitar el descubrimiento de información relevante” (ibíd.); surge la duda de quién descubre la información relevante. La respuesta se da en dos vertientes: una, puede ser que el usuario final diga “ah, qué interesante todas las cosas que encontré en este recurso electrónico” y que, efectivamente, lo

que descubra en ese momento sea relevante; la otra es información relevante descubierta por los especialistas de la información, los generadores de productos electrónicos interesados en que alguien los consuma.

De acuerdo con esto, se pueden crear dos categorizaciones de usuarios, una basada en la conciencia de los metadatos, los que están conscientes y los que no, y otra que tiene que ver con el descubrimiento de información relevante: los que la descubren y los que no. Al cruzar estas categorías con las figuras de usuarios consumidores y prosumidores, encontramos las siguientes categorías: usuario/consumidor, y especialista/productor/prosumidor.

Antes de entrar de lleno al examen de estas categorías, cabe hacer una delimitación conceptual sucinta de los términos *consumidor* y *prosumidor* en ambientes digitales. Ambos términos se acuñan dentro de espacios económicos, sociales y culturales donde el consumidor es el alter ego del capitalismo, ya que “contribuye al proceso de alienación dentro de la esfera del trabajo” (Platriniero y Esposti, 2013, p.22), aunque se le atribuye una capacidad para reapropiarse de los bienes o productos que consume y adecuarlo a sus usos simbólicamente para imprimirles un significado personal. De acuerdo con esta idea, el producto se convierte en un objeto transformado de consumo (ibíd., p. 23), lo cual se puede aplicar al espacio informativo: un consumidor de información es aquel sujeto que pretende cubrir una necesidad de información mediante su búsqueda, recuperación y utilización.

El prosumidor es aquel que conjunta las prácticas de consumo y producción, toma posesión del significado del contenido del recurso y lo reelabora, lo resignifica. El ambiente digital, concretamente la web 2.0, ha promovido mucho el trabajo de los prosumidores, recursos como Facebook, Wikipedia, YouTube, Amazon, The GeoWeb, eBay, blogs y Twitter, entre otros, promueven la interacción y el trabajo colaborativo para la creación de contenidos, entre los que están los metadatos.

CONSUMIDORES DE INFORMACIÓN

Nos referiremos aquí al usuario lego que requiere cubrir una necesidad de información para lo cual busca recursos que la cubran. Por ejemplo, la elaboración de este texto requirió hacer una búsqueda exhaustiva sobre metadatos. La intención era recuperar textos y entre los resultados apareció el sitio Welivesecurity con una entrada titulada “Metadatos: tus fotos podrían mostrar más de lo que ves” (Welivesecurity 2016):



Se hizo una revisión del sitio, de algunos enlaces y se concluyó que no era relevante. No se perciben los metadatos que hay detrás:

```

1 <!DOCTYPE HTML>
2 <html lang="es-LA" prefix="og: http://ogp.me/ns#">
3 <head>
4
5     <meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
6     <meta http-equiv="X-UA-Compatible" content="IE=edge" />
7     <meta name="viewport" content="width=device-width, initial-scale=1" />
8     <meta name="google-site-verification" content="_b2av2wz79orus1H2Dda1fNp51bPaP0H8WnkC2x3k" />
9     <meta name="msvalidate.01" content="CE77828A46C2513F660017CFCB6BA13" />
10    <link rel="stylesheet" type="text/css" media="print" href="http://www.welivesecurity.com/wp-
11    content/themes/esat/css/style-print.css" />
12
13 <!-- This site is optimized with the -->
14 <title>Metadatos: tus fotos podrían mostrar más de lo que ves</title>
15 <meta name="description" content="WeliveSecurity es un sitio de seguridad informática que cubre las últimas
16 noticias, investigaciones, amenazas cibernéticas y tendencias de malware analizadas por los expertos de ESET."/>
17 <link rel="canonical" href="http://www.welivesecurity.com/la-es/2014/05/13/metadatos-fotos-podrian-mostrar-
18 mas/" />
19 <meta property="og:locale" content="es_LA" />
20 <meta property="og:type" content="article" />
21 <meta property="og:title" content="Metadatos: tus fotos podrían mostrar más de lo que ves" />
22 <meta property="og:description" content="WeliveSecurity es un sitio de seguridad informática que cubre las
23 últimas noticias, investigaciones, amenazas cibernéticas y tendencias de malware analizadas por los expertos de
24 ESET." />
25 <meta property="og:url" content="http://www.welivesecurity.com/la-es/2014/05/13/metadatos-fotos-podrian-mostrar-
26 mas/" />
27 <meta property="og:site_name" content="WeliveSecurity" />
28 <meta property="article:tag" content="axiftool" />
29 <meta property="article:tag" content="Fotos" />
30 <meta property="article:tag" content="google maps" />
31 <meta property="article:tag" content="GPS-la-es" />
32 <meta property="article:tag" content="metadatos-la-es" />
33 <meta property="article:tag" content="privacidad" />
34 <meta property="article:tag" content="redes sociales" />
35 <meta property="article:section" content="Herramientas" />
36 <meta property="article:published_time" content="2014-05-13T15:24:28+00:00" />
37 <meta property="article:modified_time" content="2014-05-26T13:21:12+00:00" />
38 <meta property="og:updated_time" content="2014-05-26T13:21:12+00:00" />
39 <meta property="og:image" content="http://www.welivesecurity.com/wp-content/uploads/2014/05/Rubin2.jpg" />
40 <meta name="twitter:card" content="summary"/>
41 <meta name="twitter:description" content="WeliveSecurity es un sitio de seguridad informática que cubre las
42 últimas noticias, investigaciones, amenazas cibernéticas y tendencias de malware analizadas por los expertos de
43 ESET."/>

```

es decir, los datos que el especialista de la información supone importantes; entonces, ¿quién decide lo que es relevante? Los especialistas de la información, es decir, los sujetos que sólo buscan y encuentran información (usuarios/consumidores) no están conscientes de los metadatos y bien a bien no descubren información relevante.

Se llega a las primeras conclusiones al vincular esta afirmación con los dos primeros objetivos planteados; reflexionar sobre el vínculo existente entre los usuarios, consumidores y prosumidores de información con las plataformas y sistemas de información, y determinar si la percepción de los usuarios ejerce influencia sobre las actualizaciones o cambios en los sistemas. No existe un vínculo consciente entre los usuarios/consumidores y los metadatos que ofrecen las plataformas y sistemas de información, el usuario busca y recupera intuitivamente. Y si esto es así, es válido

Tendencias multidisciplinarias...

establecer que los usuarios no influyen directamente en los cambios que los especialistas hacen.

Desde la década de los ochenta, Belkin, Oddy y Brookes (1982) critican la forma en la que se elegían los términos que conforman los Sistemas de Recuperación de Información (SRI), el nombrado empate, los indizadores establecían los términos de recuperación derivados de las herramientas normalizadas, los usuarios al hacer la búsqueda digitaban las palabras que eran relevantes para ellos y el sistema trataba de casarlos con algún término indizado. Ellos proponen considerar al usuario y preguntarle directamente las palabras que le eran relevantes, a lo que le llamaron “enfoco centrado en el usuario”, mismo que hasta la fecha no es cabalmente considerado por los diseñadores de productos electrónicos.

Para probar esta premisa, se hizo una búsqueda exhaustiva sobre estudios de usuarios y metadatos, la mayoría de los documentos recuperados vinculaban al usuario y al metadato a partir del especialista de información, y no al usuario final directamente con el metadato.

Entre los pocos que lo relacionaron, resaltan dos documentos: “Photo annotation on a camera phone” y “From context to content: leveraging context to infer media metadata”, combinan elementos que permiten identificar el vínculo entre usuarios y sistemas de metadatos, y determinar los métodos y técnicas que pueden seguirse para hacerlo, a continuación se describe la investigación.

PHOTO ANNOTATION ON A CAMERA PHONE (WILHELM, A. *ET AL.*)

Objetivo. Evaluar la usabilidad de un sistema construido por los autores que permite crear anotaciones de contenido en las imágenes o fotografías tomadas.

Características del estudio. Longitudinal durante cuatro meses; muestra cuarenta estudiantes de primer año de licenciatura y quince investigadores.

Métodos, técnicas utilizadas e instrumentos utilizados:

- Prueba de interfaz.
- Lluvia de ideas.
- Observación directa. Instrumento: videograbaciones.
- Encuesta. Instrumento: cuestionario.
- Grupo focal. Instrumento: entrevista grupal.
- Entrevista.

Descripción del sistema. El sistema se llamó Mobile Media Metadata (MMM), estuvo conformado por dos componentes, un repositorio de anotaciones o metadatos alojado en un servidor antes que los usuarios cargaran las fotografías, y un buscador construido en lenguaje XHTML en los teléfonos. Esto permitiría la interacción entre el cliente y el servidor remoto y se crearía un repositorio colaborativo de imágenes anotadas con el fin de facilitar el proceso de generación de metadatos o anotación de palabras.

Esta aplicación se grabó en 55 teléfonos. Cada que un sujeto tomaba una foto en el teléfono y la almacenaba con ciertas palabras, se guardaban los datos en el servidor, lo que permitió crear la base de datos de acuerdo con el usuario.

Procedimiento: Evaluación en tres etapas:

- Primera. Ejecutaron la prueba de interfaz con cinco participantes, les mostraron tres escenarios del uso del teléfono, les pidieron que los utilizaran, los observaron, los grabaron, y después los entrevistaron para identificar sus hábitos para tomar, almacenar, compartir y recuperar las fotografías.
- Segunda. Se suministró un cuestionario semanal a los 55 participantes durante siete semanas, con el fin de establecer el uso de los teléfonos y el sistema de anotación implementado.
- Tercera. Se realizaron dos grupos focales, para discutir también sobre sus hábitos tomar, almacenar, compartir y recuperar las fotografías. Un grupo estuvo conformado por ocho sujetos de la muestra, y el otro por siete estudiantes

Tendencias multidisciplinares...

que no pertenecían a la misma. Posteriormente, juntaron los grupos para intercambiar experiencias sobre sus sistemas de anotación.

Resultados

- La integración constante de palabras para identificar sus fotos ampliaba la base de datos, de tal manera que hubo un momento que tuvieron que pasar varias pantallas con dichas palabras y se les empezó a complicar el uso de la interfaz.
- Los sujetos estudiados no estaban interesados en hacer anotaciones con palabras clave. Como las fotografías les vieran en forma inmediata imágenes, no les interesa tener tantas anotaciones.
- Hacen anotaciones como subtítulos.
- Anexan uno o dos identificadores.
- Prefieren buscar mediante los metadatos internos que ofrecen las computadoras de escritorio los paquetes que ya tienen para almacenar y registrar las fotografías.
- Compartir y echar un vistazo es más importante que generar anotaciones para identificar, buscar o recuperar sus fotos.

FROM CONTEXT TO CONTENT: LEVERAGING CONTEXT TO INFER MEDIA METADATA (DAVIS *ET AL.*, 2004)

Dos de los autores del estudio anterior (Davis y Sarvas, 2004) utilizan sus resultados para agregar posibilidades de generación de metadatos al sistema Mobile Media Metadata (MMM) mediante el uso del contexto espacial, temporal y social, y de un diseño de recuperación de la imagen basado en el contenido. Revisaron las brechas semánticas mediante el reforzamiento de la interacción del usuario con el sistema. Lo que trataron de hacer fue controlar la cantidad de palabras que los usuarios elegían al generar patrones de las anotaciones previas de fotos tomadas en lugares y tiempos

similares. El sistema busca entre sus metadatos alguno que describa la nueva foto sin tener que aumentar la base de datos.

El motivo de tratar de resolver de esta manera la generación exponencial de palabras de recuperación es de corte tecnológico:

Desafortunadamente, el bajo nivel de las características que los algoritmos actuales pueden extraer, no es suficiente para enfrentar las necesidades de cómo los humanos quieren buscar y usar el contenido de los medios. La brecha semántica es endémica a los sistemas de información multimedia actuales (Davis *et al.*, 2004, p. 188).

Se considera que la afirmación arriba expuesta por los autores no es del todo cierta, o que las tecnologías de información y comunicación han evolucionado y que diez años después de ese estudio es posible generar algoritmos y subalgoritmos de acuerdo con las necesidades del usuario; lo que importa es el diseño y la precisión en la programación.

Sobresale el hecho de que este estudio pretende controlar la cantidad de palabras que reflejan las imágenes registradas en las fotografías, pareciera que el usuario interactúa con el sistema y crea metadatos, sin embargo esto es así sólo en parte. El sujeto es más creativo y flexible para seleccionar palabras tanto para describir como para realizar búsquedas, no es consciente de que deben ser limitadas y controladas, el ser humano es cambiante, su estado de ánimo influye para elegir las. Se habla aquí del usuario promedio, no de comunidades de investigadores, profesores, alumnos o profesionales.

PROSUMIDORES DE INFORMACIÓN

Un prosumidor de información es aquel que consume, reelabora y produce recursos de información electrónicos, requiere hacer el seguimiento de lo que el usuario/consumidor busca, recupera y eventualmente utiliza. Es una sola categoría, puesto que su objetivo es muy claro: diseñar productos electrónicos para que sean consumidos. Entran aquí los especialistas de la información, y todos los que

tienen un interés por compartir o vender un producto. En esta categoría es muy claro que sí existe conciencia de los metadatos, puesto que durante todo el texto se ha enfatizado que estos actores son los que realizan el proceso de generación de datos que describan.

Existen algunas sugerencias para integrar a los autores de la información o usuarios finales en los equipos que crean los metadatos, con la condición de que sean: “[...] formalmente entrenados para entender los metadatos, los conceptos de los vocabularios controlados y el uso de las herramientas de software relacionados con los metadatos” (Understanding metadata, 2004, pp. 10-11).

Incluso existen comunidades de usuarios que están interesadas en desarrollar y enriquecer esquemas de metadatos, vocabularios controlados y lineamientos, y adecuarlos a audiencias específicas. Un ejemplo es MODS (Metadata Object Description Schema) User Guidelines que ya va en la versión 3.0. Es claro que estas comunidades ya no son sólo consumidoras entran en la presente categoría.

CONCLUSIONES

Más que conclusiones, este apartado pretende apuntalar el objetivo de reflexionar sobre la relación entre los usuarios y los metadatos. Se determinó que los usuarios/consumidores no especialistas que buscan información no están conscientes de los metadatos, y que los generadores de plataformas, sistemas y productos electrónicos no los toman en cuenta, al menos no en forma directa.

Los caminos para establecer la relación entre usuarios y metadatos corresponden al seguimiento de métodos, técnicas e instrumentos de medida que se utilizan en los estudios de usuarios: observación directa, grupos focales, entrevista, lluvia de ideas (aquí lo entenderemos como think aloud), y encuesta.

Como ya se apuntó, los metadatos describen los recursos y permiten descubrir información relevante, ya que responden a las preguntas quién, qué, cuándo, dónde, por qué, para qué y cómo del uso de los datos. Si realmente queremos cambiar el panorama presentado, como especialistas de la información tendremos

que cuestionarnos: ¿Consideramos al usuario final? ¿Nos interesa tenerlo en cuenta? ¿Analizamos los registros de búsqueda para descubrir información relevante de los usuarios? ¿Qué estamos haciendo con esos metadatos?

REFERENCIAS BIBLIOGRÁFICAS

- Belkin, N.J., Oddy R. N. y H. M. Brookes. (1982). ASK for Information Retrieval. Part I: Background and Theory; Part II: Results of a Design Study. *Journal of Information Science*, 38 (2 y 3): 61-71; 145-164.
- Davis, M. *et al.* (2004). From Context to Content: Leveraging Context to Infer Media Metadata. Proceedings of the 12th ACM International Conference on Multimedia. Pp. 188-195. Nueva York: AMC Digital Library. Disponible el 10 de octubre de 2016 en: <http://web.mit.edu/21w.789/www/spring2006/papers/davis2004.pdf>.
- Ferrater Mora, J. (2001). *Diccionario de Filosofía*. Barcelona: Ariel.
- Paltriniero, R. y P. D. Esposti. (2013). Processes of Inclusion and Exclusion in the Sphere of Prosumerism. *Future Internet*, 5 (21-33). Disponible el 28 de septiembre de 2016 en <http://www.mdpi.com/1999-5903/5/1/21>.
- Understanding metadata (2004). National Information Standards Organization. Disponible el 25 de mayo de 2016 en: <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>.
- Welivesecurity. (2016). Metadatos: tus fotos podrían mostrar más de lo que ves. Disponible el 3 de octubre de 2016 en <http://www.welivesecurity.com/la-es/2014/05/13/metadatos-fotos-podrian-mostrar-mas/>.

Tendencias multidisciplinares...

Wilhelm, A. *et al.* (2004). Photo Annotation on a Camera Phone. In Extended Abstracts of the 2004 Conference on Human Factors in Computing Systems (CHI 2004). Viena, abril 24-29. Nueva York: ACM Press, pp. 1403-1406. Disponible el 7 de mayo de 2016 en https://www.hiit.fi/files/admin/sab/SAB04/material/DCC/18_Wilhelm_PhotoAnnotation.pdf.

EDUCACIÓN A DISTANCIA
Y METADATOS

Tendencias en la literatura sobre metadatos de contenidos educativos: atención al formato, la web semántica y el aprendizaje

ALMA BEATRIZ RIVERA AGUILERA

Universidad Iberoamericana Ciudad de México

INTRODUCCIÓN

Las colecciones de material didáctico digital tienen cada vez más presencia en la red y las instituciones educativas apoyan con mayor énfasis el almacenamiento de los recursos educativos por el valor que representan en la educación presencial y a distancia. Los metadatos que dan información sobre este tipo de recursos apoyan la recuperación, la gestión y el seguimiento en el uso. Este documento tiene por objetivo identificar los temas más novedosos presentes en la literatura reciente sobre metadatos educativos. Una vez revisadas publicaciones de 2010 a 2016 en bases de datos de acceso libre y por suscripción, se encontró que los temas recurrentes son:

1. Fusión de esquema de metadatos educativos a fin de describir mejor contenidos educativos diseñados para dispositivos móviles y con elementos multimedia.
2. Articulación de los recursos educativos con perfiles de alumnos, estrategias didácticas y currículo.
3. Como asunto de fondo, la construcción de vocabularios y ontologías para asegurar la recuperación de los contenidos

Tendencias multidisciplinarias...

educativos en un ambiente de web semántica. La mayoría de los proyectos reportados en la literatura sobre metadatos educativos están asociados a la educación superior. La disciplina de los autores que destaca es Ciencias de la computación, los autores de Educación o Bibliotecología en su mayoría colaboran en obras colectivas más que interdisciplinarias.

Para la definición de metadatos educativos, se propone utilizar como encuadre de diseño el Marco de Singapur. Dicho marco propone una arquitectura para el diseño de perfiles de aplicación de metadatos a partir de los modelos de datos y vocabularios del dominio abordado, en este caso el educativo, que puedan implementarse con salidas en formato RDF (Resource Description Framework) para asegurar la integración de los metadatos en esquemas de web semántica. Las reflexiones finales señalan que existe una oportunidad para la comunidad bibliotecaria de tener un rol protagonista en la definición de metadatos y vocabularios que faciliten la recuperación de contenidos educativos. Para ello es necesario profundizar, dialogar y colaborar con los especialistas en cómputo y educación sobre la definición de perfiles de aplicación, las implementaciones de web semántica y la creación de ontologías.

EL VALOR DE LOS METADATOS DE RECURSOS EDUCATIVOS

En las sociedades contemporáneas, la educación presencial y a distancia es apoyada por contenidos digitales, y los metadatos para describir este tipo de recursos son un factor esencial para aprovechar dichos materiales en un contexto educativo cada vez más demandante. Existen cada día más productores de material didáctico. Libros de texto y materiales didácticos han sido tradicionalmente un nicho para las editoriales comerciales, pero cada vez hay más contenidos educativos de acceso abierto. En Estados Unidos, este fenómeno se reportaba como resultado de las crisis económicas y la proliferación de teléfonos inteligentes y las tabletas el consumidor

de información se está convirtiendo en el consumidor de educación en línea (OCLC, 2014).

El informe Horizonte 2015 establece, a partir de un seguimiento global de tecnologías emergentes en educación, dos puntos claves: 1. Hoy en día hay cada vez más presencia de aprendizaje híbrido (presencial y digital) en las universidades, lo cual implica que los profesores están generando constantemente material digital de apoyo a su docencia y 2. A mediano plazo habrá cada vez más Recursos Educativos Abiertos (REA) disponibles en la red (Johnson *et al.*, 2015). Para lograr la recuperación de cualquier tipo de material didáctico digital, es necesario contar con metadatos adecuados que los describan, estructuren y permitan su recuperación y preservación.

Para ganar visibilidad en un mundo interconectado, pero también como una respuesta al alto costo de la educación y la dificultad de los jóvenes de acceder a la misma, las universidades están promoviendo que los REA se almacenen en colecciones disponibles al mundo. Dichas colecciones incluyen una gran variedad de materiales de apoyo al aprendizaje: objetos de aprendizaje, presentaciones, programas de asignaturas, MOOC (Masive Online Open Courses/Cursos Abiertos Masivos en Línea), etc. El crecimiento de dichas colecciones, aunque es visto con muy buenos ojos por los actores, es complejo ya que implica recursos de infraestructura, retos personales para los docentes, cambios en la cultura institucional y metadatos adecuados. Informáticos, pedagogos, expertos en contenido, diseñadores instruccionales¹ y bibliotecarios son profesiones requeridas en proyectos relacionados con metadatos de contenido educativo. Las bibliotecas universitarias tienen con las colecciones de REA un nicho privilegiado de acercamiento a la academia (ACRL, 2016) no sólo promoviendo y diseñando los espacios virtuales que albergan los materiales u ofreciendo espacios físicos y recursos para la creación, estableciendo estrategias de preservación o apoyando a los docentes en la recuperación de recursos existentes, sino también

1 La palabra *instrucciona*l no existe en la lengua española, pero es comúnmente utilizada en el ámbito del diseño de materiales didácticos.

trabajando en la capacitación de los profesores para la descripción adecuada de sus propios recursos a través de los metadatos².

Lo anterior nos invita a estar al día en las tendencias relacionadas con los metadatos educativos. La literatura reciente sobre el tema muestra que se está transitando de un foco en el establecimiento de etiquetas, posible sintaxis e interoperabilidad, a una preocupación centrada en el aprovechamiento de los metadatos en el descubrimiento de recursos en un contexto de web semántica³ (Han y Gao, 2012 y Rodríguez González, 2012), los desarrollos de cómputo que ofrezcan mayores facilidades a los creadores de contenidos educativos en la generación de los metadatos (Casali *et al.*, 2014), el etiquetado para la descripción de formatos específicos (Correia Santos y Iano, 2015; Aichtaich, Benlahmer y Aichtaich, 2013; Hendrix *et al.*, 2012 y Mylonakis *et al.*, 2011) y fortalecer el rol de los recursos educativos en contextos educativos a través de metadatos (Duque, Tabares y Vicari, 2016; Moreno, Duque y Tabares, 2014; Silva-López, Méndez-Gurrola y Sánchez Arias, 2014; Baldiris, *et al.*, 2014; Foulonneau, Ras, Abou Zeid y Atéchian, 2013 y Enoksson, 2011). En el siguiente apartado, profundizaremos sobre las tendencias identificadas que se presentaron de forma sucinta en este párrafo.

METADATOS EDUCATIVOS: DE LA DESCRIPCIÓN E INTEROPERABILIDAD A LA ESPECIFICIDAD Y LAS RELACIONES SEMÁNTICAS

Los metadatos de material didáctico se han enfocado hasta ahora a describir material didáctico para que sea recuperable en repositorios y facilitar su integración a sistemas de gestión de cursos que

2 El tema de los metadatos es muy amplio y fascinante. Un recurso actual y muy didáctico es Mitchell (2015).

3 Para comprender gran parte del esfuerzo relacionado con los vocabularios y las ontologías en los metadatos en general y educativos en particular, es conveniente revisar el artículo fundamental sobre web semántica (Berners-Leem, Hendler y Lassila, 2001).

den seguimiento a la interrelación de un material con un alumno en particular (seguimiento de avance, evaluación, etc.). Para que la descripción e integración de los contenidos sea eficiente es necesario que los metadatos o etiquetas que están asociadas al material sean estándar,⁴ es decir, que se utilicen por las comunidades que colaboran e intercambian este tipo de material. En relación a la descripción de los recursos educativos, aunque ha habido varias propuestas, la que ha logrado mayor estandarización es Learning Object Metadata (<http://ltsc.ieee.org/wg12>). En cuanto a la integración del recurso y seguimiento de avance de un alumno en el contexto de un sistema administrador de secuencias de aprendizaje (por ejemplo, Moodle, Blackboard o aplicaciones *ad hoc*) el estándar es Shareable Content Object Reference Model (SCORM <http://www.adlnet.gov/scorm/>), el cual utiliza a su vez LOM para la dimensión descriptiva de los contenidos. Si bien ambos estándares, LOM y SCORM, están consolidados en el medio educativo, cabe mencionar que las etiquetas de Dublin Core (<http://dublincore.org/documents/usageguide/qualifiers.shtml>) que incluyen metadatos para describir cualquier recurso digital y como referente educativo específico el campo “nivel educativo” (Education Level), se utilizan o mapean cuando los repositorios o colecciones de material didáctico permiten que sus metadatos sean cosechados para ser incluidos en concentradores de metadatos para la recuperación de varias colecciones de forma simultánea.

Una propuesta de esquema de metadatos educativos relativamente reciente es Learning Resource Metadata Initiative (LRMI <http://lrmi.dublincore.net/>). Esta propuesta colaborativa inició en 2011 por la Asociación de Editores Educativos, que es parte de la Asociación Americana de Editores (<http://publishers.org/>) y Creative Commons (<https://creativecommons.org/>). Este esquema de metadatos nació desde la perspectiva de contenidos comerciales para educación básica y alineada, mas no estaba asociada al proyecto

4 Para un resumen de estándares sobre tecnología educativa, incluyendo metadatos, puede verse Association of American Publishers PreK-12 Learning Group (2015).

Schema.org (<http://publishers.org/>). Desde otoño 2014, LRMI está integrada a DCMI pero no ha publicado actualizaciones desde esa fecha, aunque en agosto 2016 se indicaba en su portal, bastante incompleto por cierto, que se está trabajando en finalizar los perfiles para cursos, credenciales y competencias. El esquema toma en cuenta los perfiles previos de contenido educativo tales como IEE-ELOM, DCMI, IMS Global Learning Consortium Learning Resource Metadata (<http://www.imsproject.org/metadata/>), ARIADNE Educational Metadata Recommendation (<http://vs.fernuni-hagen.de/methoden/ils/Organisation/ariadne.html>), SCORM e ISO/IEC Metadata for Learning Resources (<http://www.iso.org/iso/home/search.htm?qt=19788&published=on>). Un aspecto novedoso de este esquema es que supera el nivel descriptivo del mero recurso e integra etiquetas que permiten alinear el material didáctico con el contexto educativo, pero este esfuerzo todavía está en desarrollo (Barker y Campbell, 2015).

Al revisar la literatura sobre metadatos para material didáctico de los últimos años, puede verse que los asuntos pueden agruparse en tres dimensiones: tecnológica, educativa y de organización. Cabe señalar que los aspectos identificados son complementarios; sin embargo, dependiendo de la disciplina e interés particular, los autores hacen énfasis en alguna de las tres dimensiones.

ASUNTOS TECNOLÓGICOS

Esta dimensión incluye el reporte de los académicos argentinos Casali, Deco, Bender y Fontanarrosa (Casali *et al.*, 2014) que abordan la *generación automática de metadatos*. Esto ha sido implementado, al menos parcialmente, por diferentes programas de almacenamiento de objetos digitales como por ejemplo Greenstone y Mendeley. La novedad que presentan estos autores es su implementación en el DSpace. Este tema no es novedoso pero se agradece mantener la inquietud de aprovechamiento del cómputo en etiquetado que no requiere participación humana por ser derivado de las características inmersas en el recurso.

En relación a la propuesta de *metadatos específicos para materiales educativos en diferentes formatos y ambientes*, Correia Santos y Iano (2015) nos ofrecen desde Brasil una ontología que permite la interoperabilidad de metadatos para contenidos educativos de televisión digital en el marco de SCORM. Por otro lado Achtaich, Benlahmer y Achtaich (2013), de Marruecos, proponen para la descripción adecuada de materiales didácticos diseñados para cómputo móvil una extensión del conjunto de etiquetas LOM que denomina Mobile Learning Object Metadata (MLOM), la cual incluye elementos de movilidad y ubicuidad. Asimismo, Hendrix *et al.* (2012) proponen un esquema de metadatos para describir contenidos educativos de audio y video que a través de pedagogía activa, social y exploratoria promuevan el aprendizaje utilizando videojuegos. Estos autores consideran que un esquema específico es necesario por las características multimedia y complejas propias de este tipo de tecnologías. De forma similar (Mylonakis, *et.al*, 2011), ofrecen desde el Reino Unido y Francia una arquitectura integral con editor de metadatos, repositorio, interfaz de usuario e interfaz cosechador OAI-PMH (<https://www.openarchives.org/pmh/>) para contenidos multimedia. El editor de metadatos basado en LOM que ofrecen los mencionados autores permite definir diferentes perfiles de aplicación o conjuntos de metadatos adecuados al tipo de material.

ASUNTOS EDUCATIVOS

En este apartado, se abordarán algunas tendencias en el uso de metadatos educativos con claros fines pedagógicos. Sin embargo, cabe aclarar que las propuestas se valen de medios informáticos tales como el uso de ontologías o interoperabilidad de metadatos.

Duque, Tabares y Vicari (2016) proponen mapear las características de los estudiantes con los tipos de materiales con el objetivo *de ofrecer los contenidos educativos que mejor se adapten al estilo de aprendizaje, intereses y preferencias de los estudiantes*. Esta propuesta tiene como objetivo mejorar la usabilidad de los repositorios de material didáctico en el caso de Colombia. En 2014, Moreno, Duque y

Tendencias multidisciplinarias...

Tabares ofrecían un sistema basado en “razonamiento basado en casos” para generar cursos personalizados a partir de mapear los metadatos de objetos de aprendizaje con variables del perfil de estudiante. La dimensión educativa de la literatura sobre metadatos educativos también incluye el reporte de Silva-López, Méndez-Gurrola y Sánchez Arias (2014) sobre *evaluación en ambientes digitales* apoyada por sistemas que *articulan metadatos y ontologías adecuadas para cursos, evaluación de actividades, perfiles de estilos de aprendizaje, estudiantes y rutas de aprendizaje*. La propuesta de estos autores pretende aportar a la solución del fracaso escolar y atender a la diversidad, las pruebas fueron hechas con estudiantes de la Universidad Autónoma Metropolitana-Azcapotzalco de la Ciudad de México.

Enoksson (2011), en su tesis de pregrado en Suecia, propone una herramienta de edición de metadatos con un enfoque hacia la generación automática de perfiles de aplicación (denominados por este autor perfiles de anotación) que permiten *ligar las entidades material, alumno y currículo*.

Un tema clave es la reutilización de materiales didácticos digitales, la cual sin duda se asocia con las *características y calidad* de dichos materiales. Baldiris *et al.* (2014) de España y Colombia retoman la Teoría de Respuesta al Ítem para recomendar objetos de aprendizaje a docentes y alumnos. Foulonneau, Ras, Abou Zeid y Atéchian (2013) de universidades de Luxemburgo y Líbano ofrecen un componente de software libre que asigna métricas de legibilidad a textos en repositorios digitales. Estos indicadores pueden apoyar la selección de los textos más adecuados para ser reutilizados en la generación también automática de ítems de evaluación. Cualquier bandera de calidad que tenga un material didáctico a fin de cuentas se convertiría de preferencia en un metadato de recomendación para el usuario final.

ASUNTOS DE ORGANIZACIÓN Y RECUPERACIÓN

Los temas organizativos de los contenidos de materiales educativos tienen una gran cercanía con los asuntos tecnológicos. Han y

Gao (2012), desde China, nos indican que para preparar metadatos basados en ontologías es necesario partir del mapeo entre elementos de metadatos y ontologías relacionados con el contenido y el contexto. Los autores sugieren que los esquemas de metadatos son compatibles con ontologías primitivas y que es posible enriquecer la expresión semántica y comprensión por máquinas de los objetos de aprendizaje a través de vocabularios de metadatos con perspectiva ontológica. Rodríguez González (2012), en una tesis doctoral de la Universidad de Alcalá de Henares en España, define una ontología del dominio de los objetos de aprendizaje que precisamente enriquece la expresividad semántica del esquema de metadatos LOM. La propuesta de González (2012) es muy desarrollada y sugiero que se lea por quien quiera comprender mejor la complejidad, viabilidad tecnológica, diseño, implementación y evaluación de una ontología.

Como puede verse en esta breve revisión de la literatura, los metadatos educativos y sus desarrollos tecnológicos, educativos y de organización abordan asuntos interdisciplinarios; sin embargo, las propuestas e innovaciones vienen en gran medida de las ciencias de la computación y la ingeniería en sistemas.

Esto no debe asombrarnos, pues tanto las ciencias de la educación como la Bibliotecología y las Ciencias de la Información se nutren en gran medida de las propuestas de origen tecnológico. Es pertinente señalar que las publicaciones reseñadas aquí en su mayoría no tienen en la composición de sus autores carácter interdisciplinar. Algunas compilaciones que abordan el tema de los materiales contenidos educativos desde diversas disciplinas como por ejemplo *Investigación e innovación...*(2013) o *Conexión de repositorios* (2013) ofrecen capítulos diferentes por disciplina y los relacionados con metadatos generalmente tienen autores del área tecnológica. Asimismo, al revisar la literatura, vemos que el tema de los metadatos educativos tiene también un carácter internacional. Es interesante ver que en este sentido no hay países dominantes, pero el idioma que predomina en las publicaciones es el inglés. Cabe señalar que a nivel latinoamericano hay grupos trabajando sobre metadatos educativos en México, Brasil, Colombia y Argentina.

Al referirse a los contenidos educativos, algunos autores lo hacen al formato de objeto de aprendizaje, el cual se refiere por definición a un tipo de material caracterizado por la granularidad (o contenido atomizado) y reusabilidad desde su diseño. Sin embargo, muchos autores consideran como recursos educativos cualquier tipo de material de apoyo al aprendizaje. Hoy en día, la comunidad interesada en el tema de metadatos para contenidos didácticos considera la diversidad de formatos y estructuras en los materiales digitales que se conciben como recursos educativos. Dentro de esa diversidad ubican a los objetos de aprendizaje como un tipo de recurso educativo más.

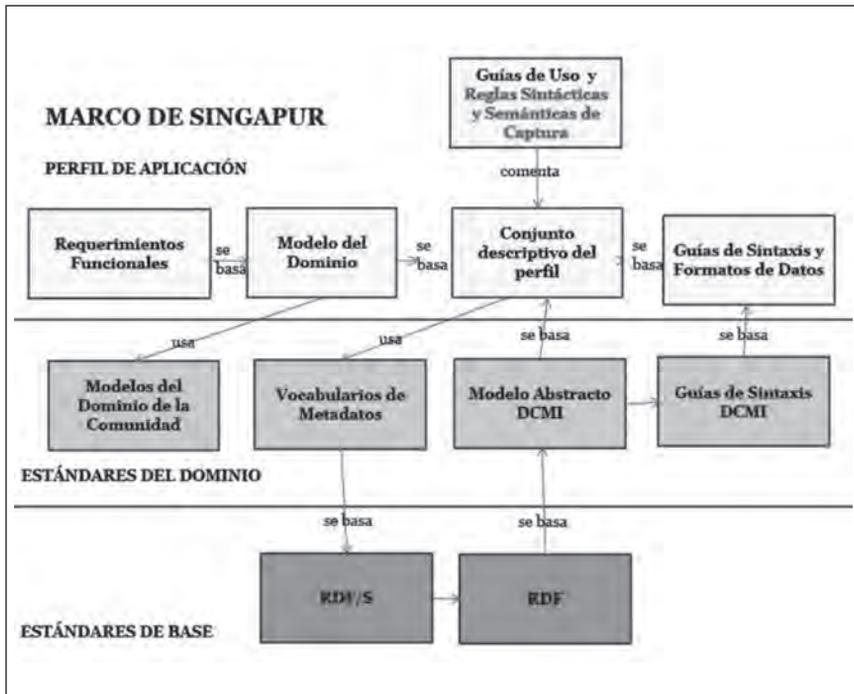
EL MARCO DE SINGAPUR: UNA ARQUITECTURA DE DISEÑO E IMPLEMENTACIÓN DE PERFILES DE APLICACIÓN DE METADATOS

Como vimos en el apartado anterior, la literatura reciente sobre metadatos educativos muestra que se está transitando de un foco en la definición de etiquetas, sintaxis e interoperabilidad, a una preocupación centrada en el aprovechamiento de los metadatos en el descubrimiento en ambientes de web semántica, perfiles que describan adecuadamente diversos formatos y los desarrollos de cómputo que faciliten la generación automática de metadatos, y posibiliten ambientes de aprendizaje dinámicos. Pero ¿qué hacer cuando tenemos que decidir sobre metadatos educativos para un proyecto específico? Un modelo de arquitectura que puede servir de base para el diseño e implementación de perfiles de aplicación de metadatos para casos específicos, como el del material educativo, es el llamado Marco de Singapur (MS).

En 2008, se publicó como documento oficial de Dublin Core el llamado Marco de Singapur o Singapore Framework (Nilsson, Baker y Johnston, 2008), el cual propone una estructura para el diseño de perfiles de aplicaciones de metadatos. Dicha propuesta establece como obligatorio para un perfil de aplicación la documentación correspondiente a requerimientos funcionales, el modelo del

dominio y el conjunto de metadatos descriptivos del perfil (ver *figura 1*). Las guías de uso y de sintaxis y el formato de datos fueron consideradas opcionales.

Figura 1. Marco de Singapur para definir perfiles de aplicación de metadatos.



Este marco sirvió de base para el desarrollo del documento *Guías para perfiles de aplicación Dublin Core* (Coyle y Baker, 2009).

Una de las bondades de dicho marco es que en su capa de estándares de dominio nos pide usar, o bien, desarrollar, vocabularios para la alimentación de los metadatos y en su nivel de base nos invita a basarse en salidas RDF para el formato de datos, lo cual posibilita el descubrimiento semántico.

Cabe aclarar que el MS, al ser un modelo conceptual, puede aplicarse a perfiles que integren tanto etiquetas Dublin Core,

como provenientes de otros estándares, e incluso etiquetas cuyo nivel de modelo de dominio corresponda a entidades de interés exclusivamente local. Para ver un ejemplo de uso del MS, puede verse Rivera Aguilera *et al.* (2011).

REFLEXIONES FINALES

El sondeo de la literatura sobre metadatos en los últimos cinco años nos invita a considerar algunos puntos:

1. Las instituciones y comunidades educativas tienen cada vez más la cultura de apoyar colecciones de materiales didácticos. Existe una gran oportunidad de posicionar a las bibliotecas universitarias como líderes en los proyectos de repositorios tanto a nivel de investigación y desarrollo, como de gestión de las colecciones. Los metadatos educativos son un factor esencial para la existencia de las colecciones y el profesional bibliotecario y de la información es ideal para desarrollar los aspectos relacionados con metadatos.
2. El campo de los metadatos educativos es dinámico y complejo. Requiere de una constante atención a su desarrollo, creatividad y trabajo colaborativo entre disciplinas (Cómputo, Bibliotecología y Ciencias de la información, Educación, Diseño instruccional, etc.) para tener una participación innovadora y efectiva.
3. Es necesario, tomando en cuenta los estándares LOM y SCORM y dando seguimiento a la evolución de LRMI, generar nuevas etiquetas y vocabularios adecuados al tipo de material y los contenidos. En este sentido, el gremio bibliotecario y de las ciencias de la información tiene una oportunidad evidente, pero para aprovecharla debe profundizar y dialogar con los informáticos en relación a la generación automática de metadatos, características de formatos digitales, conceptos de web semántica y el rol de las ontologías, y con educadores sobre currículo, estrategias didácticas, gestión

- educativo, educación a distancia, tecnología educativa, etc. Si el especialista en información no profundiza y dialoga, corre el riesgo de ser ignorado a pesar del conocimiento tradicional que posee en la descripción de recursos de información y manejo de vocabularios controlados.
4. Es conveniente trabajar las propuestas relacionadas con el punto anterior en comunidad para que sean más robustas y tengan posibilidades de consolidarse. El trabajo en grupos nacionales e internacionales es ideal para la generación de nuevos perfiles de aplicación de metadatos y sus correspondientes etiquetas y vocabularios.

REFERENCIAS BIBLIOGRÁFICAS

- ACRL Research Planning and Review Committee. (2016). 2016 top trends in academic libraries: a review of the trends and issues affecting academic libraries in higher education. *C&RL News*, junio: 274-281.
- Achtaich, K., H. Benlahmer, H. y N. Achtaich. (2013). Metadata Defining Object Learning for Mobile Language Learning. *ICERI2013 Proceedings*: 1468-1478. Disponible en <https://library.iated.org/view/ACHTAICH2013MET>.
- Association of American Publishers PreK-12 Learning Group. (2015). *A survey of the Ed Tech Standards Landscape*. Disponible en http://publishers.org/sites/default/files/uploads/PreK-12/ed_tech_standards_final.pdf.
- Baldiris, S. *et al.* (2014). Learning Object Recommendations Based on Quality and Item Response Theory. *2014 IEEE 14th International Conference on Advanced Learning Technologies*, Atenas, 2014, pp. 34-36. Disponible en <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6901391&isnumber=6901368>.

- Barker, P. y L. M. Campbell. (2015). LRMI, Learning resource metadata on the web. En A. Gangemi, S. Leonardi y A. Panconesi. *WWW'15 Companion: Proceedings of the 24th International Conference on World Wide Web*. Nueva York: ACM. Disponible en <http://dl.acm.org/citation.cfm?id=2741745>.
- Berners-Lee, T., J. Hendler J. y O. Lassila. (2001). The semantic web. *Scientific American*. 284(5): 28-37. Disponible en https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf.
- Casali, A. *et al.* (2014). Extracción automática de metadatos de objetos digitales educativos. IX Conferencia Latinoamericana de Objetos y Tecnología de Aprendizaje. Manizales, Colombia: 20 al 24 de octubre. Disponible en <http://www.laclo.org/papers/index.php/laclo/article/view/236/218>.
- Conexión de repositorios educativos digitales, educonector.info.* (2013). F. J. Mortera Gutiérrez y M. S. Ramírez Montoya, coords. México: LULU. Disponible en http://catedra.ruv.itesm.mx/bitstream/987654321/778/1/Conexion%20de%20repositorios%20educativos%20digitales_Educonector.info.pdf.
- Correia Santos, M. y Y. Iano. (2015). MPEG-SCORM: An ontological approach of interoperable metadata for digital television and e-Learning. In *Advanced Communication Technology (ICACT), 17th International Conference*, pp. 782-786. IEEE. /International Journal of Computer Systems. Diponible en http://www.ijcsonline.com/IJCS/IJCS_2015_0205012.pdf
- Coyle, K. y T. Baker. (2009). *Guidelines for Dublin Core Application Profile*. DCMI. Disponible en <http://dublincore.org/documents/profile-guidelines/#sect-1>.

- Duque, N., V. Tabares y R. Vicari. (2016). Learning Object Metadata Mapping with Learning Styles as a Strategy for Improving Usability of Educational Resource Repositories. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, 99: 1. Disponible en <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7452564&isnumber=6461616>.
- Enoksson, F. (2011). Flexible Authoring of Metadata for Learning: Assembling Forms from a Declarative Data and View Model. Tesis de licenciatura en Tecnología de Medios. Escuela de Ciencias de la Computación y la Comunicación KTH. Suecia. Disponible en <http://swepub.kb.se/bib/swepub:oai:DiVA.org:kth-32818?tab2=abs&language=en>.
- Han, Q. y F. Gao. (2012). Towards Semantic Learning Object Metadata: Mapping Standard Metadata Specifications to Ontologies. En *Teaching, Assessment and Learning for Engineering (TALE)*. *IEEE International Conference on TALE*, pp. H1C-12 a H1C-14. IEEE. doi: 10.1109/TALE.2012.6360326.
- Hendrix, M. *et al.* (2012). Defining a Metadata Schema for Serious Games as Learning Objects. En *eLmL 2012, The Fourth International Conference on Mobile, Hybrid, and On-line Learning*, pp. 14-19. Disponible en <http://www.meducator.net/dissemination.activities/files/210.pdf>.
- Investigación e innovación en sistemas y ambientes educativos: primeras contribuciones desde el espacio común de educación superior abierta y a distancia.* (2013). M. E. Chan Núñez. y V. G. Sánchez Arias, coords. Guadalajara: UDG Virtual. Disponible en http://biblioteca.udgvirtual.udg.mx/eureka/pudgvirtual/investigacion_innovacion.pdf.
- Johnson, L. *et al.* (2015). *NMC Horizon Report: 2015 Higher Education Edition*. Austin, Texas: The New Media Consortium. Disponible en <http://cdn.nmc.org/media/2015-nmc-horizon-report-HE-EN.pdf>.

Tendencias multidisciplinarias...

- Mitchell, E. (2015). *Metadata Standards and Web Services for Libraries, Archives, and Museums: an Active Learning Resource*. Santa Bárbara: Libraries Unlimited.
- Moreno, R., N. Duque y V. Tabares. (2014). PeCoS-CBR... Personalized Courses System with Case-Based Reasoning. *Novena Conferencia Colombiana de Cómputo (9CCC)*, 69-74. Pereira. Disponible en <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6955357&isnumber=6955328>.
- Mylonakis, M. *et al.* (2011). Metadata Management and Sharing in Multimedia Open Learning Environment (MOLE). En E. García-Barriocanal, Z. Cebeci, A. Öztürk, M.C. Okur (eds.) *Metadata and Semantic Research*. 5th Conference Internacional, MTSR 2011, Izmir, Turquía, octubre 12-14, pp. 275-286. Berlín/Heidelberg: Springer. Disponible en http://link.springer.com/chapter/10.1007%2F978-3-642-24731-6_29#page.
- Nilsson, M., T. Baker y P. Johnston. (2008). *The Singapore Framework for Dublin Core Application Profiles*. Disponible en <http://dublincore.org/documents/2008/01/14/singapore-framework/>.
- OCLC. (2014). *At a Tipping Point: Education, Learning and Libraries: a Report to OCLC Membership*. Disponible en <http://www.oclc.org/reports/tipping-point.en.html>.
- Rivera Aguilera, A. B. *et al.* (2011). *Propuesta de un perfil de aplicación para material didáctico basado en el Marco de Singapur*. En F.F. Martínez Arellano (comp.). *6º Encuentro de Catalogación y Metadatos: Memoria (100-114)*. México: CUIB-UNAM. Disponible en <http://eprints.rclis.org/22890/1/VI%20ENCUENTRO%20DE%20CATALOGACI%3%93N%20-%20PUBLICACION%20FINAL%20ABRIL%2014-2014.pdf>.

- Rodríguez González, M. E. (2012). Enriquecimiento semántico de metadatos educativos apoyado en ontologías de conocimiento general. Tesis doctoral en Ciencias de la Computación. Universidad de Alcalá de Henares. Disponible en <http://dspace.uah.es/dspace/handle/10017/14022>.
- Silva-López, R. B., I. I. Méndez-Gurrola y V. G. Sánchez Arias. (2014). E-assesment: Ontological Model for Personalizing Assessment Activities. *AIML (Artificial Intelligence and Machine Learning) Journal*, 14(1): 33-40. Disponible en <http://icgst.com/journals/paper.aspx?pid=P1121416325&subid=40>.

La importancia de los metadatos para el aprovechamiento de los recursos educativos en la educación a distancia

JORGE LEÓN MARTÍNEZ
EDITH TAPIA RANGEL

Universidad Nacional Autónoma de México

INTRODUCCIÓN

Las Instituciones de Educación Superior (IES) transitan hoy en día por un ambiente en constante cambio. Diariamente se genera conocimiento y se incrementa cada vez más la necesidad de las personas por actualizarse constantemente. Bajo este esquema, las IES deben realizar diferentes acciones a fin de poder satisfacer las necesidades de aprendizaje actuales y futuras de las personas, acciones que deben considerar los siguientes aspectos:

- Instituciones que diseñan opciones de aprendizaje mixtas (a distancia y presenciales).
- Educación superior que amplía su rango de participación en las sociedad.
- Tecnología que potencializa las capacidades humanas, superando las barreras espacio-temporales.
- Educación cada vez más pertinente a cada persona y su contexto apoyada de investigación basada en el aprendizaje.
- Desarrollo masivo de recursos educativos multiplataforma, reutilizables, y abiertos.

Tendencias multidisciplinares...

Aunado a los aspectos anteriores, los estudiantes o público en general requieren encontrar de una forma cada vez más rápida y adecuada los recursos educativos que les sirvan para su aprendizaje, por lo que los metadatos son esenciales para facilitar la búsqueda y recuperación del recurso educativo que atienda a la necesidad de forma más precisa.

EDUCACIÓN A DISTANCIA

De acuerdo con García Aretio (2014: 46), la educación a distancia “se basa en un diálogo didáctico mediado entre docentes de una institución y los estudiantes que, ubicados en espacio diferente al de aquellos, aprenden de forma independiente o grupal”.

El mismo autor señala que son cuatro las características mínimas de un acto de educación a distancia:

- Separación espacial y en muchos casos temporal del profesor y el alumno durante una gran parte del tiempo de formación.
- Alumnos autónomos, es decir, que estudian de forma independiente y son capaces de autorregular su estudio.
- Comunicación mediada por tecnologías entre todos los participantes y entre los participantes y los recursos didácticos, de manera que se resuelve la separación señalada en el primer aspecto.
- Soporte institucional para la planificación, el diseño y la producción del acto de educación a distancia, asimismo dispondrá de los recursos materiales, humanos y tecnológicos necesarios para el seguimiento y la motivación de los alumnos.

RECURSOS EDUCATIVOS

Los recursos educativos son un término que permite “dar cabida a una amplísima gama de agregados de objetos educativos digitales

que es heterogénea, tanto en el producto, en la cosa en sí, como en los procesos por los que son motivados y sus opciones de reutilización, rediseño o reusabilidad” (Gértrudix, Álvarez *et al.*, 2007: 16). Se considera que son heterogéneos porque recombinan recursos verbales, orales y escritos; icónicos, visuales, auditivos y audiovisuales; estáticos y dinámicos; figurativos y abstractos; íconos, índices y símbolos; analógicos y digitales (García, 2005).

Los contenidos educativos son cruciales para el aprendizaje mediado por tecnologías en cuanto que son:

- Relevantes, actualizados y auténticos.
- Explorados en muchos niveles.
- Manipulables.
- Instantáneos.
- Creativos (CEO FORUM, 2000).

Respecto al diseño de los contenidos educativos digitales, Martínez (2003) nos expone algunas consideraciones importantes que se deben cubrir

- Promover el *aprender haciendo*.
- Enseñar a partir de historias, casos y ejemplos de expertos.
- Disponer de información relevante en el momento en que se necesita —aproximación del contenido a la realidad del estudiante—.
- Generar motivación e interés por la materia puesto que son la energía para el aprendizaje.
- Permitirle al alumno cometer errores en entornos seguros de aprendizaje (PC, juegos, simulaciones).
- Permitir al estudiante controlar su proceso de aprendizaje a su propio ritmo.
- Proporcionar diferentes vías que permitan al estudiante implicar diferentes sentidos.
- Conjuntar aprendizaje y diversión.
- Incorporar emociones en la búsqueda de provocar situaciones memorables, intensas, perdurables.

Tendencias multidisciplinares...

- Promover el *enseñar a aprender* para que el alumno logre el aprender a aprender.

MODELO DE DESARROLLO DE CONTENIDOS EDUCATIVOS

La Universidad Nacional Autónoma de México (UNAM), a través de la Coordinación de Universidad Abierta y Educación a Distancia (CUAED), y como parte de sus funciones, ha apoyado diferentes entidades académicas en el desarrollo de miles de contenidos educativos digitales que beneficien tanto a la comunidad universitaria como a la sociedad. Para ello, ha desarrollado un modelo que se basa principalmente en tres rubros:

MODELO BÁSICO DE CONTENIDO: UNIDAD DE APOYO PARA EL APRENDIZAJE (UAPA)

En la CUAED, se diseñó hace algunos años un modelo básico de contenido denominado Unidades de Apoyo para el Aprendizaje (UAPA), las cuales son materiales didácticos que consideran las propiedades de los objetos de aprendizaje, que además incorporan elementos para enriquecerlos y contextualizarlos, lo que otorga sentido e integridad a la situación formativa y posibilita que el alumno tenga a su alcance lo necesario para desarrollar un aprendizaje específico, significativo y autónomo.

Gran parte de los contenidos de los programas académicos desarrollados en CUAED se basan en el modelo UAPA que incluye objetivo, contenido, actividades de aprendizaje, evaluaciones, y fuentes de información. Esta estructura y sus componentes varían de acuerdo al enfoque pedagógico que se define para el proyecto académico.

METODOLOGÍA DE DESARROLLO

Todos los contenidos elaborados con apoyo de la CUAED se implementan mediante una metodología de desarrollo propia, que consta de las siguientes fases:

1. Definición del perfil del proyecto: consta de dos etapas, una de carácter administrativo donde se establecen datos del proyecto, participantes y calendarización, y otra de carácter formativo, donde se define el modelo educativo que sustentará la propuesta y el diagrama de navegación (niveles de navegación), y las bases para definir el guión instruccional, que es el instrumento básico para esta metodología. Asimismo, se comienza el llenado del guión de información general y los componentes generales de la propuesta.
2. Análisis de contenidos: los expertos, acompañados de un asesor pedagógico, establecen los contenidos que integrarán el recurso educativo.
3. Asesoramiento pedagógico y corrección de estilo: comprende dos etapas importantes; la primera es el Asesoramiento pedagógico, que consiste en el desarrollo de las actividades de aprendizaje y evaluación, y la organización de la propuesta de manera general; la segunda es la corrección de estilo, donde se revisa la redacción del guion instruccional y se consultan dudas acerca de los términos empleados.
4. Comunicación Visual e Integración: se realizan actividades como traza de imágenes, implementación de herramientas de comunicación en plataforma, integración de contenidos, modificación de plantillas para la presentación de contenidos y actividades, creación de recursos multimedia a la medida, entre otras.
5. Implementación, evaluación y actualización: una vez integrado por completo el curso en plataforma, se realiza el pilotaje para su evaluación, de tal forma que permita detectar problemáticas que se resolverán en la actividad de actualización.

EQUIPO DE TRABAJO

La metodología de desarrollo se apoya del trabajo colaborativo de equipos multidisciplinares, quienes conjuntan sus esfuerzos, con lo cual se logra una optimización de recursos y genera un capital intelectual que incorpora los diversos pensamientos y conocimientos del equipo que participa. Los perfiles que integran estos equipos son:

- Experto en contenido. Académico designado que desarrolla los contenidos de los programas o recursos educativos y los materiales didácticos necesarios, y selecciona las fuentes de información y los medios de comunicación e información.
- Administrador de proyectos. Coordina los esfuerzos humanos y materiales para llevar a cabo el proyecto de desarrollo de los recursos educativos a buen fin; identifica los requerimientos y alcances del mismo para planificarlos y dar seguimiento a las actividades.
- Asesor pedagógico. Realiza la labor de acompañamiento al experto en contenido, analizando materiales; vigilando la congruencia entre objetivos, actividades, materiales y evaluación; identificando la interacción y comunicación; apoyando el diseño de actividades de aprendizaje, autoevaluación y evaluación; desarrollando el guion instruccional y verificando que éste último se encuentre correctamente integrado en pantalla.
- Corrector de estilo. Revisa la redacción de los guiones instruccionales y materiales ex profeso; adecua el estilo al interior del guion instruccional en relación con lo acordado en el manual de estilo para el proyecto.
- Comunicador visual. Conceptualiza la interfaz gráfica del sitio electrónico; integra los componentes generales; edita las imágenes y textos; adapta las plantillas de recursos o elementos visuales; elabora animaciones, esquemas y diagramas, y verifica la accesibilidad del sitio electrónico.

- Abogado de derechos de autor. Asesora al equipo respecto a la normatividad que debe respetarse en materia de propiedad intelectual, sobre todo cuando se hace uso de materiales de terceros.
- Experto audiovisual. Realiza la producción y posproducción de material multimedios, sobre todo de video.
- Ingeniero en sistemas. Realiza la programación de herramientas que son requeridas por el recurso educativo, tomando en consideración la solicitud del experto y del asesor pedagógico.

Ilustración 1. Equipo de trabajo



RECURSOS EDUCATIVOS DESARROLLADOS

Como resultado del trabajo conjunto entre las entidades académicas y la CUAED, se tiene

- Un bachillerato a distancia (B@UNAM) que tiene presencia a nivel nacional e internacional.
- Veinte programas de licenciatura en la modalidad a distancia, los cuales se muestran en la siguiente tabla:

Tendencias multidisciplinares...

Tabla 1. Licenciaturas en línea

| Entidad Académica | Carrera |
|--|--|
| Facultad de Ciencias Políticas y Sociales | Ciencia Política y Administración Pública |
| | Ciencias de la Comunicación |
| | Relaciones Internacionales |
| | Sociología |
| Facultad de Contaduría y Administración | Contaduría |
| | Administración |
| | Informática |
| Facultad de Derecho | Derecho |
| Facultad de Economía | Economía |
| Facultad de Estudios Superiores Acatlán | Enseñanza de Alemán como Lengua Extranjera |
| | Enseñanza de Francés como Lengua Extranjera |
| | Enseñanza de Italiano como Lengua Extranjera |
| | Enseñanza de Inglés como Lengua Extranjera |
| | Enseñanza de Español como Lengua Extranjera |
| Facultad de Estudios Superiores Cuautitlán | Diseño y Comunicación Visual |
| Facultad de Estudios Superiores Iztacala | Psicología |
| Facultad de Filosofía y Letras | Bibliotecología |
| | Pedagogía |
| Escuela Nacional de Enfermería y Obstetricia | Enfermería (Profesionalizante) |
| Escuela Nacional de Trabajo Social | Trabajo Social |

Fuente: elaboración propia.

- **Asignaturas a distancia de programas presenciales para diversas escuelas y facultades como: Filosofía y Letras, Estudios Superiores Acatlán, Estudios Superiores Iztacala, Medicina, Música, Trabajo Social, Arquitectura, y Colegio de Ciencias y Humanidades.**
- **Especializaciones a distancia, la CUAED apoya el desarrollo de seis programas para diversas facultades como: Medicina Veterinaria y Zootecnia, Estudios Superiores Zaragoza, Psicología, Ingeniería, Estudios Superiores Acatlán, Derecho, y el Centro de Enseñanza para Extranjeros.**
- **Maestrías para algunos programas de posgrado como: Maestría en Docencia para la Educación Media Superior y Posgrado en Ciencias de la Administración.**

- Diferentes recursos educativos que apoyan los distintos programas académicos que se imparten en la UNAM.

| Tipo de Recurso | Cantidad desarrollada |
|---|-----------------------|
| Libros electrónicos | 57 |
| Lecciones interactivas de inglés (EnglishMedia) | 63 |
| Apps | 26 |
| MOOC | 27 |
| Cursos Aprendo+ | 51 |
| Programas de televisión | 3,214 |
| Lecturas para aprender mejor | 12 |
| UAPAS | 182 |

Fuente: elaboración propia con base en <http://suayed.unam.mx/oferta/>.

- En cuanto a Educación Continua, la CUAED se ha vinculado con diversas entidades tanto internas como externas a la UNAM en los sectores educativo, social, productivo y gubernamental para desarrollar decenas de proyectos académicos de gran impacto a nivel regional y nacional, en espacios y ambientes educativos mediados por Tecnologías de la Información y la Comunicación (TIC).

METADATOS DE RECURSOS EDUCATIVOS PARA LA EDUCACIÓN A DISTANCIA

Los metadatos son información legible para los equipos de cómputo acerca de los recursos digitales. Dicha información describe las características del recurso, lo que facilita su manejo y recuperación. Un conjunto de metadatos combinados con un propósito específico constituye a su vez un esquema de metadatos.

Los metadatos se han utilizado durante muchas décadas (en las bibliotecas principalmente) para facilitar la descripción común de recursos. La aparición de la WWW dio lugar a la disponibilidad de millones

de recursos, lo que incrementó la importancia de los metadatos para facilitar su búsqueda y recuperación.

Si bien la adopción de un único estándar de metadatos facilitaría la reutilización de los distintos recursos digitales, así como la interoperabilidad de los mismos entre plataformas, no ha sido posible lograrlo y hoy en día no existe una gran riqueza de esquemas de metadatos capaces de cumplir con los requerimientos y las necesidades de cada institución. Algunos de estos esquemas se enfocan en metadatos técnicos, otros en educativos y algunos son muy especializados.

Considerando lo anterior, los esquemas de metadatos para el área educativa requieren la definición de los atributos especializados necesarios para describir de forma completa y adecuada los recursos educativos digitales que puedan recuperarse y reutilizarse en los programas educativos tanto en modalidad a distancia como abierta y presencial. Entre los atributos que se añadirían están, por ejemplo: estilo de aprendizaje que favorece a posibles usuarios finales y contexto de la instrucción.

Los esquemas de metadatos desarrollados hasta ahora para los recursos educativos digitales son:

- IEEE LOM (IEEE Learning Object Metadata). Desarrollado por el IEEE Learning Technology Standards Committee (LTSC). Es uno de los estándares más generalizados de metadatos educativos y se centra principalmente en su descripción. Incluye más de 60 elementos organizados en 9 categorías (general, ciclo de vida, meta-metadatos, técnicos, educativos, derechos, relación, comentarios, clasificación).
- Sharable Content Object Reference Model (scorm). Es un modelo de referencia que controla cómo organizar, describir y enlazar el contenido de aprendizaje con los Sistemas de Gestión del Aprendizaje (Learning Management Systems-lms). scorm permite la extensión de lom, por lo que permite a las organizaciones añadir nuevos elementos y mejorar los vocabularios controlados existentes.

- Dublin Core Metadata Initiative (dcmi). Ha sido desarrollado por las organizaciones con el fin de facilitar el intercambio de cualquier tipo de recursos web genéricos. En su versión inicial, el Conjunto de Elementos de Metadatos Dublin Core (dcmes), conocido como el Dublin Core (dc), se componía por quince elementos. La posterior versión Dublin Core Cualificado (qdc) se extendió con siete nuevos elementos. Sin embargo, incluso esta versión enriquecida del esquema de dc no es capaz de captar los aspectos pedagógicos de un recurso educativo.
- Iniciativa de Metadatos para Recursos de aprendizaje (Learning Resource Metadata Initiative – Lrmi: <http://lrmi.dublincore.net/>). Su objetivo es crear un esquema de metadatos uniforme para el etiquetado de contenidos educativos en Internet, de tal forma que permita a los motores de búsqueda más comunes (Google, Yahoo!) ofrecer filtros específicos en materia educativa tales como nivel educativo, público objetivo, tiempo de estudio, estilo de aprendizaje predominante que apoya el recurso, tipo de recurso, edad apropiada, y más. Esta información deberá ser proporcionada por editores o bien, por los creadores de contenido, de tal forma que el contenido sea etiquetado de forma adecuada para facilitar el acceso a contenido relevante para docentes y estudiantes.
- Lrmi extiende el trabajo iniciado por Schema.org específicamente para materiales educativos. Schema.org es una actividad de colaboración con la misión de crear, mantener y promover esquemas para datos estructurados (vocabularios) en Internet, en las páginas web, mensajes de correo electrónico, etc. Los grupos de etiquetas que componen este esquema fueron desarrollados por un grupo de trabajo comprendido por editores comerciales, generadores de recursos educativos abiertos y especialistas educativos con el objetivo de facilitar a los profesores y estudiantes la búsqueda de material educativo preciso para un acto educativo específico. En cuanto madure la adopción del esquema Lrmi por parte de los motores de búsqueda, los creadores de recursos educativos que utilicen las etiquetas Lrmi para describir el contenido permitirán que el

Tendencias multidisciplinares...

trabajo para localizar el contenido adecuado sea más rápido, eficiente y efectivo.

Los metadatos se pueden utilizar para una serie de escenarios de e-learning. Cada escenario de uso puede abordar diferentes dimensiones en términos de Sampson, Papaioannou, Karadimitriou (2002)

- Tipo de recursos educativos (digital, no digital).
- Categoría de usuarios que acceden o puedan enviar material educativo.
- Categoría de los usuarios de la edición o la gestión de los metadatos educativo de los recursos educativos.
- Estándar de metadatos para la educación / especificaciones utilizadas para describir los recursos educativos (es decir, IEEE LOM, Dublin Core, etc.).
- Las tecnologías utilizadas para representar el documento de metadatos (XML, RDF, HTML, etc.).
- Las herramientas de software utilizadas para crear el documento de metadatos educativos.

De todos los recursos educativos descritos que fueron desarrollados por la UNAM, hay todavía muchos de ellos que pueden ser abiertos para beneficio de la sociedad. Para ello es necesario que cuenten con un esquema de metadatos que les dé visibilidad y los haga localizables para los usuarios potenciales. Es por ello que se torna primordial la selección e implementación de un esquema de metadatos para los recursos educativos a nivel institucional. En este sentido, la CUAED propone y se orienta hacia el esquema LRMI, ya que ofrece compatibilidad con los principales motores de búsqueda en Internet, lo que incrementaría la visibilidad y el acceso a sus recursos para beneficio de la comunidad universitaria y de la sociedad.

CONCLUSIONES

- Existen múltiples esquemas de metadatos, lo cual complica la labor de la selección que tienen que realizar las instituciones educativas para definir su propia organización y catalogación de recursos educativos.
- No existe una política o línea institucional en la universidad que defina el o los esquemas de metadatos que deben ser aplicados a los recursos educativos que en ella se generan.
- La implementación de un esquema de metadatos para recursos educativos requiere de gente especializada para catalogar, que por lo general es difícil encontrar, o bien, a mucha gente le parece una actividad poco atractiva.
- Con el incremento del uso del Internet en nuestra sociedad, y de la gran cantidad de contenidos de todo tipo que se publican diariamente en línea, es una necesidad permanente que los recursos educativos digitales de calidad cuenten con la máxima visibilidad apoyados bajo un esquema de metadatos completo y robusto.
- A nivel institucional, se requiere establecer una cultura integral de catalogación que favorezca la recuperación de contenidos educativos que generan los académicos y los alumnos de manera ágil y pertinente.

REFERENCIAS BIBLIOGRÁFICAS

- CEO FORUM (2000). Year 3 Report. School Technology and Readiness (a Focus on Digital Learning). Washington, D.C.: CEO Forum on Education and Technology. Disponible en <http://www.ceoforum.org/reports.html>
- Fox, C. *et al.* (2013). Transforming Data to Information in Service of Learning. State Educational Technology Directors Association (SETDA). Washington, D. C. Disponible en <http://www.setda.org/wp-content/uploads/2013/11/Data-to-Information.pdf>.

Tendencias multidisciplinares...

- García Aretio, L. (2014). *Bases, mediaciones y futuro de la educación a distancia en la sociedad digital*. Madrid: Síntesis.
- García, F. G. (2005). Contenidos educativos digitales: Construyendo la Sociedad del Conocimiento. *Red digital: Revista de Tecnologías de la Información y Comunicación Educativas*, 6 (1). Disponible en http://reddigital.cnice.mec.es/6/Articulos/articulo_resumen.php?articulo=1.
- Gértrudix, M., S. Álvarez *et al.* (2007). Acciones de diseño y desarrollo de objetos educativos digitales: programas institucionales en Monográfico Contenidos educativos en abierto. *Revista de Universidad y Sociedad del Conocimiento*, 1, (4). Disponible el 20 de marzo de 2013 <http://www.uoc.edu/rusc/4/1/dt/esp/monografico.pdf>.
- Martínez, J. M. (2003). Contenidos en e-learning: El rey sin corona. *RIED. Revista iberoamericana de educación a distancia*, 5(2). Disponible en: http://www.inacap.cl/tportal/portales/tp4964b0e1bk102/uploadImg/File/TeoriaPractEns/EnsenanzaTecnologia/3_ContenidosE_learning.pdf.
- Orr, D., M. Rimini y D. Van Damme. (2015), *Open Educational Resources: A Catalyst for Innovation, Educational Research and Innovation*. París: OECD. Disponible en <http://open-educational-resources.de/wp-content/uploads/sites/4/2015/11/OECD-OER-a-catalyst-for-innovation-2015.pdf>.
- Sampson, D. (2001). Current and Future Research and Technology Developments in e-Learning. Discurso en la 2 Conferencia Internacional de nuevos Horizontes en la Industria y la Educación. Milos, Grecia, 13-14 de septiembre de 2001.

- Sampson, D., V. Papaioannou y P. Karadimitriou P. (2002) EM2: an Educational Metadata Management Tool, *Educational Technology & Society Journal of International Forum of Educational Technology & Society and IEEE Computer Society Learning Technology Task Force*, Special Issue on Innovations in Learning Technologies, 5(4).
- Solomou, G., C. Pierrakeas, A. Kameas. (2015) Characterization of Educational Resources in e-Learning Systems Using an Educational Metadata Profile. *Educational Technology & Society*, 18 (4): 246-260.

Competencia en metadatos como parte del proceso educativo en educación a distancia

MARCO BRANDÃO

Universidad Nacional Autónoma de México

INTRODUCCIÓN

Los metadatos digitales son uno de los principales recursos para facilitar la comprensión e interoperabilidad de los componentes digitales que requieren ciertas actividades, en particular aquellas de Educación a Distancia (EAD), cuyo espacio de realización depende hoy en gran medida de los Ambientes virtuales de Aprendizaje (AVA).

Los AVA son softwares para el montaje de cursos a través de Internet que median entre profesores y estudiantes en la gestión de contenidos y administración de clases, con interactividad y monitoreo del desarrollo de las partes involucradas. En EAD hoy los AVA son fundamentales. Pero la EAD cuenta con distintas tecnologías e interpretaciones. Sin embargo, dos de sus propósitos, nos señala Nunes (1993-1994), han adquirido notorio destaque desde los años 1960 de eficiencia, eficacia y calidad en la universalización de la enseñanza y de ser un medio apropiado para la actualización de los conocimientos.

En países de dimensiones continentales y con asimetrías de desarrollo como Brasil, no tardó en verse como un recurso interesante para la Educación. El Instituto RadioMonitor, en 1939, y el Instituto

Tendencias multidisciplinares...

Universal Brasileiro, en 1941, marcaron en ese país el inicio de la EAD. Diez años después, el Gobierno de la República la consideró importante para las políticas públicas de educación y creó el Movimento de Educação de Base (MEB), cuya principal preocupación era la alfabetización y el apoyo a la educación de jóvenes y adultos (Nunes, 1994). El MEB se destacó por el enfoque en la Educación Formal y por el uso de la Radio como tecnología, mientras que las dos primeras priorizaron la enseñanza profesional y el uso del sistema postal.

En este país, términos como “educação a distância”, “ensino a distância” y “teleeducação”, el último empleado desde el uso de la televisión como tecnología en los años 1970, se refieren a la misma modalidad de enseñanza, aunque actuando con diferentes metodologías y tecnologías. Hay un intento de distinguir la EAD de “educación abierta”, pero siguen parecidos los conceptos, incluso Cirigliano (1983, p. 11) observa que la educación abierta puede ser presencial o a distancia, lo que la diferencia de la educación tradicional es que cualquiera puede unirse, independientemente de la escolaridad anterior.

La tecnología ha establecido las metodologías y los comportamientos en EAD. Con el uso de las computadoras desde los años ochenta en el mundo y desde los noventa en Brasil, esto se hace más evidente sobre todo con el cambio de la dinámica en la obtención de información y conocimiento, antes centrados en el libro, el profesor, la biblioteca y la escuela.

La creación de la Universidade Aberta do Brasil (UAB), en 1992, representó un hito para los dos momentos de la EAD brasileña. Creada en el contexto de la expansión de la Internet, la UAB estableció la Educación Superior a distancia con un modelo de enseñanza que está adoptado de manera generalizada en el país.

El país ya tenía iniciativas de EAD con el uso de la Internet, incluso de instituciones de educación pública. Una de las principales de este periodo fue de la Universidade Federal do Rio Grande do Sul (UFRGS) en conjunto con la Universidade Estadual de Campinas (UNICAMP) al desarrollar el primer AVA brasileño, el TelEduc, e implementar a gran escala la enseñanza del posgrado con esta herramienta.

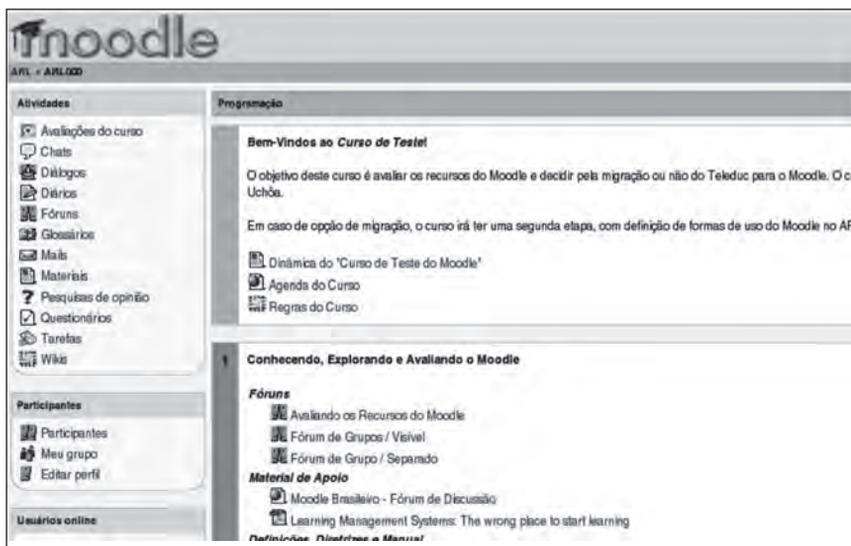
Figura 1. TelEduc.



Fuente: Home, Teleduc.org.br, 2016.

En competencia con otros AVA, el TelEduc ya incorporaba elementos adoptados como la estructura de enseñanza de la UAB y era una opción de AVA que el país comenzó a desarrollar entre otros. Sin embargo, a principios de los años 2000, la elección por la UAB del AVA creado por Martin Dougiamas en la Universidad Tecnológica de Curtin, el Moodle, trajo efectos más allá de la simple elección de esta herramienta en detrimento de la iniciativa brasileña y otras existentes.

Figura 2. Moodle



Fuente: Moodle (2016).

Desde de esa experiencia brasileña, se presenta un entendimiento de cómo las competencias necesarias para docentes y estudiantes en el proceso educativo en EAD dependen de la comprensión de los metadatos para la eficiencia, eficacia y calidad en esa modalidad de enseñanza.

COMPORTAMIENTO EN LA EAD

Como perspectiva de eficiencia, eficacia y calidad de la enseñanza, la EAD rompió con los tradicionales espacios de la escuela y con las antiguas formas de educación. Esto requiere un cambio de comportamiento del profesor —antes la referencia y el centro del proceso de enseñanza— y de los alumnos —que se son el centro del proceso de aprendizaje—.

Los modelos tradicionales de la escuela, aunque hayan evolucionado en cuanto a metodologías y conceptos, siguen manteniendo al profesor como centro del proceso de enseñanza y al estudiante como un eterno agente pasivo de su aprendizaje. Además, la existencia de espacios físicos, horarios y recursos para el aprendizaje, refuerzan la idea de que todos aprenden del mismo modo, al mismo tiempo y en el mismo espacio.

El espacio y el tiempo para enseñar eran determinados. “Ir a la escuela” representaba un movimiento, un cambio a la institución designada para la tarea de enseñanza y aprendizaje. El “tiempo de la escuela”, también determinado, era considerado como el tiempo diario que, tradicionalmente, el hombre dedicaba a su aprendizaje sistemático. Correspondía, también, en la historia de su vida, al tiempo que el hombre se dedicaba a la educación (Kenski, 2004, p. 24) [Traducción libre].

Hoy, la tecnología afecta en forma y contenido el *qué* y el *cómo* se enseña y el *qué* y el *cómo* se aprende. La información necesaria para la construcción del conocimiento está en una multitud de formatos y en su formato digital, posee una forma de presentación antes nunca posible. El primero de ellos, apunta Virilio (1993, *apud* Kenski, 2004, p. 24), es el espacio físico, en tiempo real; el segundo, es la velocidad con la que se produce:

Velocidad. Éste es el resumen del estatus espacio-temporal del conocimiento hoy en día. Velocidad para aprender y velocidad para olvidar. Velocidad para acceder a las informaciones, interactuar con ellas y superarlas con otras innovaciones. Esta “explosión de la información”, como dice Galvão, por un lado, se presta “para crear medios cada vez más eficientes para el almacenamiento y el flujo instantáneo de informaciones; por otra, desarrolla softwares y programas de búsqueda y filtro que nos ayudan a gestionar un espacio que ya está en las fronteras del infinito” (Kenski, 2004, p. 24) [Traducción libre].

El software aparece como una herramienta con la cual se presenta el potencial de la información digital. En este particular, la EAD

Tendencias multidisciplinares...

se sirve de los AVA para establecer su eficiencia, eficacia y calidad. Más que una herramienta, este tipo de software consolida una metodología de enseñanza y aprendizaje: mientras respecta la subjetividad en relación al tiempo y al espacio de aprendizaje, exige requisitos de conformidad, de calidad y plazos de las actividades; conserva el registro de la información de los materiales didácticos, de las actividades (libros, artículos, textos, evaluaciones, etc.) y de todos los cambios interactivos entre profesores y alumnos, profesores y profesores y alumnos con alumnos; son multimedia, ofrecen recursos de sonido, imagen y texto auxiliares al proceso educativo; son de extrema portabilidad (se puede acceder a ellos desde con una PC hasta con smartphones); también en un ambiente cerrado (interno) son capaces de agregar información y recursos externos de la Internet, etcétera.

Además, al retirar al profesor del espacio de la escuela, del libro y de la biblioteca, la exclusividad en la obtención de conocimiento transfiere un control al estudiante de su proceso de aprendizaje en la medida en que él pasa a actuar en el proceso de enseñanza.

Este control por parte de los estudiantes de su aprendizaje puede promover la autonomía, el interés, etc. y aprender a trabajar con comportamientos no deseados, por ejemplo, al entender que siempre prevalecerán las condiciones subjetivas o que cualquier forma de expresión, idea o actividad es válida en el ciberespacio. Otra posibilidad es que esta libertad genere un caos al proceso educativo de tal manera que pongan los objetivos de la EAD en peligro.

Sin embargo, contrariamente a las opiniones más conservadoras sobre el proceso educativo en EAD, en ese modelo de educación es necesario un padrón de comportamiento para convivir y acompañar los cursos. Al hacerlo, se tiene en cuenta la exigencia al estudiante de su autonomía e independencia para autodisciplinarse, por medio del profesor en el ejercicio de su autoridad en las concepciones tradicionales de educación. En otras palabras, la libertad hecha posible por el ciberespacio requiere conciencia y el compromiso personal de cada uno.

En el ciberespacio, esta unión de ciudadanos conectados, agrupados de modo virtual en torno de intereses específicos, pueden construir una comunidad desde el momento en que son establecidas reglas, valores, límites y costumbres, la *Netiqueta*, con la restricciones y los sentimientos de acogimiento y “pertenencia” al grupo (Kenski, 2004, p. 106) [Traducción libre.]

Es decir, lo que Kenski llama *Netiqueta* son elementos de un comportamiento necesario para el éxito en EAD. Son elementos que son potenciados en los AVA porque es un movimiento que se hace de adentro hacia afuera del individuo. Es decir, es parte de los recursos de la EAD, formar el perfil de estudiante y el manejo de la información a través de funciones multimedia marcadas por un “comportamiento” hecho posible por los metadatos digitales.

Este perfil de estudiante EAD funciona como un “metadato del estudiante” y exige otra conducta del profesor: él no necesita desarrollar actitudes y se centra en el conocimiento, en la articulación y en la evaluación de lo que el estudiante produce en un espacio meta informativo. Así es que esta comprensión de educación incorpora formas lógicas con las cuales operan los espacios virtuales. En un AVA, por lo tanto, se requiere más que las habilidades básicas de lectura, escritura, conteo o de manejo de las tecnologías informáticas: se requiere apropiarse de esta lógica multimedia y de los metadatos.

LA LÓGICA DE LOS METADATOS

Hay muchos recursos de software para la realización de tareas. En EAD se tiene igual diversidad de herramientas y formatos que en AVA. Esto nos lleva a creer que no hay restricciones de herramientas.

Sin restricciones, el comportamiento de profesores y alumnos en EAD se basa en la lógica del ciberespacio, es decir, en cómo está estructurada la información digital. Ella es *hipertextual*, multimedia e interactiva por cuenta de los metadatos.

Tendencias multidisciplinares...

Los metadatos son anteriores a la web. Su creación se atribuye a Jack E. Myers, y fueron empleados más a menudo en los años 1980 (The Linux Information Project, 2006). Hoy su uso más convencional sirve a la descripción de recursos de Internet para llevar a cabo una variedad de formas (texto, sonido, imagen etc.), por medio de la carga de información como tipo de archivo, fecha y lugar de creación, edición, localización etc., lo que extiende el conocimiento. Por ello son considerados “la información sobre la información”.

Los metadatos digitales facilitan la comprensión e interoperabilidad en la web. Como no hay restricciones de herramientas, son manejados por distintos softwares de manera que en los AVA la información debe promover un comportamiento de docentes y alumnos con esta lógica meta informativa y en esta manera de uso de la información. Esta competencia implica normas de comportamiento en los AVA.

Las normas de comportamiento deben estar presentes en cualquier actividad educacional. En EAD, el estudiante es el agente principal de su aprendizaje y es responsable por sus acciones, las cuales estarán registradas con las actividades enviadas, manifestaciones —o ausencia de manifestaciones— en el AVA, es decir, su comportamiento también es leído en el proceso educativo como un metadato. Independientemente de las herramientas, esta lógica de metadatos opera ese proceso.

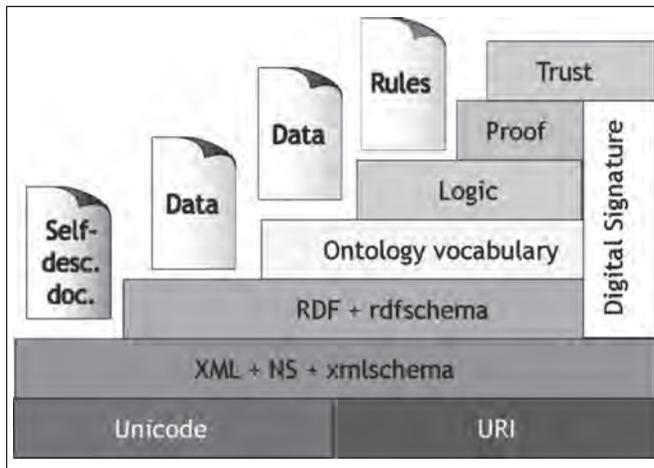
Los metadatos son presentados de diferentes maneras: como Objeto de Aprendizaje (OA) (Wiley, 2001), que son padrones para la creación de recursos virtuales utilizados por cualquier persona, accesibles, interoperables etc.; a través de la web semántica, que son tecnologías para publicar datos legibles por aplicaciones informáticas, cuyas informaciones adicionales describen el contenido, el significado y la relación de los datos para que máquinas de procesamiento puedan evaluarlas automáticamente y mejorar la interoperabilidad entre los sistemas sin operadores humanos (Castells, 2003).

Figura 3. Los Objetos de Aprendizaje



Fuente: Wiley, 2001.

Figura 4. La web semántica



Fuente: Lapuente, 2001.

Tendencias multidisciplinares...

Otra forma de entender los metadatos en EAD es a través del Sistema de Aprendizaje:

Se refiere al CON QUÉ se va a aprender. Estos sistemas permiten llevar el seguimiento del aprendizaje de los alumnos teniendo la posibilidad de estar al tanto de los avances y necesidades de cada uno de ellos. Cuentan con herramientas para colaborar y comunicarse (foros, chats, videoconferencia y grupos de discusión, entre otros), y tener acceso a recursos de apoyo como artículos en línea, bases de datos, catálogos, etc. Asimismo, hacen posible acercar los contenidos a los alumnos para facilitar, mostrar, atraer y provocar su participación constante y productiva sin olvidar las funciones necesarias para la gestión de los alumnos como la inscripción, seguimiento y la evaluación (Elementos de un Ambiente Virtual de Aprendizaje, 2016).

Estos sistemas incluyen los usuarios, el currículo y los profesionales a través de un espacio virtual al que se puede acceder de cualquier dispositivo conectado a la Internet.

Figura 5. El Sistema de administración de aprendizaje



Fuente: Elementos de un Ambiente Virtual de Aprendizaje, 2016.

El AVA tiene un papel clave para estas funciones, razón por la cual la EAD ha establecido su formato y exigido competencias para docentes y estudiantes en la comprensión de ese proceso. Antes, cuando el profesor era la estructura central del proceso, el estudiante “aprendía” el contenido. Ahora, se incorporan actividades más dinámicas como acompañar y establecer un control diario del desarrollo del alumno, de estrategias para hacerse presente en el AVA, etc. y un sistema en el cual el aprendizaje es estructurado y asistido por diferentes profesionales.

EL SISTEMA DE ADMINISTRACIÓN DE APRENDIZAJE DE LA EAD-BRASIL

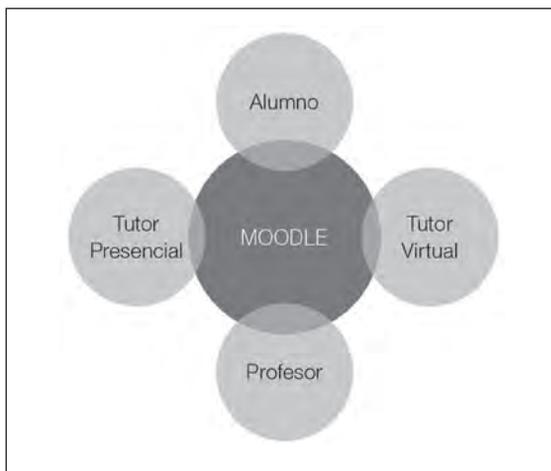
A pesar de las posibilidades de las tecnologías digitales —especialmente la Internet—, las experiencias nos hacen reflexionar sobre lo que representa la adopción de modelos basados en las herramientas disponibles. En Informática, ocurre un antagonismo observado en pocas áreas de la actividad humana: a pesar de la diversidad de software y aplicaciones para actividades disponibles, personas se ocupan en conocer y manejar solamente una, de la cual son dependientes para realizar sus tareas. Un ejemplo es el sistema operativo y sus aplicaciones básicas más usadas en el mundo. Pero podemos ver esto también en la EAD.

Brasil, país que he implementado la EAD con diferentes tecnologías y que hoy sigue con cursos consolidados, principalmente universitarios (licenciatura, posgrados) en instituciones públicas y privadas, se convierte en un escenario de lo que se cree ver en otras partes del mundo.

La implementación de EAD con los AVA, sobre todo aquellos desarrollados en las universidades brasileñas, como el TelEduc, podría representar de muchas maneras la expansión de metodologías y tecnologías de enseñanza. Sin embargo, la adopción generalizada de la tecnología Moodle representó la imposición de una sola metodología en EAD.

En general, ésta es la estructura de cursos EAD en Brasil:

Figura 6. Estructura de Cursos EAD en Brasil



Fuente: elaboración propia.

Lejos de discutir aquí la eficacia de este sistema de administración de aprendizaje, lo que emerge es la réplica de este modelo por las instituciones educativas en los más diversos cursos. Tampoco se pretende discutir la calidad del software Moodle que, es seguro, se trata de una excelente aplicación. Pero ello generó una dependencia exclusiva de la EAD en cuanto metodología, aunque existan otras opciones, una vez más en informática vemos una paradoja y en la educación una contradicción.

Paradoja porque el Moodle es ahora considerado sinónimo de AVA, cuando debería ser uno entre otros softwares; contradicción porque la comprensión del proceso educativo en EAD es el entendimiento dado por el Moodle: un cambio de software traerá dificultades o problemas de provecho de cursos dado que el aprendizaje es adicto a la herramienta en lugar de un comportamiento lógico.

CONCLUSIÓN

De la experiencia brasileña, se extrae un entendimiento de que la lógica de los metadatos configura el proceso educativo en EAD, lo que es una competencia necesaria en los AVA para la eficiencia, eficacia y calidad de esta forma de enseñanza. La adopción de un solo software como AVA puede actuar sin desarrollar estos “comportamientos meta informativos”.

De hecho, el Moodle es una excelente opción de AVA y, usado en gran escala, permite ventajas como mejorar el pronóstico de presupuesto, falta de necesidad de módulos de capacitación para el uso del AVA, mejor intercambio entre las instituciones y cursos, etc. Sin embargo, eso lo hace actuar como un monopolio de software, crea una acentuada dependencia de la metodología EAD y una menor comprensión de los metadatos para permitir un comportamiento del profesor y de los alumnos basado en la lógica del ciberespacio, no en la lógica de un software.

Los AVA deben mediar entre profesores y estudiantes en la gestión de contenidos y administración del curso, con interactividad, de manera que permita el monitoreo constante del desarrollo de las partes involucradas y no para ser “el contenido”, es decir, la única manera de enseñanza con software.

En un AVA, se requiere la comprensión de un proceso educativo distinto basado en cómo la información se comporta en el ciberespacio y cómo profesores y alumnos, además de las competencias básicas de lectura, escritura, conteo y manejo de tecnologías informáticas, deben desarrollar e incorporar esta lógica digital en EAD. Con el ejemplo de la experiencia brasileña, se tiene un entendimiento de los metadatos como una competencia necesaria para suplantar la herramienta (software) por un proceso educativo más amplio y creativo.

REFERENCIAS BIBLIOGRÁFICAS

- Castells, P. (2003). La Web Semántica. *Pablo Castells's Publications*. Disponible el 7 de junio de 2016 en <http://arantxa.ii.uam.es/~castells/publications>.
- Cirigliano, G. F. J. (1983). *La educación abierta*. Buenos Aires: El Ateneo.
- Elementos de un Ambiente Virtual de Aprendizaje (2016). *Ambientes Virtuales de Aprendizaje*. Disponible el 19 de mayo de 2016 en <https://virtualclasskms.wikispaces.com/ElementosdeunambientevirtualdeAprendizaje>.
- Kenski, V. M. (2004). *Tecnologías e Ensino Presencial e a Distancia*. Campinas, SP: Papirus.
- Lapuente, M. J. Lamarca (2001). Hacia la web semántica. *Hipertexto: el nuevo concepto de Documento en la Cultura de la Imagen*. Disponible el 7 de junio de 2016 en http://www.hipertexto.info/documentos/web_semantica.htm.
- Metadata Definition (2006). *Metadata Definition by The Linux Information Project*. Disponible el 27 de mayo de 2016 en <http://www.linfo.org/metadata.html>.
- Moodle (2016). *Moodle- Open-source Learning Platform*. Disponible el 19 de mayo de 2016 en https://moodle.org/?lang=en_us.
- Nunes, Ivonio B. (1993-1994). Noções de Educação a Distância. *Revista Educação a Distância*, 4 (5):7-25. Disponible el 19 de mayo de 2016 em http://www.feg.unesp.br/~saad/zip/OqueeEducacaoaDistancia_Ivonio.htm.
- Objetos de Aprendizaje y ROA. (2016). *Emaze Presentations*. Disponible el 19 de mayo de 2016 en <https://www.emaze.com/@AIQROLTR/Objetos-de-aprendizaje-y-ROA>.
- Teleduc (2016). Página web de Teleduc. Disponible el 19 de mayo de 2106 en <http://www.teleduc.org.br>.

- The Linux Information Project. (2006). Disponible el 19 de mayo de 2016 en <http://www.linto.org/linuxdef.html>
- Wiley, D. (2001). *Connecting Learning Objects to Instructional Design Theory: a Definition, A metaphor, and a Taxonomy*. Utah: Utah State University Disponible el 19 de mayo de 2016. En <http://www.reusability.org/read>.

Tendencias multidisciplinares del uso de los metadatos. Instituto de Investigaciones Bibliotecológicas y de la Información/UNAM. La edición consta de cien ejemplares. Coordinación editorial, Carlos Ceballos Sosa; revisión especializada, Francisco Xavier González y Ortiz y Valeria Guzmán González; cotejo y corrección de pruebas, Valeria Guzmán González; formación editorial, Mario Ocampo Chávez. Fue impreso en papel cultural de 90 g. en los talleres AGYS Alevin, S.C., Retorno de Amores, No. 14, colonia Del Valle, C.P. 03100, delegación Benito Juárez, Ciudad de México. Se terminó de imprimir el mes de diciembre de 2017.