

EL ACCESO A LA INFORMACIÓN EN LA RED. LOS METADATOS: LA DUBLIN CORE METATADA INITIATIVE.

I. INTRODUCCIÓN

Desde hace algunos años vivimos inmersos en la denominada **Sociedad de la Información**, un **nuevo contexto social y cultural** donde el **acceso al conocimiento** es considerado como un bien cada vez máspreciado en todos los sectores y ámbitos. Gran parte del éxito de esta nueva realidad social radica en las **redes de telecomunicaciones**, entre las que sin lugar a dudas destaca **Internet**, la red de redes por excelencia en nuestros días. Es, igualmente, una sociedad en la que las **NTIC** avanzan a pasos agigantados, por lo que su correcto uso y manejo requieren unos ciertos niveles de formación y normalización.

II. EL ACCESO A LA INFORMACIÓN EN RED

Internet se ha convertido en un **depósito inconmensurable de información**, en el que conviven múltiples recursos de calidad variable y cierta inestabilidad, que no siempre son fáciles de localizar. Existe una gran cantidad de directorios, buscadores o metabuscadores que permiten buscar información en la Red, pero su grado de precisión no es siempre el más deseable y el número de resultados que despliegan es, a menudo abrumador. Normalmente, junto a algunos documentos relevantes, aparecen muchos otros de menor pertinencia o nulo interés, sin haber además un control de duplicados eficiente. Por esta razón, se hace tan necesario aplicar patrones de descripción normalizados a este tipo de recursos, ya que esto facilitará enormemente la identificación y recuperación de los mismos.

Las bibliotecas son muy conscientes de que deben **seleccionar y poner al servicio de sus usuarios los recursos gratuitos de interés** que les ofrecen las nuevas tecnologías. Pero, para que dichos usuarios puedan encontrarlos y utilizarlos, es necesario que los documentos electrónicos reciban un tratamiento documental similar al de los documentos tradicionales. Así, desde hace unos años, se están realizando diversos **esfuerzos para crear o adaptar herramientas capaces de controlar este nuevo tipo de materiales**. En este sentido, contamos ya con interesantes avances, como, por ejemplo, las **ISBD (ER)**, publicadas en 1997 para normalizar la descripción de recursos electrónicos o las **ISBD (CR)**, publicadas en 2003 para los recursos continuos. El Formato **MARC**, por su parte, presenta la **etiqueta 856** para poder incluir la URL de un documento en línea, es decir, la dirección web en la se encuentra. Además, la **ISO** regula la redacción de referencias bibliográficas de recursos en línea mediante su **norma 690-2**.

En cuanto al debate sobre si es más operativo adaptar herramientas descriptivas ya existentes al entorno digital o crear nuevas alternativas, los especialistas en la materia apuestan por la utilización de formatos más simples, adaptados a las necesidades concretas de los recursos en red y creados específicamente para el tratamiento de este tipo de materiales. En esta línea se perfilan los **sistemas de metadatos**, una de las soluciones actuales más eficaces para la correcta gestión y normalización de los recursos web.

III. LOS METADATOS

Definidos de una manera simple y sencilla, los metadatos no son sino **datos estructurados sobre otros datos**. En relación con el entorno Web, se trata de datos guardables, intercambiables y procesables por medio del ordenador, organizados de tal forma que sirven para la **identificación, descripción, clasificación, localización** y, por ende, **recuperación** de un recurso web.

Desde el punto de vista bibliográfico, los metadatos conforman **documentos secundarios**, ya que resultan del tratamiento documental de un recurso primario, con el objetivo de hacerlo más accesible.

De manera general, los metadatos son informaciones descriptivas sobre un recurso dado, ya sea físico (es decir, tradicional y tangible) o electrónico. De ahí que **Tim Berners-Lee**, padre de la Web y precursor de la Web Semántica, se refiriera a ellos como «información inteligible por ordenador sobre recursos web u otras cosas». A pesar de ello, el concepto de metadato suele asociarse fundamentalmente a la descripción de documentos electrónicos y a la mejora para acceder a los mismos a través de la Red. Por esta razón, tanto su concepción como su aplicación deben abordarse desde una **doble perspectiva**: la biblioteconómica y la informática, buscándose la sinergia entre ambas.

Como señalan **Weibel y Lagonze**, dos líderes destacados en la materia, la asociación de metadatos descriptivos normalizados a objetos de Red posee el potencial de mejorar sustancialmente la localización de recursos y, por ende, el acceso a sus contenidos. La adopción de este tipo de estándares a gran escala no sólo promete garantizar una mejor recuperación genérica de la información en línea, sino también aumentar la capacidad de Internet como fuente de información y contribuir a la organización de documentos en entornos más complejos y especializados, como bibliotecas virtuales e híbridas.

IV. DUBLIN CORE METADATA INITIATIVE (DCMI)

La **Iniciativa de Metadatos Dublin Core** o **DCMI** es una organización dedicada a la promoción y difusión de normas y vocabularios interoperables sobre metadatos con el fin de describir, de forma coherente y eficaz, la información contenida en Internet. Para el desarrollo y mantenimiento de sus estándares y recomendaciones, la DCMI organiza talleres internacionales, conferencias y, sobre todo, **reuniones de grupos de trabajo**, que se forman o disuelven según las necesidades existentes. Cuenta con componentes de muy diferentes países y gran parte de ellos son, además, miembros activos del **W3C** (*World Wide Web Consortium*), organismo encargado de velar por la normalización en Internet, que ha desarrollado aplicaciones tan importantes como el lenguaje de marcas XML o la arquitectura RDF. Ello le asegura una buena infraestructura técnica.

En la actualidad, su propuesta internacional de metadatos, conocida como **Dublin Core (DC)**, es la más sólida y más ampliamente aceptada como estándar. Surgió en **1995**, en la primera reunión de la DCMI (*Dublin Core Series Workshop*), que tuvo lugar en Dublín (Ohio, EEUU), ciudad de la que tomó su nombre. Fue el resultado del trabajo cooperativo internacional de bibliotecarios, informáticos y otros expertos, auspiciado por el **OCLC** (*Online Computer Library Center*), en colaboración con otros organismos como el **IETF** (*Internet Engineering Task Force*). Desde entonces, la DCMI trabaja por el desarrollo y refinamiento de su modelo de metainformación, inicialmente desarrollado en inglés, pero traducido ya a más de veinticinco lenguas, entre ellas el español.

Las principales **CARACTERÍSTICAS del DC** son:

- **Simplicidad**, dado que los metadatos están diseñados para ser rellenados de forma sencilla, tanto por documentalistas como por los propios autores al crear sus recursos.
- **Interoperabilidad semántica**, ya que los modelos de descripción dispares interfieren con la capacidad de búsqueda de recursos.
- **Consenso internacional**, tanto en el número como en la definición de los elementos.
- **Flexibilidad**, ya que los elementos del DC son **opcionales, repetibles** y **no** establecen un **orden predeterminado rígido** ni una **longitud máxima** que limite la profundidad de las descripciones.
- **Extensibilidad**, ya que el DC ofrece mecanismos para completar sus elementos mediante otros más específicos, pensados para entornos más locales o especializados.

En cuanto a sus **ELEMENTOS**, el **Conjunto de Metadatos Dublin Core** (*Dublin Core Metadata Element Set*, DCMES) está registrado como norma **ANSI/NISO Z39.85** desde 2001 (actualizada en 2007) y como **norma ISO 15836** desde 2003. En su versión principal y más sencilla, consta de **15 elementos básicos**, perfectamente definidos, para organizar la información. Son los siguientes:

1. **Título** (*Title*): nombre dado a un recurso, generalmente por el autor.
2. **Autor o creador** (*Creator*): persona, organización o servicio responsable de la creación intelectual del recurso.
3. **Materia** (*Subject*): tema del contenido del recurso. Se recomienda expresarlo mediante un lenguaje controlado.
4. **Descripción** (*Description*): resumen del contenido del recurso.
5. **Editor** (*Publisher*): entidad responsable de difundir el recurso en la Red.
6. **Colaborador** (*Contributor*): persona, organización o servicio a los que se atribuye una contribución intelectual significativa, aunque secundaria, para la creación del recurso.
7. **Fecha** (*Date*): fecha asociada a la creación o publicación del recurso en Red. La práctica recomienda seguir el formato marcado por la ISO (AAAA-MM-DD).
8. **Tipo de recurso** (*Type*): naturaleza, categoría o género del contenido del recurso. Se recomienda igualmente describir este elemento conforme a un vocabulario controlado.
9. **Formato** (*Format*): manifestación física o digital del recurso (es decir, tipo de soporte, hardware, software o equipamiento necesario para su visualización).
10. **Identificador** (*Identifier*): referencia unívoca al recurso. Se recomienda usar algún sistema de identificación formal, como el URI (Identificador Uniforme de Recursos), URL (Localizador Uniforme de Recursos), el DOI (Identificador de Objetos Digitales) o el código normalizado ISBN.
11. **Fuente** (*Source*): referencia al posible recurso del cual se deriva el recurso presente.
12. **Lengua** (*Language*): idioma en que se expresa el contenido intelectual del recurso, preferentemente de forma codificada.

13. **Relación** (*Relation*): referencia a un recurso relacionado.
14. **Cobertura** (*Coverage*): alcance o referencia espacial y temporal del contenido del recurso (lugar de localización, jurisdicción, etc. extraídos de una lista controlada).
15. **Derechos** (*Rights*): información sobre los derechos legales que afectan al uso del recurso (PI, ©, etc.).

Como puede apreciarse, la DCIM recomienda extraer los valores de ciertos elementos de un lenguaje controlado y aceptado internacionalmente (normas ISO, clasificaciones como la CDU, LC, Dewey...), ya que así se facilita la comunicación con otros programas heterogéneos y, por tanto, la posibilidad de intercambiar metadatos.

SIMPLE DC VS QUALIFIED DC

La norma DC establece dos niveles: simple y cualificado.

El **DC Simple** comprende únicamente los 15 elementos mencionados (opcionales y repetibles), sin cualificadores, es decir, sin añadir información alguna que matice o refine las descripciones. Presenta, por tanto, un esquema básico atributo-valor.

El **DC Cualificado**, además de los 15 elementos clave, permite la introducción de otros, como la **audiencia** (a quién va dirigido el recurso), la **procedencia** (cambios significativos en la custodia del recurso) o el **depositario de derechos** (sobre quién recae la propiedad o gestión de los mismos). Pero, aparte de admitir elementos adicionales, permite el uso de cualificadores para aumentar la especificidad o precisión de los metadatos. La DCMI recomienda dos tipos principales:

- **Limitadores o cualificadores referidos al elemento** (*element refinement*), que matizan y completan el significado de éste, como si se tratara de un sustantivo concretado por un adjetivo o modificador. Por ejemplo, para el elemento "Título", un cualificador sería "Alternativo", el elemento "Relación" podría delimitarse mediante "Es parte de", "Reemplaza a", etc.
- **Codificadores o cualificadores referidos al valor** (*encoding scheme*), que aclaran y explicitan el sistema de codificación utilizado para rellenar la información, tantas veces necesario para interpretarla correctamente. Se trataría, por ejemplo, de especificar que para consignar la el valor de lengua «spa» se ha utilizado la **norma ISO 3166-1 alpha-3: 1998**, que para rellenar el valor correspondiente a la materia se ha utilizado la CDU, etc.

No obstante, no todos los elementos presentan limitadores, como tampoco todos admiten codificadores. Los cualificadores aprobados hasta la fecha no constituyen un grupo cerrado, por lo que es previsible que en el futuro aparezcan otros aplicables a dominios más especializados o localizados.

V. DUBLIN CORE EN ESPAÑA

El DC en nuestro país aún no está demasiado extendido, aunque algunos centros como la **Biblioteca Virtual Miguel de Cervantes** ya han mostrado su interés por catalogar recursos electrónicos con él. Los principales esfuerzos por conseguir una mayor presencia española en la DCMI se deben muy especialmente al **Instituto Agustín Millares**, de la Universidad Carlos III de Madrid, al **Grupo NORMAWEB**, de la Asociación Española de Documentación e Información (**SEDIC**) y a la **BNE**, pues todos ellos hicieron posible la celebración del seminario “Aplicación en España del modelo de metadatos DC” en 2003. Una importante fuente de información cercana sobre el conjunto de metadatos DC es, sin duda, **la lista DCMI-ES**, de RedIRIS, un foro abierto en español para todos aquellos interesados en la materia. A título particular, cabe destacar a la **Dra. Eva María Méndez Rodríguez**, del Dpto. de Biblioteconomía y Documentación de la UC3M, por ser la gran impulsora de los metadatos y estándares para la web en nuestro país. Su universidad fue precisamente la que, en septiembre del año 2005, acogió la Conferencia Internacional de la DCMI.

VI. CONCLUSIÓN

En definitiva, el modelo de metadatos DC constituye uno de los mayores esfuerzos internacionales e interdisciplinarios para definir el conjunto de elementos o atributos básicos necesarios para describir y recuperar información en Internet. Además de facilitar la identificación de los documentos que circulan en la Red, desempeña una importante función catalogadora y de análisis documental, ayudando al intercambio de información entre los miles de agentes vinculados con los objetos digitales. No obstante, y a pesar de todas las ventajas señaladas, las facilidades del sistema para rellenar los metadatos dan lugar, en ocasiones, a usos fraudulentos por parte de los propios autores. Así, algunas páginas web de origen comercial suelen incluir determinados metadatos, ajenos al contenido de las mismas, con el único propósito de engañar a los sistemas de indización y atraer a más visitantes de forma ilícita.

PUNTOS TRATADOS EN EL TEMA

EL ACCESO A LA INFORMACIÓN EN LA RED. LOS METADATOS: LA *DUBLIN CORE METATADA INITIATIVE*.

I. INTRODUCCIÓN

II. EL ACCESO A LA INFORMACIÓN EN RED

III. LOS METADATOS

IV. DUBLIN CORE METADATA INITIATIVE (DCMI)

SIMPLE DC VS QUALIFIED DC

V. DUBLIN CORE EN ESPAÑA

VI. CONCLUSIÓN