

Editores de metadatos desarrollados con software libre, una alternativa en la gestión de información

ALEJANDRO JIMÉNEZ LEÓN
MARÍA GRACIELA GUTIÉRREZ VALLEJO
DGSCA, UNAM, México

INTRODUCCIÓN

Si analizamos el comportamiento de los recursos informativos existentes en la web, observaremos que éstos mantienen un crecimiento exponencial, lo que provoca que al momento de buscar determinado tema los usuarios se enfrenten a una oferta excesiva de información y carente de orden para encontrar los datos requeridos. Por ello si deseamos garantizar la localización y recuperación es necesario apoyarnos en aplicaciones informáticas que permitan darle a la información un tratamiento previo para facilitar su recuperación a través de los llamados buscadores de información.

La cuestión es que para realizar estas tareas las instituciones encargadas de la gestión de información requerirían un software especializado, pero desafortunadamente ante la carencia de recursos financieros no pueden cumplir cabalmente con este fin.

OBJETIVO

Por ello el objetivo del trabajo plantea como alternativa para estructurar y normalizar esa información en un entorno digital, la implementación

de aplicaciones informáticas basadas en software libre, las cuales facilitarán la edición y gestión de metadatos en los recursos informativos, de tal forma que se pueda garantizar su localización y recuperación.

Estas aplicaciones buscan facilitar el intercambio de metadatos entre usuarios y se han desarrollado bajo el estándar de Metadatos de Dublin Core.¹

Ahora bien para lograr una adecuada integración entre los recursos informativos digitales y sus metadatos es necesario tener un equipo de trabajo multidisciplinario, un especialista en informática que identifique el funcionamiento de las aplicaciones informáticas, un gestor de información que ubique los metadatos requeridos para determinado recurso informativo, y un diseñador gráfico que permita mostrar ese entorno informativo de forma tal que sea amigable a la vista y al entendimiento de los usuarios.

QUÉ SON LOS METADATOS

Los Metadatos son etiquetas de datos que describen el contenido de un archivo digital y son utilizados para facilitar la gestión, la búsqueda y la recuperación de recursos en la Web. Su objetivo principal es facilitar el intercambio de información entre las máquinas y el usuario. Los metadatos le ofrecen al usuario una serie de datos que le permiten comprender el contenido de un recurso informativo digital, sobre qué trata éste y todo lo que tenga que ver con el interior del mismo. También permiten comprender el contexto, los aspectos relacionados con la creación del Objeto (quién, qué, por qué, cómo y cuándo), la estructura: información sobre asociaciones entre diferentes Objetos de información (que puede ser extrínseca o intrínseca); estos recursos pueden corresponder a un archivo de audio, un documento, un video, una imagen u otras formas de almacenamiento digital.

1 www.dublincore.org

SELECCIÓN DE METADATOS

La selección de los metadatos que sean incorporados a los recursos informativos es la decisión más importante porque determinará su eficaz localización y en consecuencia su capacidad de reutilización por otras personas. De hecho la creación e implementación de metadatos es un proceso intensivo que requiere una importante inversión de tiempo y recursos humanos y económicos, por lo que es necesario hacer un balance de costos y beneficios teniendo en cuenta las necesidades de los usuarios y de los administradores actuales y futuros.

IMPLEMENTACIÓN DE UNA POLÍTICA DE METADATOS

El primer paso y el más significativo para una institución que desee implementar el uso de metadatos es definir una política de metadatos. En ella se enunciarán el modelo de metadatos y los criterios y procesos que formarán el marco rector dentro del cual la institución operará la gestión de los metadatos.

CÓMO ADOPTAR UN ESTÁNDAR DE METADATOS

Si observamos los últimos años, la tendencia es hacia una convergencia sostenida de estándares, pues cada vez resulta más evidente la importancia que tiene el diseño del modelo de metadatos y el impacto que tendrá para la gestión de los archivos digitales, tanto como para el desarrollo de servicios, tales como la recuperación de contenidos, o facilitará su uso y garantizará su preservación. De hecho los metadatos están referidos a aquellos datos que describen o caracterizan un recurso digital, tanto desde el punto de vista interno como externo de modo que todo el archivo digital debe estar perfectamente identificado e incluir la información adicional que permita su contextualización. Por ello la definición del estándar de metadatos y el establecimiento de cómo se asignan los metadatos y en qué aplicaciones van a gestionarse es clave para hacer una correcta gestión de los recursos

digitales.² Uno de los puntos a abordar es entonces el desarrollo de esquemas de metadatos que proveen de reglas respecto a la sintaxis, la semántica y la obligatoriedad de los distintos metadatos.

¿POR QUÉ HACER USO DE LOS METADATOS?

En primer lugar trabajar en la organización de los recursos digitales no es como hacerlo con los catálogos tradicionales que mantienen a los datos centralizados, ni tampoco están éstos organizados por tablas de atributos como sería el caso de las bases de datos.

Por otra parte la ubicación de los recursos puede ser remota o local, y éstos estar concentrados en uno o varios servidores; a lo que hay que agregarle que pueden existir varias versiones en diferente formatos. En consecuencia los archivos están sujetos a constantes cambios y esto aumenta la dificultad para encontrarlos. Por eso es necesario encontrar un modo de organizarlos de manera que se pueda garantizar su localización y recuperación.

Bajo este panorama los metadatos son una propuesta ideal para procesar y manejar los recursos informativos, ya que bajo un uso normalizado resultan ser una solución para los gestores y creadores de información, quienes al disponer de ellos pueden agregarle a esos recursos descripciones susceptibles de ser procesadas por los motores de búsqueda.

¿QUÉ BRINDAN LOS METADATOS A LOS BUSCADORES?

Los metadatos tienen la capacidad de relacionar o establecer enlaces, indizando y clasificando inmensos volúmenes de información de diversos tipos, ya que los metadatos son datos sobre los datos, por lo que un productor de datos puede describir las características del conjunto de datos que produce, de modo que cualquier usuario pueda conocer en qué sistema de referencia se encuentran éstos, qué organismo los

2 La norma ISO 23081 es una guía para comprender, implantar y utilizar metadatos dentro del marco de ISO 15489, Otra norma recomendada es ISAD(G).

ha producido, qué fecha de creación tienen, las medidas de calidad que los evalúan, etcétera.

Con respecto a su tipología, funciones y modelos, los metadatos presentan múltiples alternativas que dependen de diversos factores como el tipo de información que describen, el nivel de estructuración de esos datos, el lugar en donde se ubican, entre otros aspectos. Cabe señalar que en la actualidad los metadatos se encuentran en una fase de construcción técnica y por eso no existe un consenso generalizado en su conceptualización o sobre los tipos y categorías existentes.

EL MODELO DUBLIN CORE

Es un modelo de metadatos elaborado y auspiciado por la Dublin Core Metadata Initiative DCMI, esta organización se dedica a fomentar la adopción extensa de los estándares interoperables de los metadatos y a promover el desarrollo de los vocabularios especializados de metadatos para describir y localizar recursos.

Los elementos que constituyen este perfil son una norma para hacer la descripción de un recurso, entendiéndose por recurso a un fichero, servicio, publicación, programa, página web, autor, fuente, organización, etcétera.

Dublin Core es muy utilizado en el ámbito mundial por su simplicidad; está formado por sólo 15 elementos básicos y generales, para describir cualquier tipo de recurso, y su independencia semántica permite una estructuración de un modo sencillo en formato XML. Su generalidad es aplicable en cualquier disciplina y ámbito semántico, así como su nivel de normalización el cual ha adquirido el rango de norma internacional y se corresponde con la norma ISO 15836: 2003^{16,17} (Información y documentación-Conjunto de metadatos Dublin Core) la cual fue complementada con calificadores, los cuales permiten añadir más detalle y precisión a la descripción de los recursos. *Information and Documentation-The Dublin Core Metadata Element Set*. Termina convirtiéndose en una infraestructura de desarrollo indispensable para la Web Semántica y una de las claves para lograr la interoperabilidad.

IV Encuentro de Catalogación y Metadatos

Los elementos que lo componen se clasifican en tres grupos que indican la clase o el ámbito de la información que se guarda en ellos: los elementos relacionados con el contenido, la propiedad intelectual y la instanciación del recurso.

En cuanto al uso, los metadatos, Dublin Core presentan ventajas que los hacen accesibles para describir recursos en un sistema de información como: simplicidad, flexibilidad, independencia sintáctica, interoperabilidad semántica, alto nivel de normalización formal, crecimiento y evolución del estándar a través de una institución formal consorciada: la DCMI, consenso internacional, modularidad y arquitectura de metadatos en la web.

¿QUIÉN DEBE DESARROLLAR METADATOS?

Esta actividad debería estar a cargo de los gestores de información que manejan los conceptos científicos o quienes conocen más sobre los propios datos para poder documentarlos apropiadamente. No debe asumirse que todos los gestores pueden estar capacitados para escribir metadatos. Por ello es necesario asegurar que hay una coordinación entre quien escribe los metadatos, y el productor de los datos; ya que el primero debe entender el objetivo del creador de información.

¿EXISTE ALGUNA APLICACIÓN QUE VERIFIQUE LA PRECISIÓN DE LOS METADATOS?

Ninguna aplicación puede realizar esta acción, como tampoco puede determinar si los metadatos incluyen elementos que el estándar especifica como obligatorios o si esto corresponde. Por lo tanto se requiere de la revisión de un experto. Pero usualmente la revisión del experto es más simple si ya se sabe que los metadatos tienen la estructura semántica correcta.

MANTENIMIENTO Y PRESERVACIÓN DE LOS METADATOS

Cuando una institución ha implementado el uso de metadatos es esencial que éstos sean sujetos a un mantenimiento periódico, de lo contrario perderán validez y su uso se irá degradando con el tiempo. Para asegurar la eficiencia de los metadatos es necesario que éstos mantengan una viabilidad operativa a largo plazo, de lo contrario el sistema perderá competitividad debido a que los datos serán o estarán incompletos u obsoletos.

Estas implicaciones pueden evitarse si la institución encargada de los recursos informativos digitales realiza las siguientes acciones por lo menos una vez al año: verificar su política de implementación y uso de metadatos, así como el estándar de metadatos implementado que no se haya modificado, reeditado o retirado, y si éste continúa satisfaciendo las necesidades de la institución o si los nuevos estándares internacionales emergentes son más eficientes para la difusión. En este punto debe evaluarse la interoperabilidad que existe entre las instituciones y la dificultad que implica cambiar de un estándar a otro; llevar un registro sobre las fechas en que fueron revisados los archivos, lo cual facilita la gestión del mantenimiento y proporciona una indicación sobre la aceptación del archivo; establecer un procedimiento organizativo que asegure que todos los archivos de nueva creación contengan sus respectivos metadatos, y procurar hacer una evaluación externa acerca de la eficacia de los metadatos de la institución.

EL SOFTWARE LIBRE EN LA EDICIÓN DE METADATOS

En la actualidad existe una infinidad de aplicaciones informáticas desarrolladas bajo software libre orientado al manejo de información digital y a la gestión de servicios de información; desafortunadamente su implementación en muchos casos se ve frenada por factores políticos y económicos. Estas aplicaciones se pueden obtener y descargar libremente de Internet, y en ellos el usuario tiene la opción de bajar sólo la aplicación ejecutable o si lo desea puede obtener además al código fuente para adaptar el desarrollo a las necesidades de la institución. A

IV Encuentro de Catalogación y Metadatos

continuación se enlistan una serie de aplicaciones para la gestión de metadatos todas ellas pertenecientes al ámbito del software libre.

JPEG y PNG Stripper³

PEG y PNG Stripper (Stripper.zip) es una herramienta portable que puede editar información de metadatos de JIFF, JPG, JPEG y archivos de imágenes PNG. El proceso de edición de metadatos no afecta la calidad de la imagen. El uso es sencillo, sólo se tienen que arrastrar las imágenes, fotos o carpetas enteras a la ventana y el software hará el resto. JPEG y PNG Stripper también ofrecen soporte de línea de comandos, tanto para archivos individuales como para carpetas.

Bitcollider⁴

El bitcollider (bc-setup-0.6.0.exe) es una aplicación que genera bitprints y etiquetas de metadatos en los archivos para facilitar su localización.

mp4UI⁵

mp4UI (mp4UI.exe) permite editar archivos de MP4. Su interfaz gráfico puede importar y exportar datos de pistas, así como editar metadatos en los archivos.

AtomicParsley⁶

AtomicParsley (AtomicParsley-win32-0.9.0.zip) es una aplicación con un editor de línea de comandos que puede leer y configurar etiquetas de metadatos para iTunes en archivos MPEG-4 y 3gp.

3 <http://www.pendrivelabs.com/jpeg-and-png-stripper-remove-metadata/>

4 http://downloads.sourceforge.net/project/bitcollider/Bitcollider/0.6.0/bc-setup-0.6.0.exe?use_mirror=voxel

5 http://downloads.sourceforge.net/project/mp4ui/mp4UI/mp4UI%201.0%20RC%202/mp4UI.exe?use_mirror=softlayer

6 http://downloads.sourceforge.net/project/atomicparsley/atomicparsley/AtomicParsley%20v0.9.0/AtomicParsley-win32-0.9.0.zip?use_mirror=softlayer

GeoNetwork⁷

GeoNetwork (geonetwork-install-2.4.2-0.exe) es una aplicación del Sistema de catálogo de metadatos desarrollado por la FAO-ONU, el PMA y el PNUMA. El sistema implementa metadatos de datos geográficos ISO19115/19139, Z39.50, CSW 2.0 y los estándares de OGC WMS, entre otros.

DC Meta Toolkit⁸

Meta de Dublin Core DSpace Toolkit (dcmataoolkit-0.2.0.zip) proporciona la capacidad de convertir grandes volúmenes de información existentes en bases de datos en DSpace metadatos Dublin Core compatibles. Esta aplicación facilita la recopilación de datos y les permite a los usuarios crear convertidores personalizados de sus bases de datos.

Creative Commons Tools⁹

Creative Commons Tools (cchost-5.1.zip), esta aplicación permite editar en los archivos metadatos para que sean recuperados por los motores de búsqueda que buscan información bajo la licencia Creative Commons.

CatMDEdit¹⁰

CatMDEdit (installer_catmdedit-4.5.exe) es una herramienta de edición de metadatos para describir recursos, principalmente de información

7 http://downloads.sourceforge.net/project/geonetwork/GeoNetwork_opensource/v2.4.2/geonetwork-install-2.4.2-0.exe?use_mirror=softlayer

8 http://downloads.sourceforge.net/project/dcmataoolkit/dcmataoolkit/dcmataoolkit%20v0.2.0/dcmataoolkit-0.2.0.zip?use_mirror=softlayer

9 http://downloads.sourceforge.net/project/cctools/ccHost/5.1/cchost-5.1.zip?use_mirror=softlayer

10 http://downloads.sourceforge.net/project/catmdedit/CatMDEdit-Installers/4.5/installer_catmdedit-4.5.exe?use_mirror=softlayer

geográfica. Desarrollado por IAAA (Universidad de Zaragoza) y GeoSpatiumLab SL, fue patrocinado por IGN España; es multiplataforma y multilingüe.

CONCLUSIONES

En la actualidad las instituciones educativas y la sociedad en su conjunto han generado un sinnúmero de recursos informativos digitales, ya sea en forma de texto, imágenes, videos, o grabaciones sonoras y programas informáticos, y todos ellos están al “alcance de amplios sectores de la sociedad”. Sin embargo pese a su difusión universal y a que constituyen un medio insustituible del conocimiento y la creatividad del hombre, los citados recursos en la mayoría de las ocasiones pasan inadvertidos y no son aprovechados debido a la carencia de metadatos que faciliten su ubicación, recuperación y gestión, de allí la relevancia de que una institución implemente una política de metadatos, pues al gestionar sus recursos informativos digitales a través de ella quedan plasmados los criterios y procesos que formarán el marco rector dentro del cual la institución operará la gestión de los metadatos. Estas acciones se traducirán en un uso competitivo de los metadatos que le proporcionaría un extenso rango de beneficios a la institución y sus usuarios, ya que les proveería de claves que facilitan la búsqueda de metadatos a través de distintas áreas de conocimiento; de esta forma la recuperación y adquisición de información sería más rápida y fácil.

Sin embargo las instituciones enfrentan carencias de recursos financieros, de allí la importancia de implementar editores de metadatos desarrollados con software libre que faciliten la gestión de información y en consecuencia la socialización del conocimiento. Además este tipo de software ofrece estabilidad y flexibilidad en su manipulación ya que se dispone del código fuente. De hecho el uso de este tipo de editores permite interoperar y potenciar el uso de información independiente del tipo de los materiales digitales existentes en las instituciones y de los diversos contextos de aplicación.

Como se puede observar existen dos líneas de acción paralelas a seguir que confluyen en un punto: la gestión de información a través de

metadatos. La primera línea es la implementación de una política de metadatos, pero para realizar ésta es necesario contar con software especializado que permita la edición de metadatos en los archivos; es allí donde entra la segunda línea, que consiste en la implementación de editores de metadatos desarrollados con software libre, de esta forma los recursos informativos tendrían los datos necesarios para ser localizados y recuperados. Por ello es importante que los desarrolladores de contenido digital mantengan un equilibrio y presten atención a ambas líneas, las cuales mantienen una dependencia mutua. Cometer el error de privilegiar a una sola línea implicaría la carencia de importantes elementos existentes en la otra.

Por otra parte, si bien los metadatos se encuentran en una fase de construcción técnica, sería recomendable observar el surgimiento de propuestas y esquemas de metadatos en otros países y sectores, y analizar cómo han subsanado sus necesidades en función de sus recursos y medios; de esta forma los actuales proyectos de gestión de información tendrían una mejor idea sobre la implementación y evolución de los metadatos.

Es claro que el éxito o fracaso de un proyecto que intente gestionar información digital dependerá de la interacción que logren los integrantes del equipo y de cómo maximicen sus recursos y minimicen sus costos, por ello el uso de software libre es una excelente alternativa para la gestión de información

BIBLIOGRAFÍA

Day, Michael. Metadata for digital preservation: an update. [en línea]. <http://www.ariadne.ac.uk/issue22/metadata> (8 may2009)

Deegan, M., Tañer, S.: 8: "Preservation. Digital Futures series". In: *Digital Futures: strategies for the information age*. Library Association Publishing, 7 Ridgmount Street, London WC1E 7AE (2002) 179-208

IV Encuentro de Catalogación y Metadatos

Guinchat Claire. *Introducción general a las ciencias y técnicas de la información y documentación*, - España: CINDOC, 1990.

Harvey, Ross. *Preservation in libraries: principles, strategies and practices for librarians*. London. Bowker Saur. 1993. (Topics in Library and Information Studies).

Martínez Comeche, Juan Antonio. *Teoría de la información Documental y de las instituciones documentales*. - España : Síntesis, 2000.

Mayor Zaragoza, Federico. "Sociedad de la Información en el siglo XXI: un requisito para el desarrollo II Reflexiones y conocimiento compartido. Capítulo 7". [en línea]. <http://www.desarrollosi.org/Volumen2/Web/PDF/Cap7_La_persona_en_la_SI.pdf> (10 ene 2009).

Moreira González, José Antonio. "Nuevas Competencias profesionales para nuevas funciones". [en línea]. <<http://www.ana bad.org/archivo/docdow.php?id=204>> (14 ago 2009).

Peig Olivé, Enric. Interoperabilidad de metadatos en sistemas distribuidos. Barcelona: Universitat Pompeu Fabra, 2003. [en línea]. http://www.tdx.cbuc.es/TESIS_UPF/AVAILABLE/TDX-0316104-132946/tepo1de1.pdf (10 abr 2009)

Ros García Juan. Políticas de Información y documentación. - España : Síntesis, 2002.