

# Evolución del Dublin Core Metadata Initiative

## José A. Senso

Universidad de Granada.  
Facultad de Biblioteconomía y Documentación  
jsenso@ugr.es

## Antonio de la Rosa Piñero

Wisdom Information Consultants B.V. Amsterdam. Holanda  
antonio@wisdom.nl

**Resumen:** *El presente trabajo trata de trazar cronológicamente la evolución de uno de los sistemas de metadatos que más impacto y repercusión ha tenido dentro del ámbito bibliotecario: el Dublin Core Metadata Initiative. Al mismo tiempo se describen algunos de los principales problemas que se encuentra este sistema para lograr su implantación así como el efecto que ha tenido dentro del entorno español.*

**Palabras clave:** *Dublin Core, Dublin Core Metadata Initiative, DCMI, Metadatos, Recuperación de información, Canberra Qualifiers, Z39.85.*

Dublin Core nace como producto del trabajo cooperativo de ámbito internacional - promovido en su primera fase por la OCLC (Online Computer Library Center) y el NCSA (National Center for Supercomputing Applications)- cuyo objetivo principal fue crear un conjunto de elementos que permitieran la descripción de recursos electrónicos con el fin de facilitar su búsqueda y recuperación (**Ortiz-Repiso Jiménez**, 1999).

Originalmente se concibió como un conjunto de etiquetas que deberían ser generadas por el autor del documento HTML con la finalidad de facilitar su identificación y posterior recuperación en Internet. No obstante, este modelo ha llamado la atención de diversas comunidades de profesionales interesados en la descripción de recursos en museos, bibliotecas, organizaciones comerciales... (**DCMI**, 1998).

Desde que en 1995 naciera este proyecto se ha utilizado el mismo esquema de trabajo. Generalmente cada pocos meses se reúne una serie de expertos del mundo bibliotecario, de la gestión de redes y comunicaciones, informáticos, etc. para estudiar su impacto, analizar modificaciones e investigar en posibles aplicaciones.

A lo largo de todas estas sesiones de trabajo se han mantenido los objetivos básicos que dieron pie al primer conjunto de metadatos Dublin Core Metadata Initiative (DCMI) (**DCMI**, 1998), (**Weibel**, 1995):

- Establecer un sistema normalizado para la descripción de documentos distribuidos en el World Wide Web, cualquiera que sea su formato: HTML, PDF, PS, SGML...
- Facilitar a autores y editores de estos documentos la incorporación de elementos que identifiquen y describan sus aportaciones.
- Corregir las deficiencias de los sistemas de recuperación basados en texto completo.
- Evitar el uso de formatos complejos (MARC, TEI, etc.).
- Crear una serie de elementos:

- Fáciles de crear y actualizar.
- Comprensibles para cualquier usuario.
- Normalizados.
- Ampliables manteniendo la compatibilidad con definiciones anteriores.
- A los que se pueda aplicar diferentes niveles de especificidad y control mediante el uso de calificadores y subtipos.
- Que faculten la integración haciendo uso de sistemas de clasificación, indización y control de autoridades con los que la comunidad bibliotecaria se encuentra familiarizada.
- Con carácter internacional en su creación, desarrollo e implementación.
- Referidos a las características intrínsecas de los documentos.
- Opcionales.
- Repetibles tantas veces como sea necesario.

En realidad, el Dublin Core no es un sistema de conversión de metadatos, tal y como afirman algunos autores (**San Segundo Manuel**, 1998), sino que es un sistema que define un conjunto básico de atributos que sirven para describir los recursos existentes (**Gimeno Montoro, et al.**, 1998) en cualquier tipo de red.

El conjunto Dublin Core fue creado en 1995. En las primeras versiones estaba formado por 13 elementos que permitían identificar las características básicas de un recurso electrónico. A pesar de que un año más tarde se le añadieron 2 elementos más, el modelo de referencia sobre el que se ha trabajado ha sido siempre el mismo. Los elementos principales eran el recurso (es decir, el objeto que se piensa describir), el elemento (definido como la característica o propiedad del recurso) y, por último, el valor (que se corresponde con la cualidad definida por el elemento). El esquema con el que se trabaja se encuentra en el gráfico 1.

El recurso puede ser un objeto que, como tal, tiene una identidad -definida como cualquiera de los elementos del DC- y ésta, a su vez, un valor que la califica -definido como una cadena de caracteres-. Así, por ejemplo, el elemento título está formado por el identificador título del recurso (en este caso, dc.title) y por el valor constituido por toda la cadena de caracteres que forman el título del recurso.

Dada la importancia de este conjunto de metadatos, así como su calado y algo grado de penetración en el ámbito bibliotecario, consideramos necesario realizar un recorrido cronológico por su evolución a lo largo de las diferentes sesiones de trabajo, incidiendo además en aquellos acontecimientos producidos fuera de estas y que han ayudado a su posterior desarrollo.

## 1.- OCLS/NCSA Metadata Workshop

La primera reunión, denominada DC1, se celebró en la sede de la OCLC en Dublín, Ohio, y de ahí toma su nombre este conjunto de metadatos. Los 52 especialistas que formaron este primer grupo de trabajo tenían como principales objetivos el análisis del estado del arte en el desarrollo de la descripción de recursos electrónicos y la creación de un modelo que facilitara esta descripción (**Lassila**, 1999).

El problema del que se partía era el mismo que ha provocado la aparición del resto de sistemas de metadatos actuales: el crecimiento de Internet -especialmente del Web- y la gran cantidad de documentos electrónicos generados. Al mismo tiempo, estos

recursos no se describen de una forma lo suficientemente rica como para que los buscadores puedan garantizar su localización de una forma precisa (Weibel, 1995).

Se consideró que la única solución a este problema pasaba por la creación de un conjunto de herramientas que permitiera, a usuarios no necesariamente especializados en tareas de catalogación, la introducción de datos referentes al contenido del documento que estaban creando. La respuesta estaba en los metadatos.

Resultaba evidente que la gran cantidad de información que había en Internet no podía ser procesada o analizada por catalogadores o indizadores profesionales. Era obvio que había que facilitar los medios para que los autores realizaran las descripciones sin tener que pasar por un proceso de formación complejo. Partiendo de este objetivo, y de que la localización del recurso era la necesidad más urgente, uno de los grupos de trabajo, el Metadata Workshop, empezó a realizar su tarea.

Este taller orientó su labor hacia el tratamiento del recurso electrónico como si se tratase de un objeto (DLO<sup>1</sup>, Document Like Object). De esta forma, el conjunto de metadatos creado sería válido para cualquier tipo de objeto (documento, imagen, sitio Web...). Si bien es cierto que esta consideración ampliaba el campo de acción<sup>2</sup>, también producía ciertos “daños colaterales”, como, por ejemplo, la eliminación de algunos elementos que permitían una descripción más exacta como los derechos de autor o la difusión del documento.

El conjunto de metadatos propuesto (Dublin Core, DC) en el DC1 pretendía recoger los rasgos esenciales de los documentos electrónicos<sup>3</sup>. A pesar de esta consideración, la idea original era que el DC complementara a otros formatos de descripción<sup>4</sup>, y no que los suplantara.

A continuación analizaremos todas las meta-etiquetas en sus dos modelos, el simple y el completo (con esquema de clasificación), tal y como se definieron en el DC1<sup>5</sup>.

**Subject:** esta etiqueta incluía tanto la descripción (resumen del contenido) del objeto como las palabras clave. En realidad éste fue uno de los primeros elementos que más adelante sufrió variaciones, ya que existían muchos problemas -especialmente a la hora de la recuperación- para que una sola etiqueta fusionase el resumen y las palabras clave.

**Title:** en aquellos casos en los que el título no sea tan obvio como en un documento textual (imagen gráfica, objetos físicos...) es posible utilizar una frase descriptiva en su lugar.

**Autor:** todas aquellas personas que tengan una responsabilidad intelectual con respecto a la creación del objeto.

**Publisher:** la persona o empresa encargada de las tareas de edición relacionadas con el objeto deben consignarse de la siguiente manera.

**OtherAgent:** la idea sobre la que se basó la creación de esta etiqueta fue la de dar cabida a todas aquellas personas que, sin estar dentro de Author o Publisher, mantengan una relación directa con la realización del objeto (ilustradores, traductores, compiladores, fotógrafos...). Para poder desarrollar convenientemente esta etiqueta se propuso la creación de un grupo de trabajo que definiera los posibles elementos que la podían formar siguiendo el modelo. No obstante, el DC1 se percató de que esta etiqueta quedaba muy por debajo, en cuanto a calidad de la información que transmite,

con respecto a similares etiquetas en otros sistemas. Por ese motivo, se contempló la posibilidad de añadir como cualificador esa misma etiqueta pero en otro conjunto de metadatos.

**Date:** fecha a partir de la cual el objeto está disponible.

**ObjectType:** tanto para especificar el género del objeto (novela, poesía), como su forma interna (diccionario, tesoro) o su propósito (borrador, versión definitiva).

**Form:** permite informar sobre el software y/o hardware necesario para mostrar u operar con el objeto en cuestión. Por lo general se trabaja siempre con la forma compleja, que incluye una definición tipo de los principales elementos que se pueden encontrar en Internet, el IMT (Internet Media Type).

**Identifier:** por lo general, este elemento suele ir acompañado del cualificador con el que se trabaja.

**Relation:** esta relación, que existe entre el objeto y otros objetos (o partes), es poco común en Internet. Tanto es así que, en el DC1, apenas desarrollaron este metadato. Tan sólo contemplaron la posibilidad de añadirle dos subelementos: type e identifier.

**Source:** el principal objetivo de este elemento era el de servir de conexión entre el objeto y su versión anterior. Puede, a su vez, incluir información de otras meta-etiquetas del mismo DC.

**Language:** especifica el lenguaje del contenido intelectual del objeto.

**Coverage:** características temporales y espaciales del objeto.

Todo este conjunto de elementos constituye la versión 0.1 del Dublin Core.

En la mayoría de los casos estos elementos se podrían utilizar en su forma más simple. Pero, además, se consideró imprescindible el establecer mecanismos que permitieran un mayor nivel de especificidad en caso necesario. Para ello se desarrolló una serie de cualificadores que complementaban la acción de las etiquetas. Junto al incremento de especificidad, que permite una descripción más precisa por medio del uso de clasificaciones (LCSH, Dewey, MeSH...), los cualificadores detallaban las reglas de codificación utilizadas, definían la sub-estructura formal de cada elemento (aportando, por ejemplo, no sólo el nombre del autor intelectual del documento, sino también su dirección de correo-e, organización para la que trabaja...) y facilitaban el control de autoridades.

## 1.1.- Principales problemas derivados del DC1

En el documento principal (**Lassila**, 1999) generado después del DC1 se reconocieron, en su apartado número siete, los siguientes problemas sin resolver después de las reuniones de trabajo:

- **Versiones:** no se pudo llegar a ningún consenso con respecto al tratamiento de las diferentes versiones de un recurso electrónico. La facilidad con la que se puede crear documentos electrónicos hace que también sea más fácil la aparición de diferentes versiones de un mismo documento con ligeras modificaciones entre unas y otras. DC no podía gestionar las diferentes

versiones de un mismo recurso electrónico<sup>6</sup>. La única solución intermedia propuesta pasaba por la utilización de las etiquetas Source y Relation para estos fines.

- **Flexibilidad:** aunque se partía de la idea de que el DC debía ser lo suficientemente flexible como para permitir la descripción de todo tipo de recursos, el DC1 dejaba constancia de la dificultad que entrañaba el hecho de que un conjunto tan heterogéneo de profesionales pudiera encontrar satisfechas sus diferentes necesidades por el uso de un mismo conjunto de metadatos.
- **Conjuntos de caracteres:** hasta la fecha, el conjunto ASCII había sido ampliamente utilizado por toda la comunidad científica en sus comunicaciones por las redes. Internet ha favorecido la aparición de diferentes conjuntos que no se tenía muy claro cómo iban a reaccionar ante la introducción de DC. Resultaba evidente que el éxito de DC dependía del estudio de la relación entre SGML/HTML y MIME<sup>7</sup>. Se instó a la IETF (Internet Engineering Taskforce) y al grupo de trabajo del nuevo protocolo HTTP que intentaran rectificar las inconsistencias producidas en MIME con relación a SGML.
- **Otros metadatos:** se contempló la necesidad de que DC permitiera la compatibilidad con otros sistemas -ya se estén utilizando en Internet o en bibliotecas tradicionales o digitales- con el fin de poder generar bases de datos comunes con información en diferentes formatos.

## 2.- OCLC/UKOLN Metadata Workshop

El segundo taller de trabajo (DC2) tuvo lugar en la Universidad de Warwick, Reino Unido, del 1 al 3 de abril de 1996, y fue organizado por la OCLC, de nuevo, y por el UKOLN (United Kingdom Office for Library & Information Networking). En la reunión celebrada el año anterior en Dublín no se había llegado a definir ninguna sintaxis para consignar los elementos de DC. Este problema se abordó en el DC2 junto al estudio de la estrategia para introducir aún más el sistema de metadatos dentro de la comunidad científica. En palabras de los organizadores, los objetivos se centraban en:

- Promover la interoperatividad semántica por disciplinas e idiomas
- Definir mecanismos para flexibilizar la descripción de los recursos

Con respecto a la sintaxis, se produjo un documento que recogía las recomendaciones a las que se había llegado en el DC2 así como las generadas por el grupo de trabajo específico (**Burnard, et al.**, 1996).

En este documento se defendía la idea de utilizar la etiqueta META de HTML para insertar los meta-elementos de DC. Al mismo tiempo se propuso una DTD que modificaba la anterior para adaptarla a esta nueva idea.

En realidad, el HTML era un lenguaje ideal para contener a este conjunto de metadatos. En primer lugar porque facilita mucho la creación de documentos electrónicos, alejándose de la complejidad de todos aquellos lenguajes que trabajaban con un conjunto mayor de etiquetas SGML. Además, satisfacía los criterios mínimos exigibles según el grupo de trabajo dedicado a la definición de la sintaxis:

- Posibilidad de utilizar secuencias repetibles y opcionales
- Robustez que sea posible verificar (por medio de un parser)

- Compacto
- Uniforme
- Que permita la introducción de cualificadores
- Procesable por el ordenador pero, a la vez, comprensible para el hombre
- Con carácter internacional
- Simple

La sintaxis ideal debería tener, además, unos requerimientos mínimos:

- Debe ser lo suficientemente simple como para que los autores puedan ser capaces de introducir los metadatos en sus documentos (directamente o a través de una plantilla).
- Debería poder representar todos los elementos de DC.
- Ser lo suficientemente flexible como para permitir que los usuarios puedan utilizarlos en su forma más simple o en la compleja (con mayor grado de descripción).
- Tendría que tener la capacidad de expresar correctamente el triple atributo-valor-esquema (attribute-value-scheme triplet).
- Que sean opcionales y repetibles.

Tras esto, y un estudio detallado del HTML y de las posibles variantes de DTDs, se propusieron las siguientes alternativas:

- Incluir los metadatos dentro de los documentos HTML haciendo uso de la etiqueta META ya existente, teniendo en cuenta que:
  - La información sobre el esquema (Scheme) y tipo (Type) propuesto en el DC1 no puede asignarse dentro del valor Name, al no estar incluidos dentro de la DTD. Este punto se desarrolla mejor en el apartado 2.1 de este trabajo.
  - Las etiquetas LINK deben escribirse detrás de todas las etiquetas META. Este punto se desarrolla más detenidamente en el apartado 2.2.
  - Había que modificar la actual DTD de HTML para adaptar la etiqueta META a las propiedades de los nuevos elementos.
- Utilizar la etiqueta LINK para unir el objeto con el documento donde se encuentren sus metadatos correspondientes. Este documento ordenaría los metadatos de dos maneras posibles:
  - Utilizando las etiquetas para la definición de listas ordenadas.
  - Diseñar una nueva DTD totalmente independiente de la especificación HTML.

Meses más tarde, en junio de ese mismo año, el consenso en el W3C Distributed Indexing and Searching Workshop (en el que participaban representantes de Dublin Core, Lycos, Microsoft, WebCrawler, IEEE, Verity y el World Wide Web Consortium - W3C-) produjo una propuesta final que proyectaba como solución el uso de la etiqueta META del HTML con pequeñas modificaciones (**Weibel**, 1996).

Se eligió por esta opción debido a:

- La dificultad para añadir nuevos atributos, e incluso etiquetas, a la versión actual del HTML.
- El poco compromiso por parte de las empresas de navegadores (browsers) a la hora de hacer los cambios necesarios.
- Que se deseaba involucrar a los robots que alimentan bases de datos.

## 2.1.- Propiedades scheme y type

Como observamos, la integración que se pretendía con HTML no resultaba posible mientras la DTD de este lenguaje no aceptara dos cualificadores necesarios para el correcto desarrollo del modelo:

- **Scheme:** indica el sistema de codificación empleado para consignar el valor a cada propiedad. Se partía de la base de que debía tratarse de un sistema ampliamente reconocido en el mundo bibliotecario. A la hora de integrar este atributo a la etiqueta determinada, y ésta a HTML.

La DTD de la etiqueta META no contempla, en la versión 2.0 del HTML (la que se utilizaba en 1995, fecha del DC1), la propiedad Scheme. La única solución posible para solventar este problema pasaba por utilizar un equivalente que no daría problemas con ningún parser ni con los navegadores. Es decir, utilizar la propiedad Content como un “cajón de sastre”:

- **Type:** describe una característica de la entidad referenciada por la etiqueta. En principio, su objetivo era únicamente complimentar a la etiqueta Relation.

## 2.2.- Casos especiales: la etiqueta LINK

Desde las primeras versiones de HTML se incluía la etiqueta LINK, que permite definir relaciones entre documentos. Esta etiqueta, que debe ubicarse dentro del encabezamiento del documento, tiene varios atributos, pero dos de ellos son especialmente útiles para el trabajo con metadatos:

- **REL:** define una relación padre-hijo entre el documento vinculante con el vinculado.
- **REV:** especifica la relación contraria: entre el documento vinculado y el vinculante.

Siempre que se hizo referencia a un sistema de codificación mediante la propiedad Scheme se recomendó el establecer un enlace con el documento en el que se explica, de forma pormenorizada, el sistema de clasificación escogido. Inicialmente, esta etiqueta se utilizaba para enlazar un objeto con documentos en los que se especificaba información sobre sus derechos de autor y/o uso.

Aunque esta aplicación no está descartada del todo, su principal objetivo consiste en ofrecer a los autores la posibilidad de registrar nuevos sistemas de metadatos sin que sea necesario establecer un directorio centralizado.

## 2.3.- The Warwick Framework

Junto a la sintaxis y la DTD para SGML propuesta, en el DC2 se llegó también al convencimiento de que era necesario compatibilizar otros sistemas de metadatos. Para ello era indispensable la creación de una arquitectura que permitiera la integración de distintos conjuntos de metadatos.

Las siguientes características deseables sobre las que se planteó el WF fueron las siguientes:

- **Modular:** que facilite la inclusión de diferentes tipos de metadatos.
- **Extensible:** para poder contener nuevos tipos de metadatos.
- **Distribuida:** con el fin de que puedan ser referenciados metadatos de objetos externos.
- **Recursivo:** y así, permitir que los objetos sean tratados como contenedores de información y, por lo tanto, disponer de metadatos asociados.

La arquitectura, que fue denominada Warwick Framework (WF), se estructura en torno a dos componentes (**Dempsey y Weibel, 1996**):

- Paquetes (packages): conjunto de metadatos con una finalidad particular. Por ejemplo, un registro basado en el DC, o en MARC, se puede considerar un paquete. Pueden ser:
  - Primitivo (primitive o metadata set): los paquetes de metadatos propiamente dichos. Contienen las descripciones en su totalidad o en parte (MARC).
  - Indirecto (indirect): son referencias a otros objetos en la estructura de la información. Se vinculan entre sí por medio de URL (Uniform Resource Locator), URI (Uniform Resource Identifier) o URN (Universal Resources Name).
  - Contenedores (containers): un paquete que, a su vez, contiene otros paquetes.
- Contenedores (containers): unidades destinadas a albergar distintos conjuntos de descripciones. Los mecanismos para asociar un contenedor WF con el objeto contenido dependen del método de implementación. Existen dos tipos:
  - Referenciados internamente al objeto que se está describiendo. El metadato contenedor es aquel seleccionado por el autor o gestor como preferido para la descripción del objeto. Este metadato se asocia con el contenido embebiéndolo dentro de la estructura que mantiene el contenedor por medio de un URI.
  - Los referenciados externamente. El contenedor es el metadato creado y mantenido por una autoridad diferente del creador o gestor del contenido del objeto. El vínculo se realiza por medio de un mecanismo desarrollado para tal propósito (con objetivos y funcionamiento similares a los de la etiqueta LINK).

Este gráfico número 2 se muestra la mayoría de relaciones posibles. Los dos contenedores con referencias externas son independientes de los objetos (y de los paquetes que se utilicen). El contenedor referenciado internamente está incrustado dentro del objeto. Todos los contenedores se referencian entre sí por medio de sus correspondientes URIs, incluido el contenedor referenciado internamente, que se puede referenciar a sí mismo por medio de otro URI.

Los contenedores (como unidades para la anexión de conjuntos de metadatos -paquetes-) puede ser transitorios (transient, cuando existen como objetos de transporte entre repositorios) o persistentes (persistent, cuando se consideran como objetos de primera clase -first class objects- en una infraestructura de información). Esta última opción, tal y como apunta Lagoze (**Lagoze, 1996**), se produce cuando el

contenedor está almacenado en uno o más servidores y se encuentra accesible para todos los servidores por medio de un URI.

Independientemente del método de implementación utilizado, la única operación definida en el WF para el contenedor es aquella que le encarga de la devolución de los paquetes que contiene (sin especificar el orden de la secuencia, con el fin de no otorgar prioridades diferentes a los paquetes). De hecho, para el contenedor, cada paquete no es más que una cadena de bits, lo que significa que cada paquete se puede describir a sí mismo dentro del contenedor por su tamaño.

Se decidió que la implementación completa de la arquitectura debería estar vinculada a SGML. Por ese motivo se organizó un grupo de trabajo que estudiaría la posible DTD. Cualquier documento creado conforme a dicha DTD debería ser representado como un elemento contenedor (con la etiqueta <container>). Cada una de estas etiquetas está formada por una secuencia de uno o más de los siguientes paquetes: dublin core, package (con uno o más de un metadato o grupo de metadatos) y packageRef (un vínculo a otro paquete). Con la finalidad de ser más precisos, <container> permite la inclusión de una serie de entidades: name (obligatorio), URI y version (optativos ambos). Obviamente, los tres tienen la categoría CDATA<sup>8</sup> (Lagoze, 1996).

Por su lado, el elemento <package> comprendía únicamente dos subelementos: metadata (un metadato) y metaGroup (un grupo de metadatos). Los dos subelementos comparten los mismos atributos: type (referido al metadato o al grupo), scheme (esquema utilizado), show (si es visible o si esta cualidad depende de su antecesor), sortKwey (por si es necesario algún tipo de ordenación entre los metadatos) e index (si debe ser indizado, si no -noindex- o si debe tenerse con él la misma consideración que con el elemento paterno<sup>9</sup> -asparent-).

Una vez finalizado el DC2 se decide crear una lista de correo electrónico<sup>10</sup> en la que discutir sobre los diferentes aspectos, tanto sobre el conjunto de metadatos como de experiencias concretas. Es precisamente en esta lista donde se propone la creación de dos nuevas etiquetas, no contempladas en la primera versión del DC:

- Description: que equivaldría al campo "abstract" tal y como se encuentra especificado en el RFC 1866.
- Rights management: en principio podría adoptar tres valores:
  - Que el campo no figurase, lo que no significa obligatoriamente que no existan restricciones.
  - Sin restricciones de uso (especificado con la frase No restriction on reuse).
  - Y la etiqueta LINK que contiene la URL de otro documento con las condiciones de copyright. Este documento podría tener información compartida por varios registros. Lo que se pretendía con esto es que, al recuperar el documento indizado mediante DC, el buscador mostrara los vínculos relacionados con la gestión de derechos.

### **3.- CNI/OCLC Workshop on Metadata for Networked Images**

El tercer encuentro (DC3) se produjo del 24 al 25 de septiembre de 1996, de nuevo en Dublín (USA), pero en esta ocasión bajo el auspicio de la CNI (Coalition for Networked Information) y la OCLC, con la finalidad de extrapolar los resultados del DC1 (DC) y

DC2 (WF) válidos para la descripción de imágenes y bases de datos gráficas en entornos distribuidos.

Teniendo en cuenta que la descripción de información gráfica se escapa a los objetivos de este trabajo, obviaremos un análisis de lo producido en estas sesiones.

#### 4.- NLA/DSTC/OCLC Dublin Core Down Under

La OCLC se unió esta vez al DSTC (Distributed Systems Technology Center) y a la NLA (National Library of Australia) para organizar, esta vez en Canberra, el DC4 del 3 al 5 de marzo de 1997.

Los objetivos de esta nueva reunión, formada por 65 participantes, se centraron en definir la estructura formal definitiva de cada uno de los elementos y, a ser posible, la de sus respectivos cualificadores; completar los trabajos iniciados en el DC2 que tenían como finalidad el compatibilizar sistemas de metadatos para facilitar su interoperatividad y, por último, precisar aún más ciertos elementos que tienen una definición poco clara<sup>11</sup> (relation y coverage, especialmente) (DSTC, 1997).

Con respecto al primer objetivo, y tal y como lo describe Weibel (Weibel, et al., 1997), se formaron dos posturas encontradas. Por un lado se encontraron los “minimalistas”, que defendían la simplicidad de la estructura sintáctica creada. Para ellos, la aspiración de alcanzar la interoperatividad semántica entre diferentes sistemas de metadatos únicamente podría alcanzarse mediante el uso de un conjunto simple y comprensible de elementos que significasen la misma cosa en todos los casos. Los “estructuralistas” afirmaban, por su parte, que si bien es cierto que esta idea proporcionaba un mayor grado de flexibilidad, la tendencia semántica podría ser demasiado peligrosa con vistas a su futuro desarrollo.

En realidad ambas posturas se unieron más tarde, al darse cuenta de que el punto común que permitía la existencia de diferentes niveles de profundidad dentro de los metadatos se encontraba en los cualificadores. De hecho, el acontecimiento más destacable de esta reunión lo encontramos en la primera propuesta seria de cualificadores. Lo que se dio en llamar The Canberra Qualifiers (CQ).

##### 4.1.-The Canberra Qualifiers

Ya desde el DC1 resultó evidente la necesidad de establecer criterios que permitieran diferentes niveles de especificidad en la descripción de los recursos electrónicos.

El modelo propuesto en el DC se puede representar por el gráfico número 3 (Weibel, et al., 1997). Este esquema de trabajo incluye ya el concepto de Package, resuelto como elemento clave del WF. Este paquete puede incluir uno o varios metadatos. A su vez, éstos pueden estar formados por diferentes elementos que, a su vez también, tienen diversos sub-componentes propios. Resulta evidente que el dato central es el *Valor del elemento*. De tal forma que es ese valor el que se consideraría como identificador del recurso.

*Nombre del elemento* corresponde a un descriptor en concreto (por ejemplo, cualquiera de las 15 etiquetas del DC). El atributo *autoridad*, asociado con el *nombre*

*del elemento* y con el *esquema* de clasificación, se utilizó en esta ocasión para dejar patente que los cualificadores podían formar parte (o no) de un “vocabulario” controlado (por ejemplo, el conjunto de etiquetas del DC) avalados por una institución encargada de su mantenimiento (algo similar a lo que ocurre entre los LCSH y la Biblioteca del Congreso), independientemente de si son documentos válidos (aprobados por un validador de etiquetas HTML).

El DC4 formalizó los cualificadores tal y como se especifica en el punto 2.1 de este trabajo para SCHEME, pero añade algunas ligeras modificaciones al resto. Por un lado, incluye una nueva propiedad: LANGUAGE, que especifica el idioma del *valor del elemento* (no el recurso en sí mismo). Además, modifica el elemento TYPE de forma especial, al utilizar la codificación por puntos (dot-encoding) para especificar diferentes grados de jerarquías.

Resulta evidente que para que esta nueva sintaxis tenga éxito es necesario establecer una subdivisión con todos los posibles elementos que pueden ser susceptibles de ser incluidos en todos y cada uno de los metadatos del DC.

Para finalizar, se propusieron dos métodos para incluir los CQ dentro de HTML:

- **Overloaded Content Approach:** unir todos los cualificadores dentro del atributo CONTENT. Esta opción tenía la desventaja de ofrecer pocas posibilidades de recuperación, ya que el contenido útil del elemento estaba “tapado” por la información que de él previamente ofrecen SCHEME y LANG. Por otra parte, sería necesario establecer un orden de acción dentro de todos los elementos.
- **Additional Attribute Approach:** ofrece una forma más clara de representar la información, ya que los elementos se convierten en atributos del HTML. Estos atributos nuevos no supondrían problema alguno para los robots, pero sí para los parsers, lo que implica una contrariedad cuando se quiere hacer volcados de gran cantidad de información utilizando macros o cualquier otro sistema similar.

No se propuso solución alguna y se dejó para posteriores encuentros la decisión definitiva.

## **5.- The 5<sup>th</sup> Dublin Core Metadata Workshop**

La quinta reunión de trabajo, celebrada en Helsinki del 6 al 8 de octubre de 1997 y organizada por la OCLC, la NSF (National Science Foundation) y la CNI, comenzó con la noticia de la futura aprobación de la versión 4.0 de HTML. En la especificación, ratificada por el W3C en diciembre de ese mismo año, se añadirían los atributos SCHEME y LANGUAGE a la etiqueta META, por lo que el modelo de sintaxis que se utilizaría para el DC sería el denominado Additional Attribute Approach, propuesto en el DC4.

Al mismo tiempo se anunció la creación de otra especificación, el ISO 8601 profile (Wolf, 1997), que afecta al formato de fecha y hora. Esta propuesta de norma fue avalada por el W3C, y su uso implica no sólo la formalización de estas dos características en todos los atributos del HTML 4.0 susceptible de incluirlas, sino también dentro del propio DC.

## 5.1.- Definiciones problemáticas

Tal y como se dejó constancia en la anterior reunión (ver apartado 4), existía una problemática en cuanto a la definición, y posterior aplicación, de una serie de elementos en el DC. Como producto del esfuerzo del grupo de trabajo encargado de ello, así como fruto de la colaboración colectiva que suponía la lista de discusión meta2, se llegó a unas medidas de consenso que se ratificaron en el DC5. Los elementos en cuestión eran:

- **Date:** la principal problemática de este elemento derivaba de la gran cantidad de fechas que pueden ser significativas en el ciclo de vida de un recurso, y la presencia de una u otra depende del uso que se le quiera dar a ese metadato. Finalmente se adoptó la decisión de que este elemento debería contener la fecha asociada con la creación o disponibilidad del recurso. Dado que esta solución todavía seguía pareciendo ambigua, se propuso la elaboración de una lista con todos los tipos de fecha que se pueden utilizar y se recomendaba su uso como cualificador (**Weibel** y **Hakala**, 1998). Su utilización todavía dependía de la aprobación general.
- **Coverage:** su uso identifica un período de tiempo (referido al contenido) y a la cobertura espacial (región física).
- **Relation:** en realidad, tal y como afirma Miller (**Miller y Gill**, 1997), este problema no es nada nuevo, ya que viene heredado de los sistemas tradicionales de catalogación bibliográfica. La única novedad la encontramos en que en el mundo electrónico las variantes que pueden existir entre un recurso y otro son mucho más complejas. Es decir, existe una gran dificultad a la hora de suministrar metadatos a un recurso que es producto de otro, así como al original. Este tema ya ha sido analizado en otras reuniones dentro del ámbito bibliotecario (**Cromwell-Kessler y Erway**, 1997) (IFLA, 1998) y aquí, en el DC5, el resultado fue producto de lo que se determinó en llamar *el principio 1:1 (1:1 principe)*, que viene a afirmar que cada recurso ha de tener un metadato que lo describa, y cada uno de esos metadatos debe incluir elementos relativos a un único recurso. Por medio de este principio, especificar una relación incluye tres componentes: la identidad del recurso base u objeto (dado que existe el elemento identifier en el DC, no es necesario su repetición en la etiqueta relation), el identificador del recurso objetivo y, por último, un nombre que especifique esa relación entre los dos. A pesar de que a estas alturas todavía no se encontraba especificado el grupo de cualificadores para este elemento, se propusieron identificadores del tipo *IsVersionOf*, *IsBasedOn*, *IsPartOf...*, mientras el grupo de trabajo generaba una lista de posibles candidatos.

## 5.2.- Z39.50 y DC

Meses antes de la celebración del DC5, el ZIP (Z39.50 Implementors Group) anunció su intención de establecer equivalencias entre los 15 elementos del Dublin Core con la lista de atributos sobre los que se establece la posibilidad de definir ecuaciones de búsqueda, BIB-1. A partir de aquí, cualquier cliente Z39.50 puede utilizar los cualificadores y esquemas del DC en las estrategias de búsqueda.

En realidad, la única posibilidad de usar DC en consultas a través de Z39.50 pasaba por equiparar los elementos del DC con los atributos del tipo Use en BIB-1 a partir de la versión 2 de la norma y, dentro de la versión 3, generar un nuevo conjunto de atributos. Estos nuevos atributos, que deberían implementarse tal y como se define en el documento creado por la agencia de mantenimiento de la norma<sup>12</sup>, contemplan la

opción de trabajar con cualificadores y esquemas. Por lo tanto, las equivalencias en las búsquedas dentro de la versión 2 serían, tal y como lo expone LeVan (LeVan, 1998), quedan reflejadas en el gráfico número 4.

Con respecto a la versión 3, nos encontraríamos ante una primera lista de atributos Use que se correspondería con los 15 elementos del DC (Title, Author, Subject, Description, Publisher, OtherContributor, Date, ResourceType, Format, ResourceIdentifier, SourceIdentifier, Language, Relation, Coverage y RightsManagement). La segunda lista de atributos no podía ser completada hasta que no finalizaran los trabajos de los grupos encargados del desarrollo de los diferentes cualificadores y esquemas. Como a estas alturas se trabajaba con un borrador (tanto para el DC como para el ZIG), los elementos con los que se contaba eran todavía embrionarios (gráfico 5).

Por otra parte, la lista de posibles esquemas (SCHEMES) se reducía y convertía en una lista de posibles atributos complementarios. Eran los siguientes: ISO-8601, ANSI-X3.30, IETF-RFC-822, URL, URN, ISBN, ISSN, SICI, FPI (Formal Public Identifier).

### **5.3.- Primeros esfuerzos para conseguir una normalización “oficial”**

El importante seguimiento que tuvo el DC desde su comienzo provocó la necesidad de establecer normas con respecto a su utilización. El primer paso se dio el 10 de febrero de 1998 con la presentación de un borrador ante la IETF que, con el título *Dublin Core Metadata for Simple Resources Discovery* (Weibel, 1998), describía los objetivos que motivaron la creación del DC, el mecanismo para codificar el sistema de metadatos dentro de HTML, la posibilidad de incluir cualificadores o el tratamiento del DC dentro de RDF (Resource Description Framework). Ese borrador se convirtió en RFC con el número 2413, bajo el título *Dublin Core Metadata for Resource Discovery*, en septiembre de 1998<sup>13</sup>. Los elementos descritos en este texto constituyen la versión 1.0 del DC.

Al mismo tiempo se iniciaron las conversaciones con NISO (National Information Standard Organization) con el fin de formalizar el DC en una norma internacional.

Para finalizar, se aprueba la creación de varios grupos de trabajo. Entre ellos destaca el encargado del desarrollo de la especificación sobre la sintaxis para los subelementos del DC, el que estudiará la conexión con RDF y el que desarrollará las versiones internacionales del DC.

## **6.- The 6<sup>th</sup> Dublin Core Metadata Workshop**

En esta nueva reunión, celebrada del 2 al 4 de noviembre de 1998 en la Biblioteca del Congreso de los Estados Unidos, se analizaron diversos aspectos concernientes a la estructura organizativa de la iniciativa. Hasta la DC5, el DCMI funcionaba de manera informal. Con el fin de lograr un alcance mayor se desarrollaron los formalismos para llevar a cabo una estructura organizativa. A lo largo de 1998 se creó una Junta Directiva, un comité asesor PAC (Policy Advisory Committee) y otro comité técnico TAC (Technical Advisory Committee).

Además se crea una sede Web que centraliza toda la documentación generada por los diversos grupos de trabajo (<http://purl.oclc.org/metadata/dc> en la actualidad <http://dublincore.org>) así como diferentes listas de correo electrónico que se constituyen como el entorno de comunicación entre los diversos grupos de trabajo.

## 6.1.- Normalización

Tras el camino abierto por la RFC 2413, los objetivos se centraron en formalizar los elementos por medio de la normalización a través de la NISO para Estados Unidos y el CEN (Center for European Normalization) para Europa. Con ese fin se consideró oportuno realizar las modificaciones necesarias en el documento base, la RFC. Éstas se orientaban hacia una redefinición de todos los elementos, con el fin de aportar claridad y consistencia, y a rehacer la especificación de acuerdo con la ISO-11179<sup>14</sup>, norma internacional para expresar formalmente la semántica de los elementos de forma consistente.

Al mismo tiempo se siguió avanzando en la formalización de DC dentro del HTML por medio de un nuevo Internet Draft (**Kunze**, 1999) que complementaba los trabajos realizados por el W3C en la nueva especificación de HTML.

## 6.2.- Cualificadores

Tras diferentes estudios se determinó que el conjunto de posibles cualificadores propios del DC debería ser limitado. Por tanto, se recomienda el uso de esquemas externos (Dewey, Library of Congress Subject Headings, Medical Subject Headings e ISO 8601 para las fechas).

Existe, no obstante, un problema generado por el uso de estos esquemas: las posibles repercusiones que puede conllevar su empleo en pos de facilitar la interoperatividad entre diferentes sistemas de metadatos. Esta interoperatividad se ve especialmente afectada en dos aspectos. Por un lado en todo lo referido al intercambio de objetos/metadatos y, por otro, en lo relacionado con las búsquedas.

Indudablemente el caso más preocupante lo encontramos en lo referente al intercambio. Cualquier sistema que desee realizarlo de forma efectiva debe considerar, previamente, la necesidad de compartir una semántica con el otro sistema así como establecer una sintaxis y estructura de la información parecida (o al menos, producir algún mecanismo capaz de efectuar las correspondencias de manera eficaz). Para Weibel (**Weibel**, 1999) el ejemplo más claro de interconexión entre sistemas se puede encontrar entre el formato MARC y las AACR2 (Anglo-American Cataloging Rules 2<sup>nd</sup> edition), ya que existe un nivel de interoperatividad total entre ambos, lo que facilita el intercambio y uso de metadatos creados de acuerdo con estas normas.

En lo que respecta a la búsqueda, la coherencia no debe ser tanta. Únicamente se debe intentar establecer una semántica común así como un protocolo que facilite la comunicación.

## 7.- The 7<sup>th</sup> Dublin Core Metadata Workshop

De esta reunión, celebrada del 25 al 27 de octubre de 1999 en Frankfurt, no existe un documento oficial que recoja lo acontecido. Tan sólo podemos contar con las presentaciones realizadas por los diferentes grupos de trabajo en sus sesiones finales<sup>15</sup>. Todas las resoluciones tomadas por estos grupos con respecto a los cualificadores fueron convertidas, poco después, en un documento que, tras las consiguientes modificaciones, se convirtió en las recomendaciones que se pueden encontrar en el apartado 9 de este trabajo.

## 8.- Dublin Core Metadata Element Set

El 2 de julio de 1999 se hizo pública la versión 1.1 del Dublin Core Metadata Element Set (DCMES), último conjunto de metadatos del DC, que actualiza el RFC 2413 (ver apartado 5.3). Tal y como se propuso en el DC5, se utilizó la ISO 11179 como apoyo en la descripción de los elementos. Siguiendo esta norma, cada definición de los elementos de DC se refiere a la descripción de un recurso concreto tal y como se entiende recurso en el RFC 2396<sup>16</sup>. Los elementos son los siguientes:

Element Title	
Nombre	Título
Etiqueta	DC.Title
Definición	Nombre dado a un recurso.
Comentario	Típicamente se refiere al nombre por el que el recurso es conocido.

Element Creator	
Nombre	Creador
Etiqueta	DC.Creator
Definición	Entidad responsable de realizar el contenido del recurso.
Comentario	Ejemplos: una persona, organización o servicio responsables del contenido intelectual del recurso.

Element Subject	
Nombre	Materia y palabras clave
Etiqueta	DC.Subject
Definición	El tema del contenido del recurso.
Comentario	Por lo general se expresa en forma de palabras clave o códigos de clasificación que puedan describir la materia del recurso. Se recomienda el uso de un vocabulario controlado para extraer los términos.

Element Description	
Nombre	Descripción
Etiqueta	DC.Description
Definición	Resumen del contenido del recurso.
Comentario	Suele incluir un resumen, índice de contenidos, una referencia a la representación gráfica del contenido...

Element Publisher	
Nombre	Editor
Etiqueta	DC.Publisher
Definición	Entidad responsable de crear el recurso disponible.

Element Contributor	
Nombre	Colaboradores
Etiqueta	DC.Contributor
Definición	Entidad responsable de contribuir al desarrollo del contenido del recurso.

Element Date	
Nombre	Fecha
Etiqueta	DC.Date
Definición	Fecha asociada con un evento a lo largo del ciclo de vida del recurso.

Element Type	
Nombre	Tipo de recurso
Etiqueta	DC.Type
Definición	La naturaleza o género del contenido del recurso.
Comentarios	Incluye también categorías, funciones... Se recomienda el uso de los tipos propuestos por el grupo de trabajo correspondiente.

Element Format	
Nombre	Formato
Etiqueta	DC.Format
Definición	Representación física o digital del recurso.
Comentario	Puede ser utilizado para especificar los requerimientos hardware y software necesarios para la manipulación del recurso. También puede usarse para describir datos como dimensión, duración, tamaño...

Element Identifier	
Nombre	Identificador del recurso
Etiqueta	DC.Identifier
Definición	Secuencia de caracteres que permita identificar unívocamente un recurso.
Comentario	Uniform Resource Identifier (URI), Digital Object Identifier (DOI), International Standard Book Number (ISBN)...

Element Source	
Nombre	Fuente
Etiqueta	DC.Source
Definición	Identifica el trabajo del que proviene el recurso actual.

Element Language	
Nombre	Idioma
Etiqueta	DC.Language
Definición	Idioma en el que está expresado el contenido del recurso.

Element Relation	
Nombre	Relación
Etiqueta	DC.Relation
Definición	Referencia a un recurso relacionado con el actual

Element Coverage	
Nombre	Cobertura

Etiqueta	DC.Coverage
Definición	Cobertura especial (region física) y/o temporal (referida al contenido) del recurso.

Element Rights	
Nombre	Gestión de derechos de autor
Etiqueta	DC.Rights
Definición	Información sobre los derechos de autor que afectan al recurso.

## 9.- Dublin Core Qualifiers (qDC)

El 17 de abril de 2000 se aprobaron los cualificadores definitivos para los diferentes elementos del DCMES (Weibel, 2000), y se hicieron públicos en julio de ese mismo año (DCMI, 2000). Se trata de una lista no cerrada que formaliza el método de utilización. Se decidió aceptar tanto cualificadores y esquemas conocidos ampliamente por la comunidad científica como otra serie de términos que, por su carácter local o especializado, son desconocidos para gran cantidad de posibles usuarios del sistema. Por este motivo es posible que se de el caso de que el programa cliente no sea capaz de reconocer el cualificador. Esto se puede deber a dos causas: o que únicamente esté configurado para soportar los 15 elementos DCMES en su forma más básica, o bien que se encuentre un cualificador concreto no comprendido por el cliente.

Para que pueda comprenderse el significado completo de un valor codificado se requiere que el cliente sea capaz de interpretar la notación. Si, por el contrario, el cliente no puede, aún resultará posible procesar el valor localizado en el atributo CONTENT mientras ignora a los atributos "conflictivos", es decir, SCHEME y LANG<sup>17</sup>. Todos los elementos refinados sin cualificador se convierten en texto perteneciente al atributo NAME. El principio que se pretende inculcar es el denominado Dumb-Down, es decir, que un programa cliente pueda ignorar cualquier cualificador y utilizar su descripción como si no tuviese cualificador.

Se reconoce que se pierde especificidad pero, por el contrario, el valor del elemento recuperado (sin el cualificador) puede utilizarse para la localización del recurso. El DCMI clasifica los cualificadores en dos grupos:

- Limitadores (element refinement): hace que el significado de un elemento sea más genérico o específico. Este tipo de elementos comparte el significado del elemento que no cuenta con cualificador pero con un alcance más restrictivo. Cualquier cliente que no sea capaz de comprender este elemento puede ignorar el cualificador y trabajar con el valor del metadato como si se tratase de un elemento genérico, sin calificar.
- Codificadores (encoding scheme): permite identificar el esquema utilizado para asignar el valor al elemento. Este esquema puede ser un vocabulario controlado, una notación formal o una regla de validación. Si el cliente no es capaz de interpretar el sistema de codificación, el valor asignado al elemento puede ser comprendido por el usuario.

Algunas de las propiedades de los cualificadores difieren respecto a las definiciones establecidas en la versión 1.1 del Dublin Core Metadata Element Set (ver apartado 9) debido, principalmente, a un intento por acercar la terminología a la utilizada en el XML

(eXtensible Markup Language) y, sobre todo, promover una integración más sencilla de los esquemas DC hacia entornos RDF.

Dado que los elementos refinados no son soportados directamente por la etiqueta META de HTML, deben añadirse al nombre del elemento DCMES separados por un punto y almacenados como si se tratase del atributo NAME, tal y como lo indica el RFC 2731 (**Kunze**, 1999).

Por su parte, el valor dependiente de un sistema de codificación dado sí tiene su correspondencia en la etiqueta META de HTML por medio de los atributos SCHEME y LANG.

El modelo de referencia 2.0 del DC estaría formado -además de por los elementos que constituyen el modelo 1.0- por:

- Elemento cualificador: añade atributos especiales al elemento con el fin de conseguir una mayor especificidad dentro de la descripción del recurso.
  - Tipo (type): define la semántica del Elemento (por ejemplo, título alternativo -DC.Title.Alternative- añadiría más cualidades al elemento DC.Title).
- Valor del cualificador: añade nuevos atributos al valor del elemento, con el fin de refinar más su descripción. Existen cuatro formas posibles:
  - Componente: representa un aspecto estructural del valor el elemento.
  - Vocabulario: define un vocabulario específico del cual se han extraídos términos para aplicarlos como valor del elemento (tesauro, diccionario, taxonomía...).
  - Codificación (Scheme): establece un sistema de codificación utilizado para representar el valor del elemento (ISO XXX, CDU...).
  - Idioma (Lang): representa el idioma en el que está expresado el valor del elemento.

Por lo tanto, la modificación del esquema del modelo de referencia 1.0 teniendo en cuenta todos estos aspectos produciría como resultado la siguiente representación (gráfico 6).

Para que este modelo de referencia tenga coherencia, es necesario que los elementos se puedan agrupar juntos (aunque no es obligatorio, sí se recomienda que el orden sea el especificado en por el DCMES y qDC) y que, en caso de utilizarse cualificadores de elementos se haga constancia de ellos, con el fin de poder interpretar correctamente los valores de los cualificadores.

## 10.- The 8<sup>th</sup> International Dublin Core Metadata Workshop

El 3 de octubre de 2000 comenzó la octava edición de la serie del DCMI. Durante los tres días siguientes unas 150 personas se dieron cita en Ottawa para dar a conocer los principales logros a los que habían llegado los diversos grupos de trabajo de la iniciativa.

En esta reunión se pusieron de manifiesto diversas necesidades (**Weibel** y **Koch**, 2000):

- Documentar de forma más exhaustiva los cualificadores. Teniendo en cuenta que la conversión en norma ANSI de los DCMES estaba cerca, era necesario empezar a trabajar con una segunda especificación que permitiera la anexión de los distintos cualificadores aprobados el 17 de abril de ese mismo año (ver apartado 9) con el fin de generalizar su uso.
- Mostrar con más claridad que la flexibilidad del conjunto de etiquetas del DCMI permite la representación de cualquier tipo de objeto. En este sentido, destaca la actuación del DC Education working group, agrupación formada a mediados de 1999 por personas relacionadas con el campo de la educación, que tiene como principal objetivo el adaptar el conjunto de elementos DC a las necesidades de representación de la información de este colectivo especializado. Uno de sus mayores logros fue la creación del DC Education Element Set (**DC Education**, 2000), un conjunto de metaetiquetas basadas en Dublin Core.
- Desarrollar un sistema que permita registrar e informar de determinados aspectos relacionados con el desarrollo de DCMI. De esta forma, se propone llevar un control sobre las diferentes iniciativas que se llevan a cabo en torno al DCMI así como establecer un mecanismo para acceder recursos de personas o instituciones que gestionen proyectos basados en este conjunto de metadatos.

Junto a esto, se presentaron las conclusiones a las que llegaron los numerosos grupos de trabajo que se encargan de discutir aspectos particulares en el devenir del Dublin Core y se presentó EOR (Extensible Open RDF), un programa informático desarrollado por la OCLC basado en Java que permite la creación, gestión y mantenimiento de bases de datos originadas sobre la base de la metainformación en formato RDF y DC.

El 24 de octubre de 2001 se celebró en Tokio un evento, el DC-2001: International Conference on Dublin Core and Metadata Applications (**Oyama y Gotod**, 2001), que, para los organizadores puede considerarse como el noveno de la serie de las reuniones de trabajo del DCMI. En realidad no se trata más que de un congreso internacional en el que se presentaron gran cantidad de ponencias y comunicaciones sobre posibles desarrollos de DC (algunas de ellas de gran interés por las posibilidades de implementación que presentan). Teniendo en cuenta que este acontecimiento no ha supuesto ningún avance en el desarrollo de la, ya norma, DC, no hemos considerado oportuno profundizar más en el tema.

## 11.- Norma Z39.85

El logro más importante, desde el punto de vista de la normalización, tuvo lugar el 10 de septiembre de 2001, cuando la ANSI norteamericana aprobó definitivamente el documento que reconocía el conjunto de etiquetas DCMES como norma internacional con la denominación ANSI/NISO Z39.85-2001 (**ANSI**, 2001).

En palabras de Weibel (**Weibel y Koch**, 2000), el esfuerzo del grupo de trabajo CEN (ver apartado 6.1) unido al estatus de norma recién logrado, debería facilitar el camino para la consecución de una norma ISO que aglutinara, tanto al conjunto de metadatos DC como a los cualificadores. Y es que hay que recordar que Z39.85 sólo formaliza las 15 etiquetas originales, y no los cualificadores. Esto se debe a que la petición de normalización es anterior al desarrollo definitivo de los cualificadores.

La norma, que tan sólo explica brevemente la historia del DC, comenta sus objetivos principales y enumera todas y cada una de las etiquetas, designa como agencia de mantenimiento al Dublin Core Metadata Initiative, ya que en esas fechas, se había constituido como organización con capacidad para ello<sup>18</sup>.

## 12.- Congresos internacionales

Conforme crecía el uso de las etiquetas DC y se ampliaba la lista de instituciones y proyectos que se embarcaban en su uso, aumentó la necesidad de convertir las viejas sesiones de trabajo en eventos de mayor calado: los congresos internacionales. Así, en octubre de 2001 se celebra en Tokio el DC-2001 (que se correspondería con el DC9), denominado International Conference on Dublin Core and Metadata Applications.

De entre todo lo que se trató, destacamos el estudio de perfiles de aplicación de las etiquetas y, especialmente, el análisis que se realizó para incluir DC dentro de RDF (**Dekkers y Weibel, 2002**).

Un año más tarde, pero esta vez en Florencia, se celebra el DC-2002 (DC10) bajo el título Metadata for e-Communities: Supporting Diversity and convergente. En él se tocan temas referentes a los mecanismos para expresar DC en XML, la redefinición de diferentes cualificadotes (especialmente para las etiquetas Creator, Contributor y Publisher) así como mecanismos para realizar citaciones dentro de textos basándose en DC y la realización de diversas guías que permitieran una mayor comprensión de este sistema de metadatos.

## 13.- Dublin Core y MARC

Desde el comienzo del DCMI se han realizado numerosos intentos de establecer un mecanismo que permita convertir registros de DC a MARC y viceversa. Las instituciones más involucradas han sido la Biblioteca del Congreso y la ALA (American Library Association).

En enero de 1993 la Biblioteca del Congreso plantea, a través de la *Proposal 93-4*, la creación del campo 856 (Electronic Location and Access) para la identificación, localización y acceso a recursos en línea, como resultado de la interacción con el OCLC Internet Resources Project.

Este trabajo en común también se originó a raíz de la aparición del DC y los primeros Workshops sobre metadatos organizados por OCLC. Desde el DC1 hasta 1997 se produjeron diversos documentos internos (Discussion Paper) a través de los cuales se invitaba a la reflexión en este sentido. Gracias a estos escritos se pudo avanzar hacia una equivalencia<sup>19</sup> (mapping) coherente entre DC y USMARC.

Casi al mismo tiempo la ALA creó, en 1998, un grupo de trabajo denominado CC:DA TFCR (Committee on Cataloging: Description and Access Task force on Metadata Cataloging Rules), que tenía, entre otros objetivos, el establecer una relación entre diversos sistemas de metadatos (TEI -Text Encoding Initiative-, DC y EAD -Encoded Archival Description-) y el formato USMARC. Si bien es cierto que los resultados de los grupos de trabajo de la ALA (CC:DA TF y CC:DA TFCR) no han sido demasiado

halagüeños -Weibel echó en cara varias veces el comportamiento defensivo así como la actitud nada innovadora de los miembros del grupo (**American Library Association**, 2000)-, se produjeron interesantes puntos de encuentro entre varios conjuntos de metadatos y el formato MARC. Destaca especialmente el avance que supuso la equivalencia que se estableció entre TEI y MARC y viceversa<sup>20</sup>.

Algunos de los primeros prototipos llevados a cabo utilizando el DC han contemplado la opción de convertir documentos de este formato a MARC. Para ello han utilizado diversos programas, en su mayoría escritos en PERL, que se encargan de ejecutar esta acción. Entre ellos destaca *d2m*<sup>21</sup>, enmarcado dentro del Nordic Metadata Project, que facilita la conversión de metadatos del DC a cualquiera de las versiones nórdicas del formato MARC.

## 14.- A modo de conclusión: el caso español

Al contrario de lo que ha sucedido en España, el conjunto de etiquetas DC se ha extendido de forma vertiginosa por varios países, destacando especialmente las áreas geográficas de Oceanía y los países nórdicos. Se puede acceder a una lista actualizada de estos proyectos en: <http://dublincore.org/projects/>

En el entorno nacional, la mirada de apatía y/o escepticismo con la que la mayoría de profesionales de la información se ha fijado en todos los sistemas de metadatos en general, y en Dublin Core en particular, ha frenado su posible uso y propagación. Salvo contadas excepciones, ninguna biblioteca española usa DC para catalogar recursos electrónicos (para eso está MARC, como diría un castizo) y son pocos los proyectos que se valen de este sistema para realizar la descripción de cualquier objeto electrónico, si bien es cierto que sí usan otros mecanismos, como RDF (en varios proyectos realizados por la editorial Anaya, por ejemplo) o variaciones sobre la base SGML (como algunos trabajos del Instituto Cervantes).

Uno de los proyectos más avanzados, desde el punto de vista de la tecnología utilizada, que usa DC es Sarac (Servicio de Acceso a Recursos de Alta Calidad). Idea fraguada entre RedIRIS y la Universidad de Granada que trata de ofrecer un buscador de recursos caracterizados por su alta calidad informativa. Para lograr ese valor añadido se sirven de un conjunto de metaetiquetas propias y otras de DC que, enmarcadas dentro de RDF, proporcionan descripciones muy exhaustivas de recursos electrónicos seleccionados por expertos nacionales de diferentes áreas de conocimiento. Puede conseguirse más información en: <http://sarac.rediris.es>

Al igual que sucede con la etiqueta META del HTML, DC tiene que combatir contra una serie de factores muy importantes:

- Editores de texto. Ninguno de los editores profesionales que se comercializan en la actualidad permite la inclusión de metadatos en las páginas HTML. Esto hace que los interesados en incluir este tipo de información en sus documentos tengan que acudir a programas complementarios o, en el mejor de los casos, plantillas que se encarguen de la conversión de la información en un formato de metadatos determinado.

- Curva de aprendizaje. La complejidad que tienen las diferentes etiquetas DC, la aparición de los cualificadores, la dificultad para introducir este sistema dentro de RDF... Un análisis de las etiquetas DC propuestas en 1995 y las actuales revela que

la complejidad para implementarlas se ha multiplicado por tres. Por el contrario, siguen sin existir guías de referencia que expliquen el uso correcto de las etiquetas y de los cualificadores o las diversas y complejas casuísticas que se pueden dar dentro del mundo electrónico.

- Instituciones implicadas en la introducción de metadatos. A decir verdad, han sido pocos los autores que hayan propuesto soluciones suficientemente eficaces para este problema. Entre otras cosas, y muy probablemente, porque no se trata sólo de un aspecto técnico o de un problema tecnológico, como se apuntaba anteriormente, sino de concienciación sobre la necesidad de las cosas.

## 15.- Notas

1. El grupo de trabajo no se detuvo en la definición exacta de DLO. Se entendía, por ejemplo, que la versión electrónica de un periódico, o de un diccionario, sí era un DLO, mientras que una colección de diapositivas no catalogadas no se podría considerar como tal. Lo que sí se afirmaba es que principalmente se refería a texto. Este concepto ha sido modificado con posterioridad (*Chilvers y Feather, 1998*) (*Drewry, et al., 1997*).

2. Algo de lo que más tarde se han beneficiado otros sistemas orientados a la descripción de imágenes o de localizaciones geográficas como GILS —Government/Global Information Locator Service— o FGDC —Federal Geographic Data Committee — Content Standard for Digital Geospatial Metadata.

3. En esta primera reunión se decidió no incluir etiquetas referentes al coste de creación del documento o la información referente a la ubicación física del objeto.

4. Concretamente se referían, por un lado, a los índices generados por robots como Lycos y WebCrawler (que aparecieron poco tiempo antes y que tuvieron gran impacto dentro de la profesión). El otro formato al que se referían era MARC. Para el DC1, los primeros eran archivos demasiado costosos de mantener y contenían poca información útil, mientras que los segundos no tenían cabida en Internet debido a su complejidad extrema, por lo que DC podría significar el término medio entre estos extremos.

5. Como se puede observar en los ejemplos incluidos a continuación, todavía no se había resuelto cómo embeber este primer conjunto de elementos dentro de HTML, de ahí que no se consideren todavía como etiquetas HTML sino como representaciones de lo que se pretendía.

6. El método propuesto (comparación del tamaño en bits del fichero) no es válido, ya que un mismo documento puede tener diferente tamaño dependiendo del formato en el que se haya almacenado.

7. MIME (Multipurpose Internet Mail Extension), cuyo funcionamiento está regulado por RFC 1590, sirvió en su origen como una forma normalizada de intercambio de mensajes electrónicos multimedia.

8. Se utiliza para especificar al parser (analizador sintáctico) SGML que ignore un determinado conjunto de etiquetas.

9. En SGML se puede establecer una relación padre—hijo entre las etiquetas, de tal forma que es factible el hecho de “transmitir” unas características de un padre a un hijo.

10. Se denominó META2 (*meta2@mrrl.lut.ac.uk*). En la actualidad se conoce como dc—general (*dc—general@mailbase.ac.uk*).

11. Dado que no se consiguió alcanzar ningún punto de consenso dentro de las futuras definiciones, así como acerca de cuáles serían los sub—elementos que formarían cada uno de los esquemas de trabajo, se decidió crear varios grupos ad hoc que se encargarían del desarrollo de estos conceptos.

12. Z39.50 Attribute Architecture. <http://lcweb.loc.gov/z3950/agency/attrarch/attrarch.html>

13. <http://www.ietf.org/rfc/rfc2413.txt>

14. ISO—11179. Specification and Standardization of Data Elements. <ftp://sdct—sunsrv1.ncsl.nist.gov/x318/11179/>.

15. <http://www.ddb.de/partner/dc7conference/results.htm>

16. En este RFC, que define los URI, se define como "cualquier cosa que tenga identidad".  
<http://www.ietf.org/rfc/rfc2396.txt>

17. Esto se puede hacer, entre otras cosas, con ayuda de por la forma de actuar de la mayoría de sistemas de búsqueda, que permite localizar la cadena de caracteres que constituye la consulta independientemente del valor que ésta tenga asociado dentro de un sistema de notación concreto.

18. La primera intención es que fuera la Biblioteca Nacional finlandesa la que encabezara la agencia de mantenimiento.

19. En términos informáticos se denomina mapping o crosswalk al hecho de establecer equivalencias entre unos sistemas y otros.

20. <http://etext.lib.virginia.edu/tei—usmarc.html/>

21. <http://www.bibsys.no/meta/d2m/>

## 16.- Bibliografía

DC Education Set Proporsal. [Página Web] 5 noviembre 2000. Consultado en: 15 2002.  
<http://dublincore.org/documents/wd/education-20001005.htm>

**American Library Association.** Association for Library Collections and Technical Services (a division of the American Library Associatin) Cataloging and Classification Section. Committee on Cataloging: Description and Access Task Force on Metadata. [Página Web] 28 agosto 2000.  
<http://www.ala.org/alcts/organization/ccs/ccda/tf-meta1.html>

**ANSI/NISO.** *ANSI/NISO Z39.85-2001. The Dublin Core Metadata Element Set.* 10 septiembre 2001. <http://www.niso.org/standards/resources/Z39-85.pdf>

**Burnard, Lou (et al.).** A Syntax for Dublin Core Metadata: Recommendations from the Second Metadata Workshop . [Página Web]  
<http://purl.org/dc/workshops/dc2conference/report19960401.htm>

**Caplan, Priscilla y Guenther, R.** "Metadata for internet resources: the dublin core metadata elements set and its mapping to usmarc.". En: *Cataloging and classification quarterly*, 1996, v. 22, n. 3/4, pp. 43-58.

**Chilvers, Alison y Feather, John.** "The management of digital data: a metadata approach". En: *Electronic library*, diciembre 1998, v. 16, n. 5, pp. 335-371.

**Cromwell-Kessler, Willy y Erway, Ricky.** Research Libraries Group Metadata Summit: meeting report. [Página Web] julio 1997. <http://www.rlg.org/meta9707.html>

**DCMI.** Dublin Core Metadata Initiative - Home Page. [Página Web] 10 octubre 1998.  
<http://purl.org/DC/index.htm>

**DCMI.** Dublin Core Qualifiers. [Página Web] 11 julio 2000.  
<http://purl.org/dc/documents/rec/dcmes-qualifiers-20000711.htm>

**Dekkers, Makx y Weibel, Stuart L.** "Dublin Core Metadata Initiative Progress Report and Workplan for 2002". En: *D-Lib Magazine*, febrero 2002, volumen 8, número 2.  
<http://www.dlib.org/dlib/february02/weibel/02weibel.html>

**Dempsey, Lorcan y Weibel, Stuart.** "The Warwick Metadata Workshop: A Framework for the Deployment of Resource Description". En: *D lib magazine*, julio-agosto 1996. <http://www.dlib.org/dlib/july96/07weibel.html>

**Drewry, Marilyn (et al.).** "Metadata: quality vs. quantity". En: *Second IEEE metadata conference*; Maryland. 1997. <http://computer.org/conferen/proceed/meta97/papers/hconover/mdrewry.html>

**DSTC.** DC4 Home Page. [Página Web], 1997. <http://www.dstc.edu.au/DC4>

**FGDC.** Federal Geographic Data Committe. [Página Web]. <http://www.fgdc.er.usgs.gov/index.html>

**GILS.** GILS (Global/Government Information Locator Service). [Página Web] <http://www.gils.net>

**Gimeno Montoro, María José; Barrueco Cruz, José Manuel y García Testal, Cristina.** "Catalogación de recursos electrónicos accesibles en Internet: revisión de propuestas para una normativa". En: *VI Jornadas Españolas de Documentación. Fesabid'98*; Valencia. 1998.

**IFLA.** Functional Requirements for Bibliographic Records. [Página Web]. <http://www.ifla.org/ifla/VII/s13/frbr/frbr-toc.htm>

**Kunze, J.** RFC 2731: Encoding Dublin Core metadata in HTML. [Fichero TXT] diciembre 1999. <ftp://ftp.isi.edu/in-notes/rfc2731.txt>

**Kunze, John.** *Encoding Dublin Core Metadata in HTML*. 18 marzo 1999. <http://www.ietf.org/internet-drafts/draft-kunze-dchtml-00.txt>

**Lagoze, Carl .** "The Warwick Framework: a container architecture for diverse sets of metadata". En: *D lib magazine*, julio-agosto 1996. <http://www.diglib.org/dlib/July96/07lagoze.html>

**Lassila, Ora y Swick, Ralph R.** *Resource Description Framework (RDF) Model and Syntax Specification. W3C Recommendation*. 22 febrero 1999. <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222>

**LeVan, Ralph.** Dublin Core and Z39.50. [Página Web] 2 febrero 1998. <http://purl.org/dc/documents/note~2.html>

**Library of Congress.** *Proposal nº 96-2. Define a Generic Author Field in the Bibliographic, Authority, Classification, and Community Information Formats*. <http://lcweb.loc.gov/marc/marbi/list-p.html>

**Library of Congress.** *Discussion Paper nº 86. Mapping the Dublin Core Metadata Elements to USMARC*. 5 mayo 1995. <http://lcweb.loc.gov/marc/marbi/dp/dp86.html>

**Library of Congress.** *Discussion Paper nº 99. Metadata, Dublin Core and USMARC: a review of current efforts*. 21 enero 1997. <http://lcweb.loc.gov/marc/marbi/dp/dp99.html>

**Miller, Paul y Gill, Tony.** "Metadata corner. DC5: the search of Santa". En: *Ariadne*, diciembre 1997. <http://www.ariadne.ac.uk/issue12/metadata>

**Ortiz-Repiso Jiménez, Virginia.** "Nuevas perspectivas para la catalogación: metadatos versus MARC". En: *Revista Española de Documentación Científica*, 1999, v. 22, n. 2, pp. 198-219.

**Oyama, Keizo y Gotod, Hironobu.** DC-2001 Proceedings of the International Conference on Dublin Core and Metadata Applications 2001. [Página Web] 24 octubre 2001. Consultado en: Mar. 2003. <http://www.nii.ac.jp/dc20001>

**San Segundo Manuel, Rosa.** "Organización del conocimiento en Internet: metadatos bibliotecarios Dublin Core". En. *VI Jornadas Españolas de Documentación. Fesabid'98*; Valencia. 1998.

**Weibel, S.;Kunze, J. y Lagoze, C.** *Dublin Core Metadata for simple Resource Discovery*. 10 febrero 1998. <http://ftp.sunet.se/pub/Internet-drafts/draft-kunze-dc-02.txt>

**Weibel, Stuart.** "Metadata: the foundations of resource description". En: *D lib magazine*, julio 1995. <http://www.diglib.org/dlib/July95/07weibel.html>

**Weibel, Stuart.** A proposed convention for embedding metadata in HTML. [Página Web] 2 junio 1996. <http://purl.org/dc/workshops/dc2conference/resources-weibel19960602.htm>

**Weibel, Stuart.** "The state of the Dublin Core Metadata Initiative April 1999". En: *D lib magazine*, abril 1999. <http://www.dlib.org/dlib/april99/04weibel.html>

**Weibel, Stuart** (weibel@oclc.org). Approval of initial Dublin Core Interoperability Qualifiers. E-mail a: dc-general (dc-general@mailbase.ac.uk). 2000 17.

**Weibel, Stuart y Hakala, Juha.** DC-5: The Helsinki Metadata Workshop: A Report on the Workshop and Subsequent Developments . [Página Web] 1998. <http://www.dlib.org/dlib/february98/02weibel.html>

**Weibel, Stuart; Iannella, Renato y Cathro, Warwick.** "The 4th Dublin Core metadata workshop report". En: *D lib magazine*, junio 1997. <http://www.dlib.org/dlib/june97/metadata/06weibel.html>

**Weibel, Stuart L. y Koch, Traugott.** "The Dublin Core Metadata Initiative: mission, current activities, and future directions". En: *D lib magazine*, diciembre 2000, v. 6, n. 12. <http://www.dlib.org/dlib/december00/weibel/12weibel.html>

**Wolf, Misha y Wicksteed, Charles.** *ISO-8601-PROFILE. Date and Time Formats*. <http://www.w3.org/TR/NOTE-datetime-970915>

**Wu, C. J.** "Metadata and future developments in cataloguing". En: *Journal of information, communication, and library science*, 1997, v. 4, n. 2, pp. 42-52.

Gráficos:

Dublin Core – GRAFICO1.gif corresponde con el gráfico 1 (en fichero aparte)

Dublin Core – GRAFICO2.gif corresponde con el gráfico 2 (en fichero aparte)

Dublin Core – GRAFICO3.gif corresponde con el gráfico 3 (en fichero aparte)

Dublin Core – GRAFICO4.gif corresponde con el gráfico 4 (en fichero aparte)

Dublin Core – GRAFICO5.gif corresponde con el gráfico 5 (en fichero aparte)

Dublin Core – GRAFICO6.gif corresponde con el gráfico 6 (en fichero aparte)

Por favor, a la hora de maquetar el artículo poned, debajo de cada gráfico, la leyenda que corresponda: Gráfico 1, Gráfico 2, Gráfico 3, Gráfico 4, Gráfico 5, Gráfico 6. Gracias.