

Metadatos : necesidad e importancia de integrar estándares

Lic. Nelida Elba García
nellygarcia@nodoalem.com.ar

Asc. Sergio Caballero
sergiocaballero@gmail.com

Universidad Nacional de Misiones. Facultad de Humanidades y Ciencias Sociales.
Carrera de Bibliotecología
biblio@fhycs.unam.edu.ar
Tucumán 1946 – Posadas – Misiones Argentina

Resumen

En el contexto global y cambiante que ofrece Internet, los espacios de registro y almacenamiento de la información académica y científica como las bibliotecas digitales y los repositorios institucionales nos presentan el desafío de integrar herramientas a los objetos digitales, como los esquemas de metadatos, con el objeto de hacerlos visibles y dinámicos.

En el marco de la Iniciativa de Archivos Abiertos (OAI) y el protocolo para la Recolección de Metadatos (PMH) se recomienda la aplicación de metadatos Dublin Core a los recursos para aportar a la interoperabilidad entre sistemas de preservación. No obstante, admitidas sus limitaciones, se indica la apertura al uso de otros formatos adicionales para alcanzar una completa descripción por metadatos.

Los repositorios acrecientan su valor cuando evidencian impacto de sus contenidos, para lo que es fundamental ofrecer metadatos precisos y completos que permitan a los objetos depositados ser identificados, localizados y reutilizados creando servicios de valor agregado.

Este trabajo, elaborado en el marco de la investigación 16/H-213; FHyCS-UNaM: *“Iniciativas de acceso abierto para la conformación de repositorios institucionales : propuestas de implementación de un espacio de la UNaM en la web”*, pretende reflexionar acerca de la importancia y la necesidad de aplicar esquemas de metadatos que posibiliten la interoperabilidad de objetos digitales en espacios de preservación de la información académica y científica.

Palabras claves: metadatos –interoperabilidad – iniciativas de acceso abierto – OAI-PMH – repositorio institucional

Introducción

Como profesionales de la información somos protagonistas de la denominada “explosión digital” originada por la rápida evolución de la tecnología y el crecimiento de Internet con sus múltiples posibilidades de acceso y socialización de la información.

Este nuevo hito informacional, ha generado problemas y oportunidades a los protagonistas de la organización documental quienes deben abocarse a crear y adaptar herramientas para un tratamiento más eficiente de la información en el contexto digital.

Las nuevas representaciones del conocimiento –publicaciones electrónicas (textos, imágenes, sonidos)- han disminuido las expectativas de relevancia en la recuperación de la información. Teniendo en cuenta que una búsqueda en Internet aún arroja resultados excesivos, carentes de calidad y fuera de contexto, es preciso orientar los esfuerzos a la búsqueda de soluciones que sean aplicables y reusables a través de estándares que permitan administrar eficientemente la información.

En este contexto dinámico que ofrece la tecnología digital, ha surgido la Iniciativa de Archivos Abiertos (OAI)¹, como una alternativa para solucionar inconvenientes que presentan los sistemas de información inmersos, ya que, desarrolla y promueve la interoperabilidad de las normas que tienen por objeto facilitar la difusión eficaz de contenidos, constituyendo una herramienta que apuntala el Movimiento del Acceso Abierto impulsado por científicos y especialistas de la información con la finalidad de colocar al alcance de toda la sociedad los contenidos académicos y científicos.

Enmarcado en este paradigma, nuestro país ya cuenta con referentes que han alcanzado logros admirables dando impulso al desarrollo estandarizado de las bibliotecas digitales y los repositorios institucionales. La visibilidad de sus recursos preservados se va incrementando gradualmente porque crece también su capacidad de interoperar con otros sistemas de información digital.

“La interoperabilidad es la capacidad de un sistema o de un producto de trabajar con otros sistemas o productos sin un esfuerzo especial por parte del cliente. Desde este punto de vista computacional, la interoperabilidad permite generar un enlace entre sistemas de trabajo para las diferentes tecnologías de información promoviendo una sana convivencia y operatividad...” (Martínez Equihua, Saúl, 2007).

Un ejemplo de ello, es el protocolo Z39.50 que con sus usos actuales y potenciales, se considera una verdadera herramienta para vincular recursos distribuidos como los sistemas que administran los catálogos de las bibliotecas.

Son varios los estándares creados y en desarrollo para dar solución a los problemas relacionados con la interoperabilidad. Estos deben ser aplicables y reusables en forma amplia en espacios que albergan información digital, siendo algunos de ellos: Transporte y Comunicación (TCP/IP y Z39.50 ; Formatos de documentos (PDF y HTML) ; Metadatos (XML, MARC 21, Dublin Core, MODS, METS) ; Enlace de citas (OPENURL, CROSSREF, DOI) ; Autenticación (Proxis y firmas digitales).

En el marco de esta ponencia se hace referencia a la necesidad e importancia de considerar la aplicación de esquemas de metadatos en combinación con herramientas como el

¹ Open Archives Initiative: <http://www.openarchives.org/>

protocolo OAI-PMH² (Open Archives Initiative-Protocol for Metadata Harvesting) que permitan la interoperabilidad de recursos depositados en espacios que abogan la organización y preservación del contenido digital académico-científico como son los repositorios institucionales.

Metadatos e interoperabilidad

El concepto de metadatos tiene sus antecedentes en la catalogación tradicional; por eso, es común encontrarse con la afirmación de que una ficha catalográfica es un ejemplo ilustrativo del término metadatos y su alcance. Como señala Hillman³ los metadatos han estado presentes desde que los primeros bibliotecarios crearon listas de recursos de información. Caplan⁴, quien documenta el nacimiento del término en las Ciencias de la Computación, establece que el prefijo “meta” se utiliza para explicar o hablar “acerca de”; así, por ejemplo, un metalenguaje es un lenguaje utilizado para describir otros lenguajes.

El alcance más difundido que se observa del término lo presenta como “**datos sobre datos**” o “**datos referentes a datos**” para identificar archivos digitales de conjuntos de datos científicos, sociales y geoespaciales. Por lo que en Internet representan la información electrónica dispersa y representan a la descripción bibliográfica de recursos electrónicos. Más concretamente en las bibliotecas digitales y los repositorios institucionales el uso de metadatos comienza de manera incipiente a partir del surgimiento de algunos estándares de metadatos como el Dublin Core.

La evolución de la web exige que los metadatos alcancen nuevas áreas de aplicación en donde juegan un papel primordial, principalmente para localizar recursos; además, constituyen un conjunto de elementos cuya semántica, comúnmente aceptada, representa la información electrónica e implica su descripción de manera adjunta al recurso o almacenada en otro lugar. Así, los catálogos pueden ser considerados como metadatos que se almacenan en un registro a parte del documento (base de datos), mientras que en un documento electrónico los metadatos se introducen formando parte de él como lenguaje de marcado invisible.

En expresiones de Rosa San Segundo, *a diferencia de las fichas catalográficas, una de las características más importantes de los metadatos es su capacidad de relación o de establecer enlaces. De esta forma, se han hecho imprescindibles en la recuperación global de la información en Internet, porque se trata de indizar y clasificar inmensas cantidades de información de diversos tipos*⁵.

Si se toma en consideración otras de las múltiples definiciones existentes:

“son información estructurada que describen, explican, localizan, o de alguna manera facilitan la obtención, el uso o la administración de un recurso de información” (Amy Brand, 2003)⁶.

² <http://www.openarchives.org/OAI/openarchivesprotocol.html>

³ Hillman, D. (2003). *Using Dublin Core*. Recuperado el 10 de marzo de 2005 de: <http://dublincore.org/documents/usageguide/>.

⁴ Caplan, P. (2003). *Metadata for Education*. Chicago: American Library Association.

⁵ San Segundo MR. (2006). Organización del conocimiento en Internet : metadatos bibliotecarios DUBLIN CORE'. Disponible en URL: http://fesabid98.florida-uni.es/Comunicaciones/r_sansegundo.htm

⁶ Brand, Amy. Daly, Frank. Meyers, Barbara (2003). *Metadata Demystified: A Guide for Publishers*. En: http://www.niso.org/standards/resources/Metadata_Demystified.pdf

“son un conjunto de elementos que se utilizan para ayudar a la identificación, descripción y localización de los recursos electrónicos por medio de una representación de la descripción bibliográfica de los mismos” (Juárez Santamaría, 2007)⁷.

... se observa que los metadatos no sólo cumplen funciones de identificación sino que también pueden contener información con fines administrativos y estructurales; son datos que admiten operaciones con objetos; atributos que agregan valor (significado, contexto, organización) a los recursos de información (texto, imagen, audio, video, etc.); tratan de representar la información electrónica dispersa; facilitan el funcionamiento y la capacidad de compartir los datos entre los buscadores; elevan su precisión para efectivizar el proceso de búsqueda y recuperación en la Web, permitiendo a los usuarios, conocer los datos sobre los recursos de su interés; por lo que en el entorno digital su aplicación es creciente presentando un panorama cada vez más complejo siendo –según Ferrer Sapena y otros⁸- múltiples sus aplicaciones:

- **Catalogación:** porque pueden describir contenidos a nivel individual (como una página web) como en colecciones.
- **Búsqueda de recursos:** su aplicación facilita el trabajo de los motores de búsqueda.
- **Firmas digitales:** formatos como RDF permiten codificar las firmas digitales, requisito fundamental a la hora de crear una web de confianza.
- **Clasificación de contenidos:** permiten establecer los contenidos que los usuarios desean ver.
- **Otras aplicaciones:** marcados de políticas y preferencias de privacidad y los derechos relativos a la propiedad intelectual.

La aplicación de metadatos **no solo permite etiquetar objetos y recursos sino también calificar procesos o acciones**; por lo que hasta representan una amenaza para el desarrollo de la práctica habitual de la catalogación y el uso del formato MARC si se consideran sus ventajas frente a los tradicionales sistemas de catalogación. Según las afirmaciones de Ortiz-Repiso, estas ventajas están a la vista:

- Permiten establecer relaciones entre registros y ficheros diferentes
- Son más flexibles en la indización y presentación del texto
- Son menos complejos y permiten la realización de diferentes búsquedas (booleanas, truncamiento).

Tipología

Los tipos y funciones de metadatos existentes son múltiples, dependiendo de diversos factores: el tipo de información que describen, el nivel de estructuración de esta información,

⁷Juárez Santamaría, Beatriz (2007): “Uso de los metadatos en el orden documental. Información, producción y servicios”, Vol. 10, n.º 42. Citado en: Martínez Equihua, Saúl. Biblioteca digital: conceptos, recursos y estándares. Buenos Aires : Alfagrama.

⁸ Ferrer Sapena, Antonia.y otros (2005). Guía metodológica para la implementación de una biblioteca digital universitaria / Gijón : Trea. p. 107

el lugar donde se encuentren los metadatos, su ámbito de aplicación, el tipo de usuarios que los utiliza y sus finalidades, entre otros.

A los fines prácticos, los tipos y funciones de los metadatos pueden clasificarse en tres amplias categorías -descriptivos, estructurales y administrativos- con límites no siempre bien definidos y a veces superpuesto. Por ejemplo, los metadatos administrativos pueden incluir una amplia gama de información que podría considerarse como metadatos descriptivos y estructurales (Tim Berners-Lee, 2000)⁹.

Metadatos descriptivos: son aquellos que sirven para la descripción e identificación de los recursos de información, permiten la búsqueda y recuperación de la información, como también distinguir un recurso de otro y entender el asunto o contenido del mismo. Se realizan mediante los estándares como Dublin Core; MARC; Meta tags, HTML, etc.

Metadatos estructurales: son los que más influyen en la recuperación de la información electrónica, facilitan la navegación y presentación de los recursos electrónicos. Así, ofrecen la información sobre la estructura interna de los recursos, estableciendo las relaciones entre ellos, de manera que pueden incluso unir los archivos de imagen y textos que están relacionados. Los estándares más difundidos para ellos son SGML y XML/RDF; EAD (Encoded Archival Description-Descripción de Archivo Codificado para Archivos).

Metadatos administrativos: son de carácter más técnico porque incluyen datos sobre la creación y control de calidad, datos sobre la gestión de derechos, requisitos del control de acceso y utilización, información sobre la preservación y permiten la gestión a largo y corto plazo. Ejemplo de los metadatos que se incluyen aquí: tipo y modelo de escáner utilizado, resolución, limitaciones de reproducción, etc.

Además de esta clasificación por tipos y funciones, también se pueden analizar sus características según otros atributos, como por ejemplo:

- el **método de creación** de los metadatos que puede ser *manual* (una persona) o *automático* (una máquina),
- la **estructura** que puede ser *simple* o *estructurada*,
- la **semántica** que puede seguir un vocabulario *libre* o *controlado*, etc.

Esta proliferación de clasificaciones, responde a que los metadatos están en fase de construcción técnica y por eso no existe un consenso generalizado en su conceptualización o sobre los tipos de metadatos existentes (Daudinot Founier, I., 2006)¹⁰.

Creación y estructuración de metadatos

Muchos de los metadatos estructurales y administrativos básicos los provee el personal técnico que se encarga de la digitalización o creación de un objeto digital, o son creados a través de un proceso automatizado. En cambio, algunas categorías de metadatos, como los descriptivos, es preferible que sea el personal creador del recurso quien provea la información, en especial cuando se trata de conjuntos de datos científicos, donde el creador es quien tiene mayor comprensión del tema y de los usos que se le pueden dar a la información. Sin embargo, si los autores o creadores de los datos no tienen el tiempo o las habilidades

⁹ Berners-Lee, Tim. (2000). Semantic Web-XML. <http://www.w3.org/2000/Talks/1206-xml2k-tbl>

¹⁰ Daudinot Founier I. (2006). Organización y recuperación de información en Internet: teoría de los metadatos. ACIMED ;14 (5). Disponible en: http://bvs.sld.cu/revistas/aci/vol14_5_06/aci02506.htm

necesarias para hacerlo es preferible que sea el personal técnico quien se encargue de su creación consultando con especialistas para confirmar su consistencia y corrección.

En relación a la estructura de los metadatos, existe una distinción entre metadatos de formato simple –tales como los representados en la sintaxis de un lenguaje de marcas e insertados dentro de la estructura del objeto digital (XML; HTML; SGML) – y metadatos de formato muy estructurado que se idean, aplican y mantienen de acuerdo con normas (internacionales) claramente establecidas (Lynne C. Howarth, 2003)¹¹.

Los esquemas de metadatos son conjuntos de elementos de metadatos diseñados para un propósito específico. Según Bodoff, un esquema estándar es un conjunto de elementos, que propone un grupo u organismo reconocido, para describir distintas funciones en las aplicaciones web¹².

El significado de los elementos en sí, es conocido como la semántica del esquema, mientras que, los valores que le asignamos a los elementos de los metadatos son el contenido.

Generalmente, los esquemas de metadatos especifican los nombres de sus elementos y su semántica. Opcionalmente, pueden llegar a especificar reglas que determinen la manera en que el contenido debe ser formulado (por ejemplo, como identificar el título principal), reglas para su representación (por ejemplo, el uso de mayúsculas) y los valores que son permitidos (por ejemplo, términos que deben ser usados de un vocabulario específico).

También puede haber reglas de sintaxis que establecen la manera en que los elementos y su contenido deben ser codificados; un esquema de metadatos sin reglas de sintaxis preestablecidas es llamado independiente de sintaxis. Los metadatos pueden ser codificados en cualquier sintaxis definida; muchos esquemas de metadatos actuales utilizan SGML (Standard Generalized Mark-up Language) o XML (Extensible Mark-up Language).

Los metadatos pueden estar estructurados por diferentes organismos y entidades como conjunto de elementos previamente definidos para aplicaciones concretas; esto implica que el número de esquemas de metadatos propuestos sea múltiple y que en un entorno de trabajo donde diferentes entidades (clientes y servidores) deban interactuar probablemente no puedan hacerlo por trabajar con diferentes modelos (los elementos en los diferentes esquemas no se llaman igual, no significan lo mismo, no permiten los mismos valores). Por tal razón, es importante garantizar la interoperabilidad entre las entidades a nivel semántico para que en la coexistencia de diferentes esquemas de metadatos se produzca el intercambio que impactará positivamente en la relevancia de los resultados de búsqueda.

En una reunión conjunta del grupo de trabajo de metadatos llevada a cabo en Ottawa en Agosto de 2001 se establecieron los **principios considerados comunes** a todos los dominios de metadatos que podrían servir de base en el diseño de cualquier esquema de metadatos o aplicación:

- **Modularidad:** implica poder combinar de manera interoperable sintáctica y semánticamente elementos de datos, vocabularios y otros elementos constructivos provenientes de esquemas diferentes. Se trata de ensamblar esquemas de metadatos preestablecidos que incluyan la funcionalidad de cada componente y que reúnan los requisitos específicos de cada aplicación.
- **Extensibilidad:** los sistemas de los metadatos deben permitir extensiones para que puedan acomodarse a las necesidades particulares de una determinada aplicación. Por

¹¹ Modelos de Metadatos para Pasarelas Temáticas / Lynne C. Howarth. Facultad de Estudios de la Información. Universidad de Toronto. Toronto, Canadá En: http://ifla.queenslibrary.org/IV/ifla69/papers/053s_trans-Howarth.pdf

¹² Bodoff, D. Hung, P. C. K., Ben-Menachem, M.: Web Metadata Standards: Observations and Prescriptions. IEEE Software. Vol. 22, no. 1 (2005) 78–85

ejemplo, el concepto de “creador” o “identificador” de un recurso de información serán comunes a la mayoría de los esquemas de metadatos, mientras que otros serán específicos a aplicaciones particulares o dominio.

- **Refinamiento:** de acuerdo con el nivel de detalle necesario los diseñadores de estándares de metadatos podrán adicionar calificadores para refinar o hacer más específico el significado de un elemento. Por ejemplo, el uso de vocabularios controlados para hacer más precisa las descripciones.
- **Plurilingüismo:** las arquitecturas de metadatos adoptadas deberán respetar la diversidad lingüística y cultural. Se trata de adaptar estándares neutrales a contextos locales.

Aplicada a la creación y uso de los metadatos la integración de las características mencionadas **permite una estandarización para generar una interoperabilidad**, esto es, la habilidad de múltiples sistemas con diferente hardware, distintas plataformas de software, estructuras de datos e interfaces, para intercambiar datos con una mínima pérdida de contenido y funcionalidad.

Aquí, en un plano técnico funcional, cobra importancia el tema de la interoperabilidad así como la necesidad de protocolos que garanticen el intercambio de información digital.

Según Gómez Dueñas¹³, es posible *definir interoperabilidad como la capacidad que presenta un sistema de información de comunicarse y compartir información efectivamente con otro mediante una interconexión libre y transparente (compartir metadatos, documentos y objetos digitales). En términos del OAI se busca crear una metodología simple, con una forma de entendimiento rápida, que sea de fácil desarrollo y aplicación que a partir del uso extendido de los navegadores de Internet, se pueda añadir a los software de sistemas de información existentes, capacidades de interoperabilidad y planear la utilización de estas capacidades como medio de transporte de la información incluyendo el manejo del metalenguaje XML, los cuales se consideran actualmente como los modelos ideales para el manejo de información digital.*

Respecto al nivel de catalogación y modelo de metadatos a utilizar, el autor expresa que *OAI se basa inicialmente en el estándar Dublin Core, en formato simplificado (sin calificadores de campo), por ser este el modelo más fácil de implementar, el cual está diseñado para trabajar principalmente con información digital y poseer un modelo de implementación fundamentado en XML. No obstante, el diseño del protocolo permite especificar cualquier modelo de metadatos existente, además permite crear nuevos modelos de metadatos mientras estos estén formulados en XML.*

Las bibliotecas digitales y los repositorios institucionales representan etapas de desarrollo hacia la Web semántica o Web 3.0, cuyo fin es la fácil localización de recursos, la comunicación entre sistemas y programas que permitan una extraordinaria efectividad en la recuperación de la información.

La base para alcanzar ese objetivo de la web semántica está en la interoperabilidad que permiten los esquemas estándares de metadatos en combinación con otros encargados de su recolección y transferencia. Por lo tanto, la construcción de los espacios de preservación digital será más eficiente si se utilizan estructuras que posibiliten organizar, localizar,

¹³ Gómez Dueñas, Laureano Felipe. (2005) La Iniciativa de Archivos Abiertos (OAI) un nuevo paradigma en la comunicación científica y el intercambio de información. En Revista Códice. Universidad de La Salle, N° 4, 2005.

recuperar, procesar, evaluar, intercambiar y reutilizar un determinado dominio o tipo de recursos, garantizando su interoperabilidad en el entorno digital.

Elementos y modelos de metadatos

La determinación del alcance, la cobertura y el contenido de ese entorno digital determinarán qué elementos y modelos de metadatos se elegirán para dar soporte a la identificación de los recursos. En general, los tipos y estructuras de metadatos que podrían requerirse para organizar objetos digitales deben incluir: metadatos administrativos, metadatos descriptivos, metadatos analíticos, metadatos de gestión de derechos, metadatos estructurales o técnicos, otros según necesidad.

La elección de uno o más modelos de metadatos a emplear en la creación de los registros para identificar a los objetos digitales dependerá del objetivo del servicio y los tipos de metadatos que contenga.

Iniciativas que impulsan la interoperabilidad

Existen varios estándares de metadatos para describir distintas funciones. En particular, para describir recursos se ha propagado el uso del esquema Dublin Core, que también desarrolla una línea de aplicación para la anotación específica de recursos educativos y su interoperabilidad con estándares de metadatos del sector educativo: **IEEE-LOM** (Learning Object Metadata), Metadatos para Objetos de Aprendizaje¹⁴; **RDF** (Resource Description Framework), Infraestructura para la Descripción de Recursos¹⁵; **Open Citation Project –Citebase**¹⁶; **Paracite**¹⁷; **EAD** (Encoded Archival Description)¹⁸; **UK Learning Object Metadata Core**¹⁹; **ARIADNE**²⁰; **Learning Federation Metadata Application Profile**²¹; **DC-educational**²², entre otros.

En el contexto del Movimiento del Acceso Abierto (OA) y la Iniciativa de Archivos Abiertos (OAI), uno de los estándares que estriba en los conceptos de interoperabilidad más utilizados es el Protocolo para la Recolección de Metadatos **OAI-PMH** (Open Archives Initiative Protocol for Metadata Harvesting), que será abordado con mayores detalles por encontrarse esta ponencia enmarcada en un proyecto de investigación que también aboga las iniciativas de libre acceso a la información.

En Internet, un protocolo es un método establecido para intercambiar datos. Con la finalidad de facilitar el intercambio de datos entre los distintos servidores nace en la Convención de Santa Fe en 1999 el protocolo OAI-PMH (Open Archives Initiative – Protocol

¹⁴ Berlanga, Adriana J., López, Clara, Morales, Erla, García, Francisco J. (2005). Consideraciones para reforzar el valor de los metadatos en los objetos de aprendizaje (OA). Salamanca : Universidad de Salamanca. Departamento de Informática y Automática e Instituto Universitario de Ciencias de la Educación.

¹⁵ Recomendación del W3C 22 febrero 1999 (español) traducida por Eva Méndez. En:

<http://www.sidar.org/recur/desdi/traduc/es/rdfrdfesp.htm>

¹⁶ Hitchcock et al, 2002. En: <http://citebase.eprints.org>

¹⁷ En: <http://paracite.eprints.org/>

¹⁸ Peis, Eduardo y Ruiz-Rodríguez, Antonio A.8 *EAD (Encoded Archival Description): Desarrollo, estructura, uso y aplicaciones* [on line]. "Hipertext.net", núm. 2, 2004. <<http://www.hipertext.net/>> [Consulta: 22/09/2009]. ISSN 1695-5498

¹⁹ UK LOM: UK Learning Object Metadata Core. Draft 0.2.

²⁰ ARIADNE: Foundation for the European Knowledge Pool. <http://www.ariadne-eu.org> (2004)

²¹ The Learning Federation: Metadata Application Profile v. 1.3.

http://www.thelearningfederation.edu.au/repo/cms2/tlf/published/8519/Metadata_Application_Profile_1_3.pdf (2003)

²² Education Working Group. DC-Education Application Profile. http://www.ischool.washington.edu/sasutton/dcmi/ed/04-05/DC-Education_AP_06-20-05.html (2005)

for Metadata Harvesting) que podía implementarse sobre cualquier sistema que requiriese la comunicación de metadatos. Esto hizo afirmar, a algunos de los creadores del mismo, que OAI-PMH estaba llamado ser a las bibliotecas digitales lo que HTTP había sido a la web²³.

El Protocolo para la Recolección de Metadatos **OAI-PMH**, es una iniciativa apoyada financieramente por la Federación de Bibliotecas Digitales (DLB) y la Liga para la Información en Red (CNI) y la Fundación Nacional para el Desarrollo de la Ciencia (NSF), que desarrollan y promueven la aplicación de estándares de interoperabilidad para facilitar la eficiente diseminación de contenidos (Gómez y Arias, 2002)²⁴.

Se desarrolló como una alternativa que busca promover el libre acceso a la información científica de manera ordenada y normalizada mediante la implantación en los repositorios institucionales de documentos y bibliotecas digitales existentes en el mundo, de unas metodologías comunes que persiguen estandarizar la interoperabilidad entre éstos, con el fin de facilitar de forma eficiente y transparente la diseminación de contenidos mediante el desarrollo e implantación de una metodología y protocolos sencillos y prácticos que permiten una rápida aplicación de las normas y estándares internacionales relacionados con manejo de información digital (Lynch, 2003)²⁵.

Este protocolo orientado a la recolección de metadatos, se trata de una simple interfaz que permite el acceso a los metadatos de contenidos o recursos en formato XML provenientes de distintas fuentes, plataformas y repositorios; se encarga exclusivamente de la comunicación de metadatos, no de los textos completos de los documentos que se referencian; requiere adhesión a los metadatos Dublin Core, al cual se adapta porque es muy sencillo, flexible y puede ser soportado por la mayoría de los proveedores de datos y servicios; pero permite y apoya otros formatos de metadatos.

Existen dos clases de participantes en el marco de **OAI-PMH**:

- **Proveedores de datos:** que son los propios archivos de depósitos de documentos que proporcionan los metadatos relacionados con los documentos que almacenan.
- **Proveedores de servicios:** recopilan los metadatos de diferentes archivos para proporcionar servicios de valor añadido a los usuarios finales: sistema de búsqueda e identificación, filtrado, alertas temáticas, medición del uso e impacto de los documentos, etc.

Cabe destacar que el protocolo OAI-PMH:

- Genera y promueve estándares de interoperabilidad que facilitan la difusión, intercambio y accesibilidad a documentos de diferente naturaleza,
- Permite almacenar los metadatos en un solo punto de la red donde se realizan las diferentes consultas.
- Se ocupa de la gestión de la información, por lo que no define la creación de los metadatos, ni da los parámetros para realizar una consulta.

²³ SEDIC Asociación Española de Documentación e Información: http://www.sedic.es/autoformacion/acceso_abierto/4-iniciativa-ficheros-abiertos.html

²⁴ Gómez, ND. y Arias, OM. (2002). El cambio de paradigma en la comunicación científica. Información, cultura y sociedad, <http://eprints.rclis.org/archive/00003649/>

²⁵ Lynch, Clifford A. (2003). "Institutional repositories: essential infrastructure for scholarship in the digital age". *ARL bimonthly report*, no. 226 (Febr.).

- Utiliza transacciones HTTP (Hyper Text Transfer Protocol) utilizado en la transferencia de contenidos web, donde está definida la sintaxis y la semántica que utilizan los clientes y servidores para comunicarse entre sí.
- Está basado en un modelo cliente/servidor que transmite preguntas y respuestas entre un proveedor de datos y un proveedor de servicios.

Ejemplificación del funcionamiento del protocolo OAI-PMH:

- Cuando un usuario realiza una búsqueda en un servidor web, envía una solicitud de información a un proveedor de servicios (servidor web).
- El proveedor de servicios, solicita a un proveedor de datos (sistema de información digital, biblioteca digital, repositorio institucional, bases de datos, etc.) que le envíe registros de metadatos de los diferentes recursos que dispone.
- La respuesta del proveedor de datos llega al proveedor de servicios como un conjunto de registros de metadatos en formato XML.
- El proveedor de servicios enviará los resultados al servidor web que efectuó la consulta para que el usuario, por medio de una interfaz y a través de las URLs, llegue al documento.

Explica Arencibia²⁶ que la ventaja de la búsqueda reside en que el usuario selecciona los registros de su interés a través de los metadatos que los describen. Es decir, que el usuario no utiliza el protocolo OAI-PMH sino que lo hacen los proveedores de datos y los proveedores de servicios para comunicarse en función de responder al usuario su demanda. Precisamente, es el proveedor de servicio el que se conecta con varios proveedores de datos a través del protocolo OAI-PMH para que estos hagan disponibles la información que requiere el usuario.

Consideraciones actuales y locales acerca de los metadatos y la interoperabilidad

La razón de ser de los metadatos radica en facilitar el descubrimiento de información relevante; actualmente el terreno propicio es el de las bibliotecas digitales y los repositorios institucionales, espacios educativos y académicos para los que deben ser creados y aplicados los metadatos con el objeto de organizar la información digital dotándola de la identificación adecuada, facilitando la interoperabilidad y la integración de recursos para alcanzar la preservación.

En la web, existen una diversidad de Directorios y Servidores que se ocupan de indizar las iniciativas OAI, entre las más conocidas se pueden destacar:

- Open Archives Initiative, <http://www.openarchives.org/>
- Proveedor de Datos Registrado – Registered Data Providers, <http://www.openarchives.org/Register/BrowseSites>
- OpenDOAR, <http://www.opendoar.org/countrylist.php>
- ROAR, <http://archives.eprints.org/index.php?action=browse>
- RSP Blog Directory, <http://rsp.ac.uk/blogs/>
- Repositorios dinámicos, <http://repositoriosdinamicos.wordpress.com/category/acceso-abierto/>

Los espacios de preservación de la información académica y científica de nuestro país (repositorios institucionales argentinos, etc.) pueden ser visualizados en algunas de las

²⁶ Arencibia, Jorge Ricardo (2006). Las iniciativas para el acceso abierto a la información científica en el contexto de la web semántica. En: *Biblos*, vol. 7:25-26

herramientas de indización mencionadas como OpenDOAR y ROAR donde pueden ser observados también los detalles técnicos que caracterizan a cada uno (software utilizado, modalidades de acceso, inscripción a la iniciativa OAI, etc.).

Actualmente suman 7 (siete) y se espera que en un breve plazo se incremente dicha cantidad, lo que evidencia en Argentina una conciente comprensión de la necesidad y la importancia de sumar iniciativas que den sentido a los esfuerzos de producción intelectual académica y científica.

La Universidad Nacional de Misiones, conserva pendiente el planteo de una política formal relacionada con la implementación de un repositorio institucional y biblioteca digital, así como la reglamentación suficiente para atender todas las dimensiones del problema (jurídica y legal, tecnológica, institucional y social).

Según algunos aspectos del diagnóstico²⁷, cuya evaluación se halla en proceso, se pueden destacar algunas situaciones relativas al estado de la UNaM con vistas a un proyecto de digitalización:

<i>Algunos aspectos observados de una evaluación en proceso</i>	
Información digitalizada en Internet	<ul style="list-style-type: none"> ● Reglamentos ● Planes de estudios ● Disposiciones, Resoluciones, etc. ● Artículos de divulgación ● Material didáctico ● Catálogo de bibliotecas
Espacios de depósito de la información digital	<ul style="list-style-type: none"> ● Sitios web ● Aulas virtuales ● Blogs ● Catálogos en línea
Herramientas utilizadas para crear los documentos (programa/software)	<ul style="list-style-type: none"> ● MS Word ● MS Excel ● MS PowerPoint
Formatos utilizados para publicar el/los documento/s digitalizado/s	<ul style="list-style-type: none"> ● PDF ● DOC ● HTML ● XLS
Sistemas de Gestión/Softwares/Formatos utilizados para publicar recursos digitalizados	<ul style="list-style-type: none"> ● Sistema KOHA ● Sistema Aguapey ● Software WinIsis ● Softwares propietarios ● Formato MARC (Aguapey y Koha) ● Formatos propietarios

En general, la información que albergan los sitios webs de las distintas unidades académicas de la UNaM se hallan desprovistas de la codificación con metadatos estructurados y, no se avizoran previsiones técnicas, al corto y al mediano plazo, que luego faciliten la construcción de un repositorio institucional según estándares.

Desde el proyecto de Investigación 16/H-213; FHyCS-UNaM: *“Iniciativas de acceso abierto para la conformación de repositorios institucionales : propuestas de implementación*

²⁷ Realizado mediante encuestas en el marco de la Investigación *“Iniciativas de acceso abierto para la conformación de repositorios institucionales : propuestas de implementación de un espacio de la UNaM en la web”* 2008-2009.

de un espacio de la UNaM en la web”, que da lugar a esta presentación, se pretende plantear alternativas de solución a la problemática de la producción académica y administrativa generada en la Universidad mediante la elaboración de una guía metodológica que prefigure este espacio, siendo uno de sus objetivos específicos **“proponer un diseño de repositorio institucional que garantice la visibilidad y la accesibilidad de la producción académica y administrativa que se origina en la Universidad”**.

En relación a los metadatos y la interoperabilidad y aportando al mencionado objetivo se mencionan algunas recomendaciones a tener en cuenta para el diseño de un repositorio institucional UNaM²⁸:

- Adherir a OAI como motor del proyecto teniendo en cuenta que es la mejor alternativa para unir todas las fuentes de información sin alterar su independencia de trabajo, a la vez que permite construir un sistema común centralizado útil a todos los usuarios.
- Seleccionar esquemas de metadatos que permitan la interoperabilidad en el contexto OAI y fuera de él, por ejemplo Dublin Core.
- Inscribir el repositorio en servicios como OpenDOAR, ROAR, etc. para que el mismo sea indexado permitiendo su visualización y explotación desde los servicios de búsqueda.

Conclusiones

En la creación de los metadatos convergen elementos bibliotecológicos e informáticos que involucran a una diversidad de profesionales: especialistas de la información, diseñadores de sistemas, proveedores de datos, editores, vendedores, archivistas, ingenieros y especialistas en codificación de texto electrónico; este trabajo interdisciplinar potenciaría el desarrollo de las interfaces correspondientes para la localización, manipulación, recuperación y uso de la información digital.

- Los repositorios institucionales constituyen una alternativa gratuita para hacer visibles los resultados de investigaciones frente a los modelos de publicación, difusión y preservación de contenidos académicos tradicionales caracterizados por la burocracia y el alto costo.
- La Iniciativa de Archivos Abiertos (OAI) proporciona la arquitectura y especificaciones técnicas necesarias para que productores y distribuidores de documentos de acceso abierto lo hagan visibles y accesibles a los investigadores a través de sus servidores. OAI desde el punto de vista del investigador, hace posible la integración de documentos de una misma temática, área geográfica, tipología, etc.
- Si como responsables de escenarios académicos y científicos nuestra misión es lograr que se visualicen los resultados del esfuerzo permanente debemos trabajar para alcanzarlo y el camino es el acceso abierto a la comunicación científica.
- Para ello es necesario demostrar a los actores institucionales la necesidad y la importancia de instalar una política de digitalización que motive e involucre a los profesionales interdisciplinariamente para iniciar sin demora la construcción de un espacio para la UNaM en la web; de lo contrario se notará su ausencia en el mapa de la visibilidad.

²⁸ Se denominan repositorios institucionales al conjunto de servicios web centralizados, ofrecidos por una institución, creados para organizar, gestionar, preservar y ofrecer libre acceso a la producción científica, académica o de cualquier otra naturaleza cultural, en soporte digital, generada por los miembros de esa institución.

Referencias

- Abián, Miguel Ángel (2007).** La web semántica: además de metadatos (ontologías), lógica y confianza / Por. URL: <http://www.wshoy.sidar.org/index.php?2007/01/30/37-la-web-semantica-metadatos-ontologias-logica-y-confianza> [Consultado: 01/11/2009].
- Arencibia, Jorge Ricardo (2006).** Las iniciativas para el acceso abierto a la información científica en el contexto de la web semántica. En: *Biblos*, vol. 7:25-26
- Ayllón Bonet, Julio César (2006/2007).** Metadatos y documentos XML/RDF para Recuperación / Recuperación y acceso a la Información Curso. Universidad Carlos III de Madrid. URL: <http://metadatos-xml-rdf.awardspace.com/docs/MetadatosDocumentosXML-RDF.pdf> [Consultado: 01/11/2009].
- Berlanga, Adriana J., López, Clara, Morales, Erla, García, Francisco J. (2005).** Consideraciones para reforzar el valor de los metadatos en los objetos de aprendizaje (OA). Salamanca : Universidad de Salamanca. Departamento de Informática y Automática e Instituto Universitario de Ciencias de la Educación.
- Berners-Lee, T. (2000).** Tejiendo la Red: el inventor del World Wide Web nos descubre su origen. Madrid: Siglo XXI de España.
- Bodoff, D. Hung, P. C. K., Ben-Menachem, M. (2005).** Web Metadata Standards: Observations and Prescriptions. IEEE Software. Vol. 22, no. 1 78-85
- Brand, Amy. Daly, Frank. Meyers, Barbara (2003).** Metadata Demystified: A Guide for Publishers. Sheridan Press & NISO Press, July.
- Caplan, P. (2003).** Metadata for Education. Chicago: American Library Association.
- Daudinot Fournier I. (2006).** Organización y recuperación de información en Internet: teoría de los metadatos. *ACIMED* ;14 (5). En: http://bvs.sld.cu/revistas/aci/vol14_5_06/aci02506.htm
- Directory of Open Access Repositories.** En: <http://www.openoar.org/> [Consultado: 01/11/2009].
- Ferrer Sapena, Antonia y otros (2005).** Guía metodológica para la implementación de una biblioteca digital universitaria / Gijón : Trea. p. 107
- Gartner, Richard (2008).** Metadata for digital libraries: state of the art and future directions / JISC Technology & Standards Watch En: http://www.jisc.ac.uk/media/documents/techwatch/tsw_0801pdf.pdf [Consultado: 01/11/2009].
- Gómez Dueñas, Laureano Felipe. (2005)** La Iniciativa de Archivos Abiertos (OAI) un nuevo paradigma en la comunicación científica y el intercambio de información. En *Revista Códice*. Universidad de La Salle, N° 4.
- Gómez, Nancy D. y Arias, OM. (2002).** El cambio de paradigma en la comunicación científica. *Información, cultura y sociedad*, <http://eprints.rclis.org/archive/00003649/> [Consultado: 01/11/2009].
- Hernández Ramírez, Haliuska; Sáiz Noeda, Maximiliano (2007).** Ontologías mixtas para la representación conceptual de objetos de aprendizaje / Haliuska Hernández Ramírez, Maximiliano Sáiz Noeda. En: *Procesamiento del lenguaje natural*, ISSN 1135-5948, N°. 38. pags. 99-106 URL: <http://dialnet.unirioja.es/servlet/busquedadoc?db=1&t=repositorios+and+metadatos&td=todo> [Consultado: 01/11/2009].
- Hillman, D. (2003).** Using Dublin Core En: <http://dublincore.org/documents/usageguide/>. [Consultado: 01/11/2009].
- Howarth, Lynne C.** Modelos de Metadatos para Pasarelas Temáticas / Facultad de Estudios de la Información. Universidad de Toronto. Toronto, Canadá En: http://ifla.queenslibrary.org/IV/ifla69/papers/053s_trans-Howarth.pdf [Consultado: 01/11/2009].
- Jornadas Españolas de Documentación (2009).** Datos y metadatos: la normalización dinámica de los elementos y de los procesos constituyentes de una Biblioteca Virtual / Tachi Hernando de Larramendi, Jesús L. Domínguez Muriel, Andrés Viedma Pélaez, Francisca Hernández Carrascal, Xavier Agenjo Bullón. En: *Interinformación: XI Jornadas Españolas de Documentación* : 20, 21 y 22 de mayo de 2009, Auditorio Palacio de Congresos de Zaragoza, ISBN 978-84-692-2663-6, pags. 109-116 . URL: <http://dialnet.unirioja.es/servlet/busquedadoc?db=1&t=repositorios+and+metadatos&td=todo> [Consultado: 01/11/2009].
- Juárez Santamaría, Beatriz (2007).** “Uso de los metadatos en el orden documental. Información, producción y servicios”, Vol. 10, n.º 42. Citado en: Martínez Equihua, Saúl. *Biblioteca digital: conceptos, recursos y estándares*. Buenos Aires : Alfagrama.

Lynch, Clifford A. (2003). "Institutional repositories: essential infrastructure for scholarship in the digital age". ARL bimonthly report, no. 226 (Febr.).

Open Archives Initiative. En: <http://www.openarchives.org/> [Consultado: 01/11/2009].

Open Archives Initiative Protocol for Metadata Harvesting. En: <http://www.openarchives.org/OAI/openarchivesprotocol.html> [Consultado: 01/11/2009].

Peig Olivé, Enric (2004). Interoperabilidad de metadatos en sistemas distribuidos. En: Tesis doctoral dirigida por Jaime Delgado Mercé. Universitat Pompeu Fabra. [Consultado: 01/11/2009].

Peis, Eduardo y Ruiz-Rodríguez, Antonio A. (2004) EAD (Encoded Archival Description): Desarrollo, estructura, uso y aplicaciones [on line]. "Hipertext.net", núm. 2, <<http://www.hipertext.net>> [Consulta: 22/09/2009].

Registry of Open Access Repositories (ROAR). En: <http://roar.eprints.org/> [Consultado: 01/11/2009].

San Segundo María Rosa. (2006). Organización del conocimiento en Internet : metadatos bibliotecarios DUBLIN CORE. Disponible en URL: http://fesabid98.florida-uni.es/Comunicaciones/r_sansegundo.htm

Torres Pombert A. (2006) ¿Catalogación en el entorno digital?: una breve aproximación a los metadatos. ACIMED 2006;14(5). Disponible en: http://bvs.sld.cu/revistas/aci/vol14_5_06/aci09506.htm [Consultado: 01/11/2009].