

# CATALOGACIÓN/ORGANIZACIÓN DE DOCUMENTOS DIGITALES: ESTADO DE LA CUESTIÓN, TENDENCIAS Y PERSPECTIVAS DESDE ESPAÑA

*Eva M<sup>a</sup> Méndez Rodríguez*

Dpto. Biblioteconomía y Documentación  
Universidad Carlos III de Madrid  
<[emendez@bib.uc3m.es](mailto:emendez@bib.uc3m.es)>

## Resumen:

Este artículo refleja el estado de la cuestión de la organización y recuperación de documentos digitales. Se revisan las iniciativas de estandarización internacional sobre catalogación de recursos electrónicos en sentido tradicional (desde las ISBD (ER) hasta la posición de MARC21, frente a la información digital y los nuevos estándares de representación de la información como XML), así como las normas relacionadas con el nuevo concepto de metainformación, destacando su concepción disciplinar. Se analizan también tendencias, iniciativas, proyectos y perspectivas de la gestión y organización de la información electrónica en nuestro país, prestando especial atención al ámbito informativo catalán, así como a los retos que presenta para España, la interoperabilidad entre los distintos sistemas y servicios de información nacional y la perspectiva de interoperabilidad global fundada en metadatos, protocolos y convenciones.

## Descriptores:

Objetos de información digital / Catalogación de Internet / Metadatos / Organización de recursos electrónicos / Estándares para la descripción / España / Cataluña.

## 1. Introducción: de “procesos técnicos” a “procesos tecnológicos”

*I'm all for progress; it's change I don't like.*  
Mark Twain

En los primeros años del nuevo milenio que hemos empezado, vivimos con entusiasmo la *revolución tecnológica*, *la sociedad de la información*, *la sociedad red...* y otras denominaciones plausibles que llenan la boca de iniciados y profanos, sin que nadie sepa ciertamente adónde nos lleva esta vertiginosa, a veces incierta, combinación de tecnologías informáticas, avances en las telecomunicaciones y cambios en los procesos y formas de comunicación.

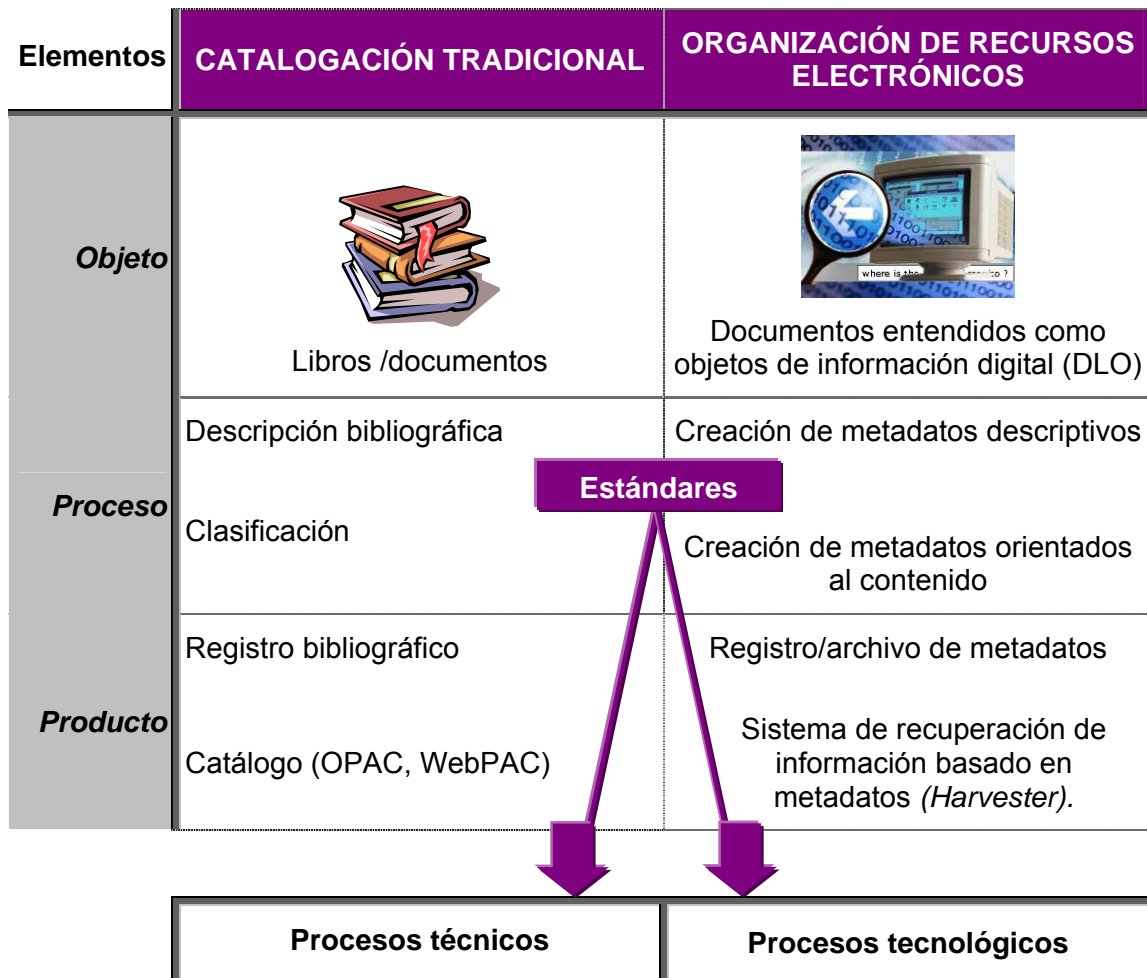
A finales del siglo pasado, todas las contribuciones científicas en el ámbito de la Biblioteconomía tenían un tono prospectivo y e incluso futurible, como si la gestión de la información digital fuese a cambiar sus presupuestos, principios y tendencias de la noche a la mañana, al comenzar el año 01, y documentalistas y bibliotecarios esperábamos esa proyectada mutación kafkiana en informáticos. Entrado ya el nuevo milenio y, habida cuenta de que las necesidades de información de los usuarios son las mismas aun cuando la satisfacción de esas necesidades cambia continuamente por la desintermediación creciente que provocan las tecnologías, los profesionales de la información hemos aprendido a vivir y trabajar en ese proceso de aceleración constante en la generación, difusión, uso y obsolescencia del conocimiento. Estamos aprendiendo a adaptar los procesos técnicos, en los que se basa el tratamiento de la información, a los procesos tecnológicos en los que se funda actualmente el desarrollo social, intelectual, e incluso económico.

Las representaciones que han evidenciado tradicionalmente la innovación tecnológica, como los ciclos de Kondratieff o la *S de Foster*<sup>1</sup>, se ven superadas por la sublimación del factor tiempo que es el que da validez al conocimiento en la sociedad de las TICs, sobre todo ante la súbita evolución de Internet y de la Web. En este contexto de violenta innovación, partiendo de la clave del progreso tecnológico que prevalece desde que el hombre aplica la informática a la gestión de la información, la documentación digital, la información para la Web difiere sustancialmente en términos de variedad, volumen creciente y volatilidad a la que está sujeta, de la documentación tradicional impresa. Todos estos cambios están cuestionando los métodos tradicionales de procesamiento de la información (sobre todo, catalogación, clasificación e indización), o al menos, han hecho necesaria su revisión y adaptación, debido al carácter horizontal de las TICs, promotoras del desarrollo de otras muchas tecnologías en torno a los documentos digitales entendidos como objetos<sup>2</sup>, a los procesos de conocimiento y a la recuperación de recursos en nuevos entornos informativos multidimensionales.

El conjunto de tareas tradicionales tales como la selección, adquisición, descripción y conservación de los registros de conocimiento, conocidas en la Biblioteconomía tradicional como “procesos técnicos”, se someten en la actualidad a un complejo panorama de gestión y recuperación de la información. En este nuevo y no menos complejo mundo informativo conviven, y convivirán aún por mucho tiempo, el procesamiento técnico del libro y el procesamiento tecnológico de los objetos de información digital, dando lugar a toda suerte de servicios de información: unos tradicionales y tangibles, otros digitales y teñidos de la noción de virtualidad y otros híbridos que combinarán, en su planteamiento procedimental, la técnicas documentales y tareas de carácter informático, predicando la idea de interoperabilidad.

A pesar de lo mucho que se ha hablado de estos temas en la literatura profesional y que la tinta electrónica y la universalidad de la Web han contribuido a difundir las nuevas tareas y retos profesionales en torno a la información digital, tenemos ante nosotros un panorama aún difícil de analizar. El fenómeno de la Web, amén de otras interpretaciones, suele asociarse a la infraestructura tecnológica que la hace posible, de igual forma, el procesamiento, análisis y organización de la información electrónica debe de mutar, en su explicación y ejecución, a las tecnologías que facilitarán tales procesos, a saber principalmente, el marcado de documentos, los esquemas de metadatos y al conjunto de protocolos habilitadores entre ellos.

La catalogación convencional combina tres elementos: documentos (en soporte tangible), procesos (descripción bibliográfica y análisis de contenido, principalmente) y productos fruto de esa actividad (los registros bibliográficos y la conjunción de todos ellos en forma de catálogo). Mientras, la organización de recursos electrónicos basada en metadatos implica: documentos entendidos como objetos de información, la asignación, automática o no, de atributos de esa información (metadatos) y registros de metadatos, legibles por máquina, codificados bien con el propio objeto que describen, bien a través de bases de datos (Fig. 1).



**Fig. 1.** Catalogación tradicional vs. organización de recursos electrónicos: Comparación de elementos constitutivos [elaboración propia].

Nuestra disciplina evoluciona progresivamente desde la práctica acumulada en las bibliotecas tradicionales hacia una dimensión más innovadora basada en la arquitectura de la información, la representación estructurada del conocimiento y en la visión, aún profética, de la Web semántica, donde los estándares, algunos antiguos y otros nuevos y en constante evolución (como los esquemas de metadatos), cumplirán su cometido ancilar al servicio de la organización documental. Este desarrollo se traduce a nivel técnico en una evolución en las normas y procedimientos documentales que implica un cambio, y éste a su vez un esfuerzo y adaptación. Contra la idea de Mark Twain (*estoy completamente a favor del progreso, lo que no me gusta es el cambio*), que utilizamos con intención provocadora para iniciar esta reflexión, los gestores de la información tienen que aprender a adaptarse al cambio como una de sus destrezas profesionales, quizás hoy la más importante. El ámbito de la descripción, representación y organización de la información está sufriendo profundas mutaciones en el cómo de sus procedimientos y tendremos que hacerle frente con una actitud innovadora pero sobre todo con una introspección hacia nuestras competencias, no sólo técnicas sino también tecnológicas. Por todo ello, en este documento se pretende dar una visión integral y sistemática sobre el estado de la cuestión en la catalogación y la organización de los recursos de información digital, prestando especial atención a las normas, prácticas y principales iniciativas, e introduciendo elementos de reflexión y perspectiva para su aplicación en nuestro país.

## **2. La Escuela del control bibliográfico y la Escuela estructuralista: convivencia de dos tendencias en la organización de recursos electrónicos**

Para defender esta idea de dividir las tendencias en la descripción y organización de la información digital en escuelas, nos basamos en Matthew Dovey, quien utiliza el término “escuelas” para explicar las distintas concepciones de los metadatos. Este autor, en su elocuente artículo “Cosas sobre cosas: diferentes significados de metadatos” distingue tres interpretaciones de ese concepto que denomina escuelas (de la catalogación, estructuralista y de la estructura de datos, respectivamente) [Dovey, 1999]. Sin embargo no vamos a tratar aquí estrictamente el concepto de metadatos, si bien es cierto que su aprehensión influirá, de forma determinante, en la segmentación que hacemos de las prácticas profesionales en torno a la organización y recuperación de recursos electrónicos. Utilizamos “escuela” pues, con un triple significado:

- a) Por un lado, en su sentido habitual como doctrina, principios, normas y prácticas específicas que adopta un grupo particular de profesionales y/o autores ante un problema, concretamente, en el ámbito que nos ocupa, el problema de la catalogación o del proceso técnico de los objetos de información digital.
- b) También en el sentido de que, cada una de estas escuelas o corrientes, de algún modo, alecciona o da ejemplo y experiencia ante la comunidad bibliotecaria, ganando adeptos en el uso e interpretación de los estándares implicados en cada caso.
- c) Y por último, hablamos de escuelas, en cierta forma, por la preocupación general que tenemos de etiquetar o nominar determinadas conductas profesionales de tal manera que nos permita agrupar y sistematizar la realidad, aún versátil e interpretable, de los procesos en torno a las bibliotecas digitales u otros sistemas de información electrónica.

A partir de esta diada de escuelas que hemos llamado del control bibliográfico y estructuralista, podemos revisar el estado de la cuestión en materia de descripción y organización de recursos digitales, por una parte, desde los formatos catalográficos tradicionales —ISBD, Reglas de Catalogación y MARC, fundamentalmente— y, por otra, desde la aparición de diferentes modelos, esquemas o estándares de metainformación. De esta forma también podemos analizar la confluencia e interrelación de ambas tendencias o escuelas en pos de la interoperabilidad que, desde un punto de vista técnico, se materializa en la existencia de esquemas íntimamente relacionados con los elementos de metadatos como el propio estándar MARC21, en los ensayos de mapeo y conversión de elementos descriptivos entre formatos, o en la aparición de nuevos protocolos (como OAI, *Open Archives Initiative*) y en la adaptación de los ya existentes vinculados, hasta ahora, estrictamente al mundo bibliográfico (como Z39.50).

### **2.1 La escuela del control bibliográfico: la catalogación en la Era digital**

La primera de las escuelas o tendencias que queremos reflejar aquí, la hemos denominado *del control bibliográfico* por evocación del antiguo CBU (*Control Bibliográfico Universal*) promovido a finales de los 70 del siglo pasado por la IFLA y la UNESCO a fin de unificar los catálogos, definiendo las normas para la creación de asientos bibliográficos, y de hacerlos compatibles para el intercambio automatizado de publicaciones.

Dentro de esta escuela incluimos todas aquellas iniciativas que tratan de ratificar, a partir de los estándares tradicionales de catalogación o de la creación de otros nuevos como la DCMI (*Dublin Core Metadata Initiative*), la idea confinada al

control bibliográfico. En una sociedad globalizada, donde el carácter universal de la Web pretende democratizar el acceso a la información, es más necesario que nunca resucitar la idea de internacionalización e intercambio normalizado de publicaciones. quizás, cambiando la B por una W: **Control Web Universal (CWU)**<sup>3</sup>.

Resulta difícil hablar de la escuela del control bibliográfico, tal y como la concebimos aquí, sin hacer una reflexión de carácter histórico en torno a sus orígenes. A mediados de los 90 el contexto Web era puramente HTML (*Hypertext Markup Language*) —ni siquiera había llegado a ser recomendación su versión 3.2<sup>4</sup>—, Netscape no había recibido todavía el obstinado ataque de Microsoft en el reino de los agentes de usuario, aún no había aparecido Google <<http://www.google.com>> y la confianza en los metadatos (metaetiquetas) que empezaba a digerir Altavista<sup>5</sup> <<http://www.altavista.com>> era escasa debido al *spamming*<sup>6</sup>. Las distintas comunidades informativas de la Red intentaban organizar su información electrónica a través de estándares de metadatos (hasta entonces, basados generalmente en el complejo metalenguaje SGML, *Standard Generalized Markup Language* (ISO 8879-1986); por ejemplo ICPSR (*Inter-University Consortium for Political and Social Research*) para la descripción de conjuntos de datos de ciencias sociales, FGDC (*Federal Geographic Data Committee*) en el ámbito de la información geoespacial, o TEI (*Text Encoding Initiative*) para las publicaciones de carácter literario y/o humanístico. La comunidad bibliotecaria, más que ninguna otra, comenzaba también a tomar sus decisiones sobre la ya urgente organización de la Web.

En un principio, se pensó que seguir con las mismas herramientas catalográficas que tantas satisfacciones habían dado a los bibliotecarios de la tinta y el papel en el control e intercambio bibliográfico en la era pre-Internet. Así lo expresaban diversos autores en el Coloquio sobre Catalogación celebrado en 1996 en el seno del proyecto InterCat (*Internet Cataloging*) de OCLC<sup>7</sup>, defendiendo, entre otras cosas que el formato MARC y las Reglas de Catalogación Angloamericanas (AACR) eran herramientas totalmente adecuadas para la descripción de recursos de Internet. En esa misma línea continuista, y quizás siguiendo el viejo principio de “más vale lo malo conocido que lo bueno por conocer”, comenzaron a adaptarse los estándares tradicionales al nuevo medio de la información electrónica distribuida en Internet. Así, se introduce en el USMARC el campo para la localización electrónica y la información sobre el acceso (856), en el que se fundarían los primeros conatos de catalogación de la Red; en 1994 comienza al la revisión de las ISBD para ficheros de ordenador —ISBD (CF)— que finalmente se publicarían en 1997 bajo la denominación de ISBD (ER) —*Electronic Resources*—; en 1998 se revisan las Reglas de Catalogación Angloamericanas diez años después de su segunda edición (AACR2)<sup>8</sup>; también en 1998 la IFLA publica la recomendación *Functional Requirements for Bibliographic Records* (FRBR) a fin de reestructurar las bases de datos catalográficas, de tal forma que reflejen una estructura conceptual de los recursos de información<sup>9</sup>; a tenor de todos estos cambios se publican guías de catalogación y manuales adaptados al nuevo medio electrónico<sup>10</sup>.

Paralelamente a todos estos cambios en las normas y herramientas, convencionales hasta entonces, surge, también en el seno de OCLC, el esquema de metadatos Dublin Core, en un principio, como reacción bibliotecaria lógica a la inconsistencia de la recuperación de carácter general en la Web a través de servicios de búsqueda. Un conjunto sencillo de, en un principio, trece elementos, y desde finales de 1996, quince, que permitiría describir cualquier tipo de objeto de información de la Red, coadyuvando de esta forma a mejorar la localización de recursos a través de motores de búsqueda. Desde su primera reunión en marzo de 1995 en Dublin (Ohio, EE.UU), de la que recibe el nombre, el Dublin Core ha evolucionado hasta convertir su conjunto de elementos (DCMES, *Dublin Core Metadata Element Set*) en un estándar primero nacional en el contexto americano (Z39.85-2001) y, desde marzo de este año

en un estándar internacional (ISO-15836-2003). Asimismo, en sus años de desarrollo el DC ha ido “cualificándose”, esto es, introduciendo subelementos que sirven para refinar o cualificar las descripciones de los DLOs (*Document Like Objects*<sup>11</sup>) La Iniciativa de Metadatos Dublin Core (DCMI, como se denomina desde el 2001) es más que un mero formato de metainformación y se ha convertido en poco tiempo en el mayor estándar de metadatos a nivel internacional, convirtiéndose en el objetivo de las transformaciones y la compatibilidad de otros esquemas, tanto de la comunidad bibliotecaria, como de otras comunidades de información de la Web, facilitando además el desarrollo de conjuntos de metadatos específicos<sup>12</sup>. Si bien es cierto que el DC surge en el contexto o bajo la escuela general del control bibliográfico, y los primeros experimentos de aplicación e implementación de este formato se llevaron a cabo en bibliotecas (principalmente universitarias) o en instituciones de investigación para crear sistemas de recuperación de información académica basados en recursos Web, hoy en día esta escuela ve al Dublin Core, como el resto de las comunidades, como un estándar que hay que compatibilizar.

A diferencia de lo que ocurre con los registros MARC, el uso del DC no requiere un conocimiento especial, incluso se pueden crear automáticamente, a través de diversas aplicaciones, registros basados en la semántica de la DCMI. Como ya hemos discutido en otras ocasiones [Méndez, 2002, p. 84-87] el formato MARC no es el más adecuado para la descripción de objetos de información digital. Sin embargo, al igual que el desarrollo del DC convivió con la adaptación del MARC, ambos formatos conviven en la actualidad en pos de ese control bibliográfico de las publicaciones científicas en Internet y comparten, en muchos proyectos bibliotecarios, la utilización de esquemas de organización del conocimiento —también tradicionales, aunque concebidos igualmente como metadatos, en este caso orientados al contenido y a la recuperación por materias— como la Lista de Encabezamientos de Materia de la Biblioteca del Congreso (LCSH), o la Clasificación Decimal de Dewey (DDC), etc.

A pesar de todo esto, la organización de objetos de información digital, desde el punto de vista de lo que llamamos la escuela del control bibliográfico, no es sólo un problema de estándares y de descripción, sino también una cuestión de selección. De igual forma que las bibliotecas tradicionales no catalogan todos los impresos que se publican, sino sólo aquellos que se adecuan a los objetivos de la institución, en el mundo digital, no todos los objetos de información deben catalogarse. Esta escuela de la “catalogación de Internet” debe desarrollar un papel proactivo no sólo en los estándares de descripción, sino en pautas de evaluación y selección de recursos Web, que pueden incluirse en sus catálogos y colecciones.

Todos los aspectos que hemos apuntado sobre los estándares, formatos de catalogación y metadatos así como la forma de trabajar para formar colecciones de recursos electrónicos desde la perspectiva del control bibliográfico, confluyen en el proyecto que OCLC comenzó a finales de 1998 *Cooperative Online Resource Catalog* (CORC). En CORC añaden además, dos aspectos fundamentales, desde nuestro punto de vista, en la organización de objetos digitales: por un lado, la cooperación, como actividad necesaria —simplemente explicada desde el volumen de la Red— para abordar este tipo de trabajo técnico, y por otro, la tendencia general a la estructuración de la información o, lo que es lo mismo, el influjo de la escuela estructuralista y del XML en el control bibliográfico ante el desarrollo de la Web semántica. Basándose en el modelo que sigue en WorldCat para el material impreso, CORC tenía como objetivo contar con diferentes estructuras de registro, principalmente MARC y DC, en la misma base de datos. Para ello utiliza el *Resource Description Framework* (RDF), compatible con XML, HTML y MARC, importando y exportando los registros alternativamente, con semántica MARC o Dublin Core. CORC ha sido absorbido por el servicio integral de OCLC denominado *Connexion*<sup>13</sup> que permite la creación y edición de registros bibliográficos y de autoridad facilitando la

catalogación integrada en línea. El sistema de CORC ha servido además de referente para el desarrollo de una nueva generación de OPACs o WebPACs<sup>14</sup> o para nuevos sistemas de acceso al documento electrónico, donde tratan de convivir y de compatibilizarse registros en distintos modelos de metadatos. En este sentido destacan los esfuerzos que las distintas empresas de software están haciendo por incluir y compatibilizar registros de información relativos a recursos electrónico, como por ejemplo Hyperion<sup>15</sup> de la empresa Sirsi, para organizar el *archivo de medios digitales* y hacerlo interoperar con el OPAC bibliográfico en Unicorn; o DigiTool que ha desarrollado Ex-Libris en relación a su producto Aleph 500 para crear colecciones digitales como parte de una estrategia digital global; o MetaLib, también de Ex-Libris, que basándose en la tecnología SFX y en distintos formatos de metainformación, permite la organización, disseminación, recuperación y acceso de información científica en un entorno heterogéneo de catálogos bibliográficos de bibliotecas y bases de datos Web<sup>16</sup>.

## **2.2 La escuela estructuralista: XML, el estándar, también para bibliotecas.**

De igual forma que la escuela del control bibliográfico ha hecho evolucionar sus estándares o ha creado otros nuevos para facilitar su propia concepción del control universal de la información Web, al mismo tiempo y de manera similar, como fruto de un proceso evolutivo, se desarrolla otra escuela que hemos denominado estructuralista, cuyo momento más significativo podríamos datarlo en 1996<sup>17</sup>. En esta escuela podemos considerar todas aquellas tendencias en la organización de la información relacionadas con el marcado estructural de los documentos electrónicos en XML y con la construcción de la **Web semántica**. Cuando Dovey analiza las corrientes principales a la hora de interpretar y aprehender el concepto de metadatos, se refiere al estructuralismo como la *invención* del marcado textual como *la habilidad de embellecer un fichero de texto ASCII con comentarios adicionales y señales en lo que se refiere a sus estructuras subyacentes e interpretativas* [Dovey, 1999, p. 10]. Sin embargo en la división que hacemos aquí (más allá de lo que son estrictamente los metadatos), bajo la escuela estructuralista en la organización de la información electrónica podemos recoger todas las orientaciones que utilizan el metalenguaje XML (*eXtensible Markup Language*) como un lenguaje universal para definir estructuras de datos, y cada vez más, como estándar universal para editar, gestionar y almacenar información en el contexto electrónico.

En la Web semántica, el planteamiento de universalidad en la recuperación de recursos electrónicos es distinto al que reflejábamos en la escuela del control bibliográfico. Sus principales promotores, Tim Berners-Lee y James Hendler, persiguen la creación de una Web de datos con significado, de tal forma que un programa de ordenador pueda aprender tanto acerca de lo que quieren decir los datos, como acerca de la información necesaria para procesarlos. Sin embargo, el objetivo es el mismo, un procesamiento global e interoperable de la información Web que, en este caso, descansa en el procesamiento tecnológico del marcado semántico de XML y en estructuras de metadatos.

XML es un subconjunto de SGML que tienen su misma potencialidad pero, más flexible y menos complejo, permite estructurar todo tipo de datos a través del marcado o etiquetas definidas previamente en una DTD (Definición del Tipo de Documento) o en un esquema. HTML, todavía el lenguaje por excelencia de la WWW, no tiene la versatilidad de estructurar los contenidos bibliográficos que presenta XML. Como HTML, XML también usa etiquetas tipo <title> o atributos como lang=en, sin embargo es el propio lenguaje el que especifica el valor de estos atributos y etiquetas, no estando limitado, como HTML, a un conjunto finito de etiquetas con valores predefinidos. Esta es una de las características más importantes de XML en cuanto a la concepción de la organización de la información, que ha hecho posible que cada

comunidad informativa adapte la manera en que quiere estructurar sus datos en la Web, y que supone una verdadera revolución al plantear una descripción particular, en términos de metadatos, de los datos que conforman el contenido de un documento. A pesar de esa posibilidad de que cada tipo de información pueda describir su propio lenguaje de metadatos, sus propios esquemas o DTDs para codificar con sintaxis XML los datos sobre los datos, el propio Consorcio Web ha desarrollado un Marco para la Descripción de Recursos denominado RDF<sup>18</sup> que constituye un modelo genérico de representación de las propiedades de los elementos de descripción de metadatos, del valor de sus atributos, y de las relaciones que se establecen entre ellos. RDF es el lenguaje por excelencia sobre el que descansa el proyecto de la Web semántica y de que las máquinas puedan entenderse entre sí en la expectativa de una Web global en XML.

No es nuestro objetivo explicar aquí como funcionan XML o RDF, sino más bien destacar su vocación de *lingua franca* para la representación y organización de la información cuya aplicación en todo tipo de servicios y sistemas de información digital, incluyendo las bibliotecas digitales, está dando lugar a una nueva forma de concebir nuestra profesión. En el contexto de las tecnologías XML, un documentalista puede convertirse perfectamente en algo mucho más amplio que un simple catalogador, en un creador de metadatos, en un gestor de perfiles informativos de un portal Web o en un sindicador de contenidos.

Por supuesto, la idea de usar un formato universal para representar e intercambiar la información no es nueva y podemos reconocerla ya en la norma ISO 2709 para el intercambio universal de registros bibliográficos. XML hereda de SGML esa idea de universalidad, así como la de separar los datos de su representación gráfica, circunstancias que lo convierten en un estándar clave en el ámbito de las bibliotecas digitales y de las bibliotecas en general. Son muchas las contribuciones que podríamos destacar en este sentido, que van desde las reflexiones de Sperberg-McQueen<sup>19</sup> [1998], al libro de reciente aparición de Roy Tennant [2002] o al breve manual de aplicación de XML en bibliotecas que ha incluido Eric Lease Morgan en su sitio Web *Musings on Information and Librarianship* [Morgan, 2003], pasando por autores de diversos contextos geográficos y/o procedencias [Kim y Choi, 2000]. Hoy nadie duda de la potencia de XML para la representación de la información y por ende su validez en el ámbito documental, y así se ratifica día a día en la lista de distribución XML4Lib<sup>20</sup> sobre el uso de este metalenguaje en bibliotecas. De todas formas, en esta escuela estructuralista podemos situar aquellas tendencias en la organización y catalogación de objetos de información digital que, por distintas circunstancias, creen que los formatos MARC no consiguen responder a la necesidades del contexto informativo de la Red, ya que el MARC no es un lenguaje que pueda aplicar cualquiera y que al mismo tiempo puedan interpretar los navegadores, circunstancias ambas que sí se dan en XML/RDF. El panorama pues, en la escuela estructuralista, podemos definirlo teniendo en cuenta una serie de afirmaciones básicas que nos sirven para clarificar el estado actual (2003) de esta tendencia:

- XML proporciona un formato básico para documentos estructurados sin dependencia de una semántica particular.
- Los esquemas XML y RDF son tecnologías ya maduras en la actualidad (2003) y permiten la utilización de un marcado específico para el almacenamiento, consulta, búsqueda y recuperación tanto de cualquier objeto de información, como de cualquier descripción sobre ese objeto.
- Se han desarrollado distintos estándares de metadatos de propósito particular (por ejemplo, EAD, *Encoded Archival Description* para la documentación del patrimonio archivístico) y el Dublin Core, como conjunto de elementos de propósito general aplicable a cualquier objeto Web. Todos estos esquemas semánticos de



metadatos han encontrado en la sintaxis XML el mayor bastión para organizar y representar la información en las bibliotecas digitales.

La generalización de las tecnologías XML, así como la potencial o pretendida universalidad de la Web, han hecho que la escuela del control bibliográfico y la escuela estructuralista confluyan en pos de la interoperabilidad. Así han surgido algunos esquemas de codificación para el mundo bibliográfico en XML, como por ejemplo BiblioML. BiblioML es un proyecto promovido por Martin Sévigni con el apoyo del Ministerio de Cultura y Comunicación de Francia, y está formado por una DTD basada en el formato bibliográfico internacional normalizado UNIMARC. Con la misma intención, Desde junio de 2002, la Biblioteca del Congreso de los Estados Unidos, a través de su Oficina de Normalización MARC y desarrollo de la Red ha establecido distintos estándares que implican la aproximación de XML al mundo MARC, o aún mejor, la necesidad de vinculación de MARC a la sintaxis XML, a saber:

- Un esquema XML para codificar registros MARC21<sup>21</sup> que permitirá la codificación de metadatos en sentido estricto con semántica MARC, bien a partir de la conversión desde otros formatos de metadatos, bien en su creación original directamente en MARC.
- Un esquema para conjuntos de elementos bibliográficos que pueden utilizarse con distinta finalidad, pero sobre todo en el contexto de las aplicaciones bibliotecarias, denominado MODS (*Metadata Object Description Schema*), que incluye algunos de los campos MARC pero con etiquetas alfabéticas en lugar de las numéricas del MARC (p. ej. `titleInfo`, `typeOfResource`, etc.). Si tuviésemos que definir el tipo de estándar que es MODS diríamos que está, tanto en su complejidad como en su nivel de detalle descriptivo, entre la DCMI y el formato MARC.
- Un esquema específico de metadatos para la concepción de bibliotecas digitales, denominado METS (*Metadata Encoding and Transmission Standard*), y que tiene la finalidad de codificar todo tipo de metadatos (estructurales, descriptivos y administrativos). A pesar de que METS tiene el refrendo y la clave de su desarrollo en la Library of Congress, surge a finales de 2001 en el seno de la iniciativa americana *Digital Library Federation* para la descripción e intercambio de repositorios de metadatos en la segunda parte del Proyecto MOA (*Making Of America*).

MARXML, MODS y METS, implican el reconocimiento en el ámbito bibliotecario de la necesidad de datos estructurados sobre los datos, para mantener una colección de objetos digitales, así como la confluencia entre la escuela del control bibliográfico y la estructuralista en el ámbito de la gestión de la información electrónica. Concretamente METS reconoce la necesidad de metadatos técnicos y administrativos en la organización de bibliotecas digitales, que faciliten el acceso a, por ejemplo, la información técnica sobre los procesos de digitalización, y no sólo a la información descriptiva del documento que podría conseguirse con un simple registro MARC o MARXML.

Esta confluencia demuestra que, en realidad, hablar de escuelas no es más que la pretensión, como decíamos antes, de sistematizar tendencias o de agrupar actitudes ante la organización/catalogación de la Web. La escuela del control bibliográfico se puede ver como una interpretación, una aportación o simplemente una visión de la Web semántica desde el mundo de la catalogación y de la biblioteconomía que trata de evolucionar sus estándares hacia una interoperabilidad que va más allá del mero concepto de biblioteca. Mientras, la escuela estructuralista se puede entender como una visión de la organización del conocimiento y de la catalogación desde la Web semántica, que trata de convertir, en su caso, desde la flexibilidad del XML, la Web en un verdadero sistema de información<sup>22</sup>. El enclave de las dos tendencias son, sin duda alguna, los metadatos. En la primera de ellas, entendidos

como códigos (o esquemas, en una terminología más Web) de catalogación, y en la segunda, entendidos como esquemas semánticos subsidiarios del marcado de documentos para describir y recuperar la información.

### 3. Dimensión disciplinar de la organización de la información electrónica:

Al margen de la división que hemos hecho en “escuelas” en el apartado anterior, otro aspecto destacable en la organización de objetos de información Web es la dimensión disciplinar que adquieren los nuevos procesos de carácter tecnológico realizados sobre el documento electrónico. Es decir, cada comunidad informativa, cada grupo disciplinar o profesional, cada tipo de información Web, e incluso podríamos decir, cada materia o dominio informativo, ha desarrollado sus propios estándares de codificación para estructurar, organizar, difundir y recuperar su información particular. El desarrollo temático o especializado de metadatos va desde la definición de DTDs específicas en XML —en la mayoría de los casos (por ejemplo EAD, TEI) originalmente en SGML— a veces construidas o adaptadas *ad hoc* para un proyecto concreto, a la definición de esquemas y de su cobertura y uso, e incluso, al estudio de los valores posibles y vocabularios específicos que pueden tener estos atributos (metadatos orientados al contenido, como ontologías, tesauros temáticos, o por ejemplo, los valores del elemento `type` en la DCMI<sup>23</sup>),

Este cariz disciplinar y/o especializado que tienen los procesos de descripción y organización de objetos de información digital, es la consecuencia lógica o se debe, al menos, a las siguientes circunstancias que envuelven el mundo informativo de la Red:

- a) La heterogeneidad de recursos de información, y las distintas formas de representarlos en la WWW. Así encontramos en la Web: información textual (libros y revistas electrónicos), imágenes (subrogados digitales de objetos de arte, mapas, planos, fotografías comerciales), objetos multimedia (películas y video digital), etc.
- b) Los procesos diferentes de creación de objetos digitales, bien a través de la digitalización de documentos tradicionales, bien a través de la creación directa en formato digital.
- c) La “verticalización” de la información. Esto es, la tendencia general a la especialización temática de los servicios y sistemas de información electrónica en la Red, dando lugar tanto a portales comerciales temáticos como a colecciones virtuales de recursos científicos especializados en una materia, que a su vez desarrollarán un formato de metainformación específico y adecuado para dicha materia. A este tipo pertenecen los sistemas de información especializada en el ámbito de la educación superior o las denominadas *subject gateways* o *information gateways*<sup>24</sup>.
- d) La necesidad que tienen las distintas instituciones públicas y/o privadas no sólo de dar información, sino también de generar servicios. Dentro de estos servicios podemos destacar el gobierno electrónico y la teleeducación, en el ámbito público, o toda la suerte de sistemas de comercio electrónico en el contexto privado. En ambos casos necesitan describir y recuperar los objetos digitales que conforman su actividad o tramitación o bien, su modelo de negocio.
- e) La propia versatilidad de la sintaxis XML/RDF que, a través de los espacios de nombre (*namespaces*) permite adaptarse a las necesidades de cada comunidad de interés y codificar su información consistentemente y de manera independiente a cualquier semántica específica.

En muchos casos, los esquemas de metadatos cobran importancia, al margen de sus procesos de formalización como estándares internacionales, justamente porque

pueden ser diseñados para servir a comunidades específicas de usuarios, y porque una vez que existe un consenso sobre su estructura, son relativamente sencillos de implementar, lo cual permite experimentar con los esquemas y refinarlos. Este es por ejemplo el caso de la DCMI que, a pesar de ser un estándar de propósito general, esto es, dirigido a organizar y describir cualquier objeto de información Web, debe su éxito, entre otras cosas, a:

- Su capacidad de adaptación a necesidades de descripción particular información. Así, ha servido de base para otros esquemas de metadatos como el AGLS (*Australian Government Locator Service*) o el modelo del EdNA (*Education Network Australia*).
- La posibilidad de cualificación y matización de sus elementos —bien a través de cualificadores o de esquemas [*schemes*] de contenido— permite utilizar este formato con mayor precisión descriptiva, adecuándolo a las necesidades de una comunidad específica o un tipo de objeto de información particular<sup>25</sup>.
- Su independencia de la sintaxis. Así, se puede codificar el conjunto de los quince elementos Dublin Core (DCMES) en HTML, XML, XHTML, o RDF, o a través de una definición de campos en una base de datos, ajustándose así, tanto para su uso en colecciones virtuales, como en bibliotecas digitales que editan electrónicamente sus colecciones.
- Y sobre todo, a la posibilidad de crear perfiles de aplicación específicos que, basados en esquemas [*schemas*] RDF, permiten definir un conjunto de entidades que sirven para declarar qué elementos, de qué espacios de nombre (*namespaces*) subyacen a un esquema local utilizado para organizar la información de una temática concreta o de un proyecto en particular [Heery y Patel, 2000].

Se pueden distinguir prácticamente tantas aprehensiones de la organización de la información digital, tantos modelos “catalográficos” o mejor, tantos esquemas de metadatos particulares, como tipos y temáticas posibles. Sin embargo, podríamos distinguir tres comunidades y sectores informativos destacados por su especial trabajo en la descripción y recuperación de recursos<sup>26</sup>:

- 1) La información gubernamental. En este dominio destacan todas las iniciativas, normalmente con un cariz nacional o geográfico, de descripción de objetos de información digital fruto de los sistemas de gobierno electrónico o de la administración electrónica que están desarrollando distintos países. Así, destacan modelos y esquemas de metadatos como: GILS en el ámbito americano, e-GMS (*e-Government Metadata Standard*) para la organización de recursos digitales en el ámbito del gobierno electrónico en el Reino Unido, o AGLS y NZGLS, basados ambos en el DC y utilizados para la información gubernamental en Australia y Nueva Zelanda, respectivamente.
- 2) El sector de la educación. Los metadatos en la educación implican la descripción de lo que se denomina en este ámbito “objetos de aprendizaje”, es decir, cursos, materiales de aprendizaje, etc. y existen igualmente diversos esquemas especializados a tal efecto, como: IMS (*Instructional Management Systems*) GEM (*Gateway to Educational Materials*), LOM (*Learning Object Metadata*), o el ya citado EdNA.
- 3) La información geoespacial donde la descripción de recursos geográficos tiene una larga tradición normativa que se inicia en el *Content Standard for Digital Geospatial Metadata* (CSDGM), más conocido como FGDC, y culmina en la ISO 19115, estándares de metadatos que sustentan en XML/XSL distintos

Sistemas de Información Geográfica, e incluso Infraestructuras de Información Espacial.

A estas tres percepciones particulares de la organización sectorial de la información electrónica a través de modelos descriptivos de metadatos, añadiríamos además la comunidad bibliotecaria, donde podrían incluirse todas las tendencias y normas que hemos denominado “escuela del control bibliográfico” (MARCXML, METS, MODS, etc.) y otros trabajos relacionados con la organización de recursos electrónicos en el ámbito general de las bibliotecas digitales, como por ejemplo el esquema de aplicación específico del Dublin Core en bibliotecas<sup>27</sup>.

#### **4. Normalización y tendencias hacia la interoperabilidad.**

Partiendo de todas las diversas dimensiones en la descripción y organización de contenidos disciplinares para la Web, e independientemente de que la semántica funcional de cada uno de los modelos particulares tienda a representarse y codificarse en XML, los distintos metadatos necesitan interoperar con otros esquemas para facilitar la búsqueda integral en la Web, no sólo a nivel sintáctico, sino también semántico. La interoperabilidad es el gran reto en la gestión de la información de la Web, tanto desde la perspectiva de la Web Semántica, como desde la perspectiva más bibliotecaria que hemos denominado Control Universal de la Web (CWU); y no es otra cosa que la capacidad que tiene un esquema de metadatos de trabajar con otro modelo de metainformación distinto, obteniendo resultados globales.

Para explicar muchos de los procesos que acontecen en el mundo de la información digital, necesitamos establecer símiles e incluso metáforas constantes que lo relacionen con el mundo mejor aprehendido de la documentación impresa o “tradicional”. De manera elemental se podría decir que, si los metadatos son, en cierto sentido, la evolución de la idea en la que se funda la catalogación, la interoperabilidad es, en cierto sentido también, el resultado evolutivo de la idea del intercambio de registros en el mundo UNIMARC, ISO2709 y Z39.50, e igualmente se basará en la normalización, en la creación de protocolos o en la adaptación de los existentes y en la búsqueda de correspondencia semántica (mapeo) entre los elementos de distintos modelos de metainformación. Esto que parece tan sencillo encierra una gran complejidad y ocupa el centro de atención de todas las comunidades de metadatos, y por supuesto también a bibliotecarios y documentalistas. Ahora el reto es que la pantalla del ordenador que presenta los catálogos en línea sea también el dispositivo para visualizar los objetos digitales y conectar con toda la información del mundo electrónico.

En el contexto informativo distribuido, heterogéneo y cambiante de la Web, no sólo es suficiente contar con estándares robustos —ya sean formales, como las ISO 15836 ó 19115 a nivel semántico, o *de facto* como XML a nivel sintáctico—, sino también con la posibilidad de que éstos interoperen, es decir, que permitan compartir información descrita según distintos esquemas de codificación y recuperar información global procedente de repositorios digitales diversos. En una Sociedad de la Información Global que trata de cimentarse sobre el principio del “acceso universal” a la información, la creación de sistemas de datos interoperables no se puede prorrogar, al menos a los niveles o contextos que señalamos a continuación, como primer paso fundamental hacia la idea de Web semántica; donde la Web alcanzará todo su potencial si se convierte en un entorno donde los datos estén definidos y enlazados de tal forma que se puedan usar para una recuperación de información más eficaz, integrando recursos entre diferentes aplicaciones.

El primer contexto donde debe conseguirse la interoperabilidad es entre sistemas de información del mismo tipo, que utilizan un mismo esquema de metadatos pero que tienen sistemas de recuperación de información locales. Por ejemplo, que

dos colecciones distintas de datos geoespaciales (*clearinghouses*) puedan intercambiar sus objetos de información o que permitan realizar una búsqueda integral en ellas, bajo un mismo interfaz y de manera distribuida. En este caso, la interoperabilidad estará relacionada, sobre todo, con el nivel de estandarización y con el hecho de que ambos sistemas utilicen sintaxis, estructuras y vocabularios compatibles para describir sus datos. Este nivel de interoperabilidad no es ajeno al mundo bibliotecario, y es el que fundamenta la búsqueda en catálogos MARC distribuidos basada en el protocolo Z39.50 (ISO 23950).

El siguiente paso en la interoperabilidad, más complejo, es el que se establece entre sistemas completamente distintos y también en una arquitectura Web totalmente distribuida. Al mismo tiempo que aumentan exponencialmente los contenidos de la Web, aumenta la cantidad y variedad de metadatos para describir y gestionar esos contenidos heterogéneos. Si además, a esto le añadimos, tal y como explicamos en el apartado anterior, que cada dominio científico o informativo utiliza esquemas de metadatos (vocabularios o conjuntos de elementos) diferentes, tanto en lo relativo a su cobertura semántica como a su complejidad, el panorama para interoperabilidad se complica, y aún más, si tenemos en cuenta la implementación local que hacen de ellos, incluido el idioma de los valores que aplican para cada elemento. Pensemos por ejemplo, en un usuario que busca información sobre Leonardo da Vinci. En la actual percepción de la recuperación de información global, el usuario puede dirigirse a un buscador tipo Google, y resignarse a discriminar, de entre los 402.000 recursos o enlaces que aparecen, la información de calidad (en función de diversos criterios como la fiabilidad de la fuente, o la cobertura intelectual de su consulta en ese momento) de la información irrelevante. Sin embargo, la idea de la Web semántica y de la interoperabilidad en que se funda es que el usuario sea capaz de encontrar, con la misma búsqueda sobre "leonardo da vinci" distintos objetos de información como: una revista científica en que hablen del pintor italiano (metadatos: PRISM), una fotografía de da Vinci (metadatos: DIG35); la Gioconda (metadatos: CIMI o VRA), un registro bibliográfico de un libro que hable sobre Leonardo (MARC), un vídeo sobre la pintura del Cinquecento (metadatos: MPEG7), un curso virtual (metadatos: LOM o IMS), etc. Lo primero que podríamos pensar es que sería estupendo que todos los sistemas de información utilizaran MARC para describir los recursos informativos de la Web; sin embargo ese planteamiento es elemental, ya que la robustez de cada uno de los sistemas de información digital radica justamente en el esquema específico de metadatos que utilizan. La posibilidad de realizar una búsqueda integral como la del ejemplo, será posible gracias a los avances en torno a la interoperabilidad de los cuales vamos a señalar las cuatro iniciativas o tendencias más importantes, basadas, en todo caso, en la normalización.

- 1) La evolución de los protocolos de búsqueda existentes. El caso más importante en este sentido es Z39.50 que en los últimos años ha avanzado hacia "una nueva generación" del protocolo (ZING, *Z39.50 International Next Generation*). En esta nueva generación de Z, se recogen diversas iniciativas de los implementadores de Z39.50 para que el contenido semántico e intelectual de este protocolo sea más accesible y más atractivo para los proveedores, vendedores y usuarios. Con objetivos semejantes, es decir, aumentar la capacidad de Z39.50 aplicada a los recursos electrónicos de la Red, se han desarrollado múltiples perfiles de aplicación de Z39.50<sup>28</sup> y se sigue investigando en la interoperabilidad entre Bib-1 y los distintos esquemas de metadatos. Asimismo existen proyectos semejantes recientes como Zinterop (*Z39.50 Interoperability Testbed Study*) que dirigido por el profesor Moen pretende realizar un banco de pruebas de interoperabilidad de Z39.50, inicialmente realizado sobre 419.657 registros MARC21 procedentes de WorldCat [Moen, 2003]

- 2) La creación de nuevos protocolos de interoperabilidad como OAI-PMH (*Open Archives Initiative – Protocol for Metadata Harvesting*). OAI surgió originalmente en el contexto de la Convención de Santa Fe, celebrada en 1999 y en los trabajos de Van de Sompel, que definió una interfaz que permitiera que un servidor de publicaciones electrónicas diera a conocer los metadatos de los artículos que tenía almacenados, de tal forma que dichos metadatos se pudieran recoger bien por servicios de búsqueda unificados, bien por otros repositorios que quisieran identificar las publicaciones para copiarlas [Lynch, 2001]. Posteriormente a Santa Fe, y en poco tiempo, OAI se ha convertido en el protocolo por excelencia entre distintos esquemas de metainformación que describen DLOs en distintas colecciones de distintas bibliotecas o repositorios de información digital. Esencialmente es un protocolo basado en la arquitectura cliente-servidor y construido sobre HTTP diseñado para distribuir, recopilar y compartir metadatos. Expresado en XML convierte los metadatos a la semántica DC sin cualificar, de tal forma que se pueden compartir metadatos almacenados en distintos sistemas. Actualmente existen dos tipos de participantes en OAI: los proveedores de datos que administran el sistema que soporta OAI-PMH como una forma de divulgar los metadatos; y los proveedores de servicio que utilizan los metadatos recopilados por el protocolo OAI como base para construir servicios de valor añadido. Son muchas las instituciones que están implementando e investigando el potencial de OAI. OCLC, por ejemplo, ha desarrollado un software de fuente abierta para soportar OAI<sup>29</sup> y están trabajando para construir nuevos modelos de servicio.
- 3) Otra iniciativa o grupo de iniciativas de interoperabilidad la constituye lo que podemos llamar “la armonización de esquemas de metadatos”. En este sentido cabe destacar la resolución CORES del 18 de noviembre de 2002, donde las distintas instituciones participantes en el proyecto CORES —financiado por el programa IST (*Information Society Technologies*) del VI Programa Marco de la UE— llegaron a un consenso para asignar URIs (*Uniform Resource Identifiers*) a cada uno de los elementos de metadatos de distintos esquemas, (concretamente, GILS, ONIX, MARC21, CERIF, DOI, IEEE/LOM y DCMI, a los que se unieron también, en junio de 2003 los gestores de UNIMARC) [Baker y Dekkers, 2003]. La asignación de URIs a los elementos de metadatos, junto con el desarrollo de una política eficaz que garantice la persistencia y estabilidad de tales URIs, parte del reconocimiento de que la unidad constitutiva de cualquier modelo de metainformación es el elemento, y en tanto que éstos elementos son unidades de significado pueden compararse y mapearse a otros estándares, constituyendo así también, un paso hacia la interoperabilidad.
- 4) La última tendencia que queremos destacar en materia de interoperabilidad entre distintos formatos son los registros de esquemas de metadatos. Los registros de metadatos son básicamente, *bases de datos de esquemas que pueden trazar la línea histórica para compartir diccionarios de datos y el proceso de registro impulsado por la comunidad ISO/IEC 11179* [Heery y Wagner, 2002], aunque realmente el desarrollo de registros de esquemas de metadatos está íntimamente asociado a la creación y a la noción de Web semántica, y concretamente se impulsaron desde el ámbito de la normalización para la WWW y desde la comunidad de gestores del conocimiento. Debido a la naturaleza distribuida de la Web, se pueden facilitar de muchas formas índices de términos, es decir de registros de metadatos, principalmente local o distribuida; simplificando en cualquier caso, la indización de esquemas relevantes de un dominio científico o disciplina particular y facilitando la navegación entre diversos esquemas. Si bien esta práctica de construcción de registros de metadatos no es equiparable a ninguna de las iniciativas que hemos señalado antes, facilita la normalización hacia la interoperabilidad, no sólo en un ámbito multidisciplinar, sino también

multilingüe<sup>30</sup> a través del conocimiento profundo de los distintos modelos de metadatos para la organización de la información digital.

## **5. La organización de la información electrónica en España y Cataluña: iniciativas, tendencias y proyectos.**

Es ahora el momento de revisar cómo afectan todas las tendencias en la organización de recursos electrónicos en nuestro país. En España, como en el resto de los países, los bibliotecarios y documentalistas se han visto desbordados por la cantidad, variedad y valor de la información producida y difundida a través de Internet, y desde hace varios años ya, ensayan propuestas de descripción y catalogación de recursos electrónicos, que van desde la inclusión de registros MARC en sus catálogos, a la creación de sistemas de recuperación locales constituidos por documentos electrónicos. Para analizar la situación en España distinguiremos, sin ánimo de ser exhaustivos, sino más bien de reflejar las tendencias y realidades, dos ámbitos principales: por un lado, las reacciones bibliotecarias en relación al procesamiento de la información Web en las bibliotecas españolas en general, y por otro la situación de la implantación de metadatos y otros estándares relacionados en nuestro país, reflejando asimismo aquellos proyectos o sistemas que han surgido específicamente a tenor de la Red y que entrarían dentro de lo que intuitivamente se entiende por bibliotecas digitales o sistemas de información digital.

### **5.1. La organización de recursos electrónicos en las bibliotecas españolas**

La mayoría de las bibliotecas en España no describen como parte de su colección los recursos electrónicos, o al menos no lo hacen integrándolos en los catálogos MARC. La mayoría de sus páginas web cuenta con un enlace tipo: *recursos electrónicos, enlaces de Internet, otros recursos*, simplemente *Internet* e incluso *recursos de interés bibliotecario* dando un énfasis especial a aquellas páginas de la red de interés para la profesión bibliotecaria (obras de referencia, normas, editoriales y librerías, etc.). En cualquier caso, este tipo de “servicio” incluido en los sitios Web de nuestras bibliotecas, nunca es más que una lista, directorio o índice clasificado con enlaces a páginas seleccionadas por su potencial interés informativo para los usuarios de cada biblioteca, en el mejor de los casos, revisado o mantenido periódicamente.

En cuanto a la Biblioteca Nacional de España, a diferencia de otros países donde se marcan las pautas de catalogación y de gestión de metadatos de información digital<sup>31</sup>, la única iniciativa en relación a la constitución y organización de la biblioteca digital es la reciente creación del *Sistema de Recuperación de Recursos Electrónicos*, que ha desarrollado en colaboración con la Fundación Telefónica dentro del proyecto *Atril Virtual*. Este sistema de confusa denominación<sup>32</sup>, permite el acceso a los propios fondos digitalizados de la Biblioteca Nacional, facilitando la visualización (a través de la descarga de un *plug-in* de Java) del documento completo, bien desde la consulta del catálogo Ariadna o bien a través de las exposiciones virtuales de la propia Biblioteca Nacional. Este sistema, no implica ninguna innovación con respecto a la descripción de documentos digitalizados, simplemente se cataloga, de manera convencional en formato IBERMARC el documento impreso y se añade un acceso a su versión digital, no realizándose ningún proceso de marcado ni de descripción de metadatos sobre este fondo digitalizado. En el caso de la Biblioteca Nacional de Cataluña, tampoco existe ningún sistema específico de descripción y recuperación de recursos electrónicos, como en la mayoría de las bibliotecas españolas, bajo el enlace “recursos Internet” clasifica en seis categorías de enlaces, a listas poco exhaustiva de páginas de interés bibliotecario (general, organismos bibliotecarios, bases de datos bibliotecarias, bibliotecas, bibliotecas digitales, otros recursos).

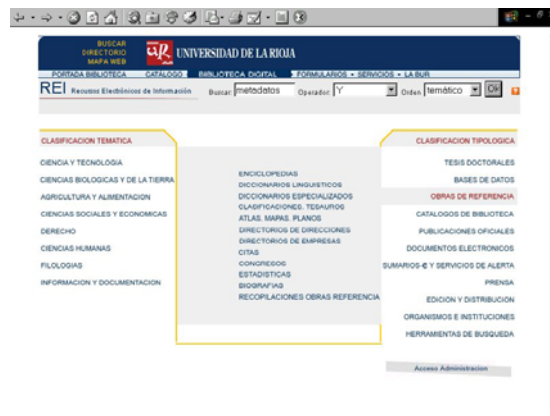
Una reflexión aparte, en cuanto a las tendencias en la catalogación y organización de recursos Web, merecen las bibliotecas universitarias o de investigación. Algunas bibliotecas académicas incluyen en sus catálogos ediciones digitales de actas de congresos y revistas científicas, principalmente, disponibles en Internet (Fig. 2). La inclusión de este tipo de publicaciones en los WebPAC de las bibliotecas universitarias españolas se debe a dos circunstancias principales:

- Por un lado, al convencimiento creciente en el ámbito bibliotecario de que los catálogos deben ampliar su papel tradicional y no ser sólo el inventario de los recursos (bibliográficos, etc.) que la biblioteca posee, sino más bien el depósito descriptivo de los recursos a los que la biblioteca tiene acceso (incluyendo por supuesto las suscripciones a revistas electrónicas, y también a recursos de acceso libre en la Red).
- Por otra parte, se debe a que la mayoría de las bibliotecas universitarias españolas han implantado Sistemas Integrados de Gestión Bibliotecaria de “segunda generación” (Unicorn, Aleph500, Virtua, Innopac/Millennium, etc.) fundados sobre MARC21 y a que en la mayoría de los casos (Hyperion de Sirsi, MetaLib y DigitoolLibrary de Ex-Libris, p. ej.) incluyen extensiones de software para gestionar tanto las colecciones digitalizadas localmente, como las colecciones virtuales de recursos académicos en línea.

Además de esta circunstancia, también es habitual en las bibliotecas universitarias, la inclusión de un directorio o índice de recursos basado en las materias o disciplinas que se imparten en las distintas universidades, llegando en algunos casos a crear un sistema de recuperación específico con una estructura de información local, para estos recursos. Este es el caso, por ejemplo, del servicio REI (Recursos Electrónicos de Información) que mantiene la biblioteca de la Universidad de La Rioja (Fig. 3), donde sobre una base de datos en FileMaker, se clasifican por materias los recursos de carácter científico, convirtiendo el tradicional “enlaces de interés”, que repiten casi todas las bibliotecas universitarias, en un servicio de búsqueda por palabras clave.



**Fig. 2.** Registro bibliográfico relativo a una publicación electrónica. Biblioteca de la Universidad Carlos III de Madrid  
<http://ieeexplore.ieee.org/servlet/opac?punumber=2192>



**Fig. 3.** Base de datos REI (Recursos Electrónicos de Información) Universidad de La Rioja  
<http://pellinor.unirioja.es:8880/biblio/recursos?sub=1>

En el caso de las bibliotecas universitarias catalanas, hay algunos aspectos destacables en lo que a descripción y recuperación de documentación electrónica se refiere. En cuanto a la organización y búsqueda de objetos de información digitales de acceso remoto, destaca, como no podría ser de otra forma dado su carácter “virtual”,



la biblioteca de la Universidad Oberta de Catalunya (UOC) donde además del catálogo convencional, permite la consulta de su “fondo digital”, entendiendo por tal bases de datos, revistas a texto completo, enciclopedias, estadísticas y otros recursos de Internet y recursos propios digitales: por ejemplo, conferencias de Manuel Castells (en texto o audio), artículos de su personal docente e investigador, libros digitalizados, etc. Las búsquedas se pueden acotar a los campos de autor, título, materia y resumen, lo que implica que, aunque no exista un modelo de metadatos formal, los recursos digitales de acceso remoto se seleccionan y se describen en función de estos campos, pudiendo distinguirse: proceso de catalogación y el catálogo tradicional, y proceso de descripción de recursos electrónicos, y catálogo de fondo digital.



Fig. 4. Búsqueda en la Colección digital de la biblioteca de la UOC

<http://xina.uoc.es/esp/index.html?cdigital/index>



Fig. 5. Búsqueda en TDX (Tesis Doctorals en Xarxa).

[http://www.tdx.cesca.es/index\\_tdx\\_cs.html](http://www.tdx.cesca.es/index_tdx_cs.html)

En cuanto a la descripción de información local digitalizada, es importante destacar el reciente servicio TDX (*Tesis Doctorals en Xarxa*), coordinado por el *Consorci de Biblioteques Universitàries de Catalunya* (CBUC) (Fig. 5). El CBUC, ha mostrado desde hace mucho tiempo la necesidad de plantear políticas de tratamiento de la información electrónica (así lo demuestran por ejemplo estudios como el de Núñez [2000] y la existencia desde 1999 la Biblioteca Digital de Cataluña), pero en este caso TDX implica el tratamiento de documentos digitalizados, concretamente las tesis doctorales leídas en las universidades de Cataluña principalmente. Uno de los aspectos más importantes de este servicio es que pertenece al proyecto mundial NDLTD (*Networked Digital Library of Theses and Dissertations*) donde, además de la búsqueda en el catálogo albergada por VTLs, se está ensayando la búsqueda “federada” (es decir, la búsqueda unificada de recursos distribuidos) a través del protocolo OAI<sup>33</sup>.

## 5.2. Implantación de metadatos en España para la organización de objetos de información digital

Al hablar de la implantación o uso de esquemas de metadatos en nuestro país, obviamente queremos referirnos a los sistemas y servicios de información electrónica, a los proyectos, etc. que utilizan dicha tecnología para la descripción, organización y recuperación de los objetos de información digital, sean éstos digitalizados *ad hoc*, o bien recursos de la Web que se describen de manera individual formando colecciones virtuales de recuperación y/o acceso cualificado. Sin embargo, nos parece inevitable, dado que en muchos ámbitos profesionales aún se presume la validez de estas estructuras de información en las cabeceras de los documentos HTML para mejorar la

búsqueda global en la WWW, hacer una breve referencia al uso de metaetiquetas en las páginas web de bibliotecas de nuestro país.

```
<meta name="title" content="Universitat Autònoma de Barcelona. UAB. L'Autònoma.">
<meta name="DC.Title" content="Universitat Autònoma de Barcelona. UAB. L'Autònoma.">
<meta http-equiv="title" content="Universitat Autònoma de Barcelona. UAB. L'Autònoma.">
<meta name="keywords" content="UAB, Autònoma, Autonomous, Barcelona, Bellaterra, universitat, estudios, investigacion, diplomatura, master, ingenieria, doctorado, licenciatura, Universitat, Estudis, recerca, llicenciatura, enginyeria, doctorat, University, Studies, investigate, doctorate, undergraduate, postgraduate, graduate">
<meta name="Description" content="Universitat Autònoma de Barcelona, Universidad Autònoma de Barcelona, Autonomous University of Barcelona">
<meta name="abstract" content="Universitat Autònoma de Barcelona, Universidad Autònoma de Barcelona, Autonomous University of Barcelona">
<meta name="author" content="Universitat Autònoma de Barcelona">
<meta name="revisit-after" content="10 Days">
<meta name="DC.Creator" content="Universitat Autònoma de Barcelona">
<LINK REV=made href="mailto:webmaster@uab.es">
<meta http-equiv="keywords" content="UAB, Autònoma, Autonomous, Barcelona, Bellaterra, universitat, estudios, investigacion, diplomatura, master, ingenieria, doctorado, licenciatura, Universitat, Estudis, recerca, llicenciatura, enginyeria, doctorat, University, Studies, investigate, doctorate, undergraduate, postgraduate, graduate">
<meta http-equiv="description" content="Universitat Autònoma de Barcelona, Universidad Autònoma de Barcelona, Autonomous University of Barcelona">
<meta http-equiv="DC.Description" content="Universitat Autònoma de Barcelona, Universidad Autònoma de Barcelona, Autonomous University of Barcelona">
<meta name="VW96.objecttype" content="Homepage">
<meta http-equiv="Content-Type" content="text/html; ISO-8859-1">
<meta name="distribution" content="all">
<meta name="resource-type" content="Homepage">
<meta http-equiv="Pragma" content="no-cache">
<link rel="shortcut icon" href="favicon.ico">
<meta name="robots" content="all">
<title>Universitat Autònoma de Barcelona</title>
```

**Fig 6.** Código fuente-metaetiquetas de la cabecera de la página principal de la Universidad Autónoma de Barcelona <<http://www.uab.es>>

En cuanto al uso de metaetiquetas en las páginas web de bibliotecas, podemos decir en líneas generales que<sup>34</sup>:

- Se usan los valores `keywords` y `description`, dirigidos supuestamente a aumentar la visibilidad de estas bibliotecas en la Web a través de buscadores de carácter global. Es habitual que las palabras claves aparezcan en castellano y en inglés y en el caso de las bibliotecas de Cataluña, también en catalán (Fig. 6).
- No existe una política coherente de aplicación de metaetiquetas en las páginas web de las bibliotecas españolas. En el caso de que se usen, cada una adopta arbitrariamente la forma de consignarlas.
- Normalmente se mezclan códigos o esquemas, por ejemplo, valores del atributo `name` del Dublin Core, con otros valores generales de HTML, e incluso se utiliza el atributo `http-equiv` con carácter igualmente descriptivo (Fig. 6).
- Aunque la finalidad de estas etiquetas es mejorar el *ranking* de relevancia en motores de búsqueda generales, su uso no garantiza la visibilidad, debido sobre todo, como ya hemos apuntado, a la evolución de los sistemas de recuperación de tipo “*all-the-web*”, que no tienen en cuenta estos valores de metainformación.

Al fin y al cabo, el uso actual de metaetiquetas podría evaluarse más en términos de arquitectura de la información, o como parte del diseño Web de las páginas de los servicios de información, que del aspecto descriptivo de la información en sí misma. No obstante la veleidat detectada en el uso de tales metadatos básicos, revela, en cierto sentido, la poca utilización-implantación y superficial aprehensión de la tecnología/tendencia de los metadatos en nuestro sector profesional. No obstante, en cuanto al tratamiento, indización y búsqueda a través de metaetiquetas, hay que mencionar los trabajos del grupo iris-index de RedIRIS que ha investigado, desde 1997 la puesta en marcha de una estructura de intercambio de información indexada, primero a través de metadatos SOIF en Harvest y posteriormente Dublin Core simple, desarrollando incluso herramientas de indización específicas (Webber) o plantillas para introducir metaetiquetas en las páginas (MetaWebber)<sup>35</sup>.

Por otra parte, aunque existe un interés creciente por los esquemas formales de metadatos en España, aún es muy tímida su implantación. No obstante podemos comentar algunas experiencias tanto en el ámbito nacional como catalán. Se pueden señalar al menos dos grupos de iniciativas: aquellos proyectos de ámbito supranacional que en los que puntualmente participa España, y aquellas iniciativas netamente nacionales donde se implementan elementos descriptivos de metadatos. Dentro del primer grupo, destacamos el proyecto COVAX que, financiado por el Programa IST del VI Programa Marco, cuenta con la participación de tres instituciones españolas (la Residencia de Estudiantes del CSIC, la UOC, y la biblioteca Menéndez Pelayo) y la empresa Software AG. COVAX tiene como objetivo: la difusión del patrimonio cultural europeo facilitando el acceso a los ciudadanos europeos a las fuentes del patrimonio intelectual, cultural y científico conservado en archivos, bibliotecas y museos; la explotación a través de Internet de las infraestructuras culturales existentes; la aplicación de normas en el campo de la estructura y la recuperación de información y la interoperabilidad entre sistemas a través de distintos modelos de metadatos aplicados al patrimonio (EAD, TEI, etc.) codificados en XML. En esta misma línea, esto es, proyectos internacionales con participación española, se encuentra también, por ejemplo LINGU@net, un centro de recursos multilingüe que ofrece acceso gratuito a un directorio de más de 1.500 recursos relacionados con la enseñanza y el aprendizaje de unos 45 idiomas, donde la catalogación y organización de recursos se hace en función del esquema del Dublin Core [Bellido y Pastor, 2002]. Finalmente también nos parece importante revelar la participación financiera del Ministerio de Educación y Cultura de España en el reciente proyecto E-LIS (*Eprints in Library and Information Science*) para crear un archivo multilingüe y multinacional de la información científica digital en nuestra disciplina, basado en el funcionamiento del protocolo OAI para la búsqueda distribuida basada en metadatos (Fig. 7).

Fig.7. Formulario de búsqueda avanzado de E-LIS <http://eprints.rclis.org/perl/advsearch>

Fig. 8. Catálogo de metadatos del IDEC <http://www.geoportal->

En cuanto a iniciativas específicas de metadatos a nivel nacional resaltan, en el ámbito de la digitalización del patrimonio archivístico, los proyectos de la Subdirección General de Archivos Estatales: AER (Archivos Españoles en Red) y Censo Guía de Archivos de España e Iberoamérica, en cuyo seno se ha desarrollado su propio esquema descriptivo EAG (*Encoded Archival Guide*) una DTD de XML específica basada en el estándar EAD para la descripción de objetos de información de archivo. Y también es sumamente importante el trabajo desarrollado en metadatos geoespaciales por el consorcio TEIDE (Tecnologías para la Infraestructuras de Datos Espaciales<sup>36</sup>) en la utilización de FGDC e ISO 19115 para el tratamiento y recuperación de información pública cartográfica de una forma estandarizada; uno de los últimos avances de TEIDE en el ámbito de los metadatos a este nivel es el estudio de la interoperabilidad o realización de pasarelas para compatibilizar metadatos tan específicos, con el modelo de propósito general Dublin Core.

En Cataluña, también existen iniciativas de descripción de información electrónica en ámbitos específicos, como cabría esperar por el nivel de normalización, también en el terreno de la información geoespacial, así existe el IDEC (*Infraestructura de Dades Espacials de Catalunya*), donde se mantiene un catálogo de metadatos (Fig. 8), e incluso han desarrollado un software específico para la edición y exportación de metadatos según un subconjunto de elementos de la ISO 19115 para la descripción de la información geográfica.

En nuestro sector profesional, existen varios grupos de investigación en este campo, de entre los que podemos destacar el trabajo del grupo del Servicio de Documentación Multimedia, Multidoc, de la Universidad Complutense, donde, entre otras cosas, se está desarrollando una aplicación web para la creación de colecciones documentales multimedia en formato electrónico, denominada *WebCollector*. *WebCollector* cuenta en la actualidad con un núcleo, un módulo de administración y control de acceso, dos módulos de búsqueda y un módulo que implementa en protocolo OIA-PMH para la distribución de metadatos Dublin Core, lo que permite el acceso a los recursos y a su descripción desde cualquier parte de la Red. También es importante en el ámbito profesional catalán, el trabajo del grupo de la Facultad de Biblioteconomía y Documentación dirigido por Assumpció Estivill donde, a través del proyecto *Recursos web i metadades*, se planteó el objetivo de comprobar las dificultades de aplicación de metadatos DC a las páginas del sitio web de dicha facultad [Estivill, 2001]. En la actualidad continúan trabajando en esta línea, actualmente aplicando dicho formato de metadatos cualificado a revistas científicas como *BiD*.

A pesar de que todavía no contamos con una masa crítica equiparable a otros países en la aplicación de metadatos, este tema suscita un gran interés en el sector de los profesionales de la información en nuestro país y son muchas las iniciativas, seminarios, cursos y reuniones profesionales que dan cabida a este tema. Merece mención especial, el *Seminario sobre la aplicación en España del modelo de metadatos Dublin Core*, celebrado en Madrid el 5 de junio de 2003 donde se presentó el mirror en español de la DCMI <<http://es.dublincore.org>>, que surgió originalmente en el grupo sobre Normalización para la Recuperación de la Información en la Web de SEDIC, pero que finalmente se ha realizado en el seno del Instituto Agustín Millares de Documentación y Gestión de la Información en la Universidad Carlos III de Madrid y que pretende divulgar en español los principales documentos, etc. relacionados con la iniciativa más importante de metadatos DCMI ISO 15836. El DCMES (*Dublin Core Metadata Element Set*) había sido traducido al catalán ya en mayo de 2001, *Vid.* <http://www.gencat.es/bc/dces.htm>. Y asimismo es importante el documento realizado en el seno del Consorci de Biblioteques Universitàries de Catalunya: *Recomanacions*

del CBUC per a la catalogació en Dublin Core de recursos web en bases de dades multidisciplinàries mantingudes per una institució del CBUC, en: <http://www.cbuc.es/5digital/DCdoc2-WebCBUC.doc>

Fruto de dicho seminario celebrado este año (2003), se ha creado también la lista DCMI-ES, Foro español de la *Dublin Core Metadata Initiative* <<http://www.rediris.es/list/info/dcmi-es.es.html>> que cuenta con más de 370 suscriptores<sup>37</sup>. Quizás tengamos que esperar, a nuestro juicio poco ya, para que las bibliotecas realicen sistemas de información de calidad basados en modelos de metadatos y en sintaxis XML. Mientras tanto, conviven en nuestro país muchas percepciones relativas al procesamiento de los recursos electrónicos; en ellas están involucrados, sin ningún lugar a dudas, distintos sectores profesionales —además del bibliotecario o catalográfico— que, a nuestro parecer, tendrán que hacer un esfuerzo de confluencia de intereses en la recuperación de información electrónica, de tal forma que las distintas iniciativas de metadatos se realimenten y compatibilicen de forma interdisciplinar.

## 6. Conclusiones y reflexión.

A lo largo de este artículo hemos revisado distintos esfuerzos e iniciativas, procedentes de diversas comunidades de desarrollo del Web (bibliotecarios, documentalistas, informáticos, e incluso otros ámbitos profesionales) que se interesan por la descripción de recursos electrónicos y pretenden mejorar la organización de la información en Internet y proporcionar métodos más efectivos para su recuperación. Todas estas propuestas que van de la catalogación “tradicional” en el mundo bibliotecario, a la tendencia de los metadatos y el desarrollo de un Marco para la Descripción de Recursos (RDF) que permita construir una arquitectura donde puedan coexistir diferentes estructuras de datos, procedentes de distintos colectivos informativos. El binomio: comunidad de información particular-esquema de metadatos específico, que genera una dimensión eminentemente disciplinar en la gestión de la información en entornos digitales debe de confluir a favor de la interoperabilidad; así pues el avance de OAI, Z39.50, la creación de registros de esquemas de metadatos y de otras iniciativas semejantes, son decisivos para compatibilizar y recuperar información de la gran cantidad de depósitos de información digital existentes en la Red.

Con todo lo que hemos apuntado aquí, a pesar de la variedad de tendencias en la organización de los objetos de información digital, a pesar de la rémora de los viejos estándares bibliográficos o de la complejidad informática de las recomendaciones del W3C, está claro que la Web nos está llevando a una nueva forma de organizar y poner a disposición de los usuarios la información, en definitiva, a una nueva forma de organizar, representar, recuperar y preservar el conocimiento. No hace mucho tiempo que una colega australiana me dio su tarjeta de visita donde debajo de su nombre ponía claramente “*metadata librarian*”, cargo que contemplé conmovida, desde la credibilidad que tengo, desde hace mucho tiempo ya, en esta tecnología, y desde la legitimación que dicha mención suponía de los cambios que está sufriendo nuestro trabajo, en lo relación a las nuevas destrezas que debe desarrollar el profesional de la información ante la WWW<sup>38</sup> para pasar de *procesos técnicos* a un *procesamiento tecnológico* de la información. Dublin Core y otros esquemas de metadatos, unidos a la potencia instrumental de XML, nos proporcionan herramientas nuevas —no sólo a los bibliotecarios, sino a una comunidad mucho más amplia— para la descripción y el acceso de la Web, esto es, nos dirigen a lo que antes denominaba CWU (*Control Web Universal*) y nos abren una puerta de participación muy amplia en la construcción de la Web Semántica.

Si queremos rentabilizar nuestra experiencia en la gestión del conocimiento, los profesionales de la información tendremos que vincularnos de forma activa al desarrollo de estándares de metadatos y a todas las innovaciones que nos permitan, ulteriormente, mejorar la recuperación de información en el contexto real de nuestro trabajo que hoy en día es, sin ninguna duda, la Web.

En cuanto a nuestro país, si bien es cierto que todo apunta a que esta situación cambie en los próximos años, sobre todo en el ámbito de las bibliotecas científicas y universitarias, podemos decir que hasta hoy (fin de 2003) no se ha afrontado, o al menos no se ha hecho de forma notable y mucho menos unificada, la organización de recursos electrónicos. La implantación de esquemas y/o modelos de metadatos es igualmente mínima y carece de planteamientos uniformes disciplinares a nivel nacional. A pesar de los pequeños avances e iniciativas desarrolladas en los últimos años en torno a la formación, divulgación y aplicación de estándares internacionales para la organización, recuperación y preservación de la documentación digital, queda aún mucho camino por recorrer. Quizás haya que esperar algunos años todavía para que debajo de nuestros nombres en la tarjeta de visita ponga "bibliotecario/a de metadatos", sin embargo es muy necesario que la formación y la evolución de nuestros profesionales vaya en esta dirección. Y se deben divulgar en España todas las iniciativas que trabajen con metadatos, así como incentivar la investigación multidisciplinar (bibliotecarios-documentalistas/otras disciplinas y comunidades informativas) en la organización y recuperación de objetos de información digital; involucrando en estas investigaciones, imitando lo que ocurre en otros países, tanto a la Administración Pública como a las más importantes bibliotecas a nivel nacional.

## Bibliografía<sup>39</sup>

- AGUDO PLANS, Sílvia. Avaluació de les metadades com a eina d'indexació i recuperació al web de la Biblioteca de la Universitat de Barcelona [documento HTML] *BiD: Biblioteconomia i Documentació*, diciembre 2002, nº 9. Disponible en: <http://www.ub.es/biblio/bid/09argudo.htm>
- BAKER, Thomas, Makx Dekkers. Identifying Metadata Elements with URIs: The CORES Resolution [documento HTML]. *D-Lib Magazine*, July/August 2003, vol. 9, nº 7/8. Disponible en: <http://www.dlib.org/dlib/july03/baker/07baker.html>
- BANERJEE, Kyle. Describing Remote Electronic Documents in the Online Catalog: Current Issue. *Cataloging and Classification Quarterly*, 1997, vol. 25, nº 1, p. 5-20.
- BELLIDO, Luis, Encarna Pastor. Lingu@net Europa: Un catálogo de recursos multilingüe con garantía de calidad [documento HTML]. *Cultura & Política @ Ciberespacio. 1er Congreso Online del Observatorio para la Cibersociedad*, septiembre 2002. Disponible en: <http://cibersociedad.rediris.es/congreso/comms/g02bellido.htm>
- BERNERS-LEE, Tim. Metadata Architecture [documento HTML]. *Design Issues*. W3C, 6 de enero de 1997, rev. 21 de septiembre de 2000. Disponible en: <http://www.w3.org/DesignIssues/Metadata.html>
- BERNERS-LEE, Tim. *Tejiendo la red: el inventor del World Wide Web nos descubre su origen*. Madrid: Siglo veintiuno, 2000.
- BRISSON, Roger. The World Discovers Cataloging: A Conceptual Introduction to Digital Libraries, Metadata and the Implications for Library Administrations. *Journal of Internet Cataloging*, 1999, vol. 1, nº 4, p. 3-30.
- CAPLAN, Priscilla. *Metadata Fundamentals for All Librarians*. Chicago: ALA, 2003
- CATALOGING *Internet Resources: A Manual and Practical Guide* [documento HTML], Nancy B. Olson, ed. 2<sup>nd</sup> ed. OCLC, 1997, versión electrónica del 23 de mayo de 2003. Disponible en: <http://www.oclc.org/support/documentation/worldcat/cataloging/internetguide>
- CATALOGING *the Web: Metadata, AACR, and MARC 21*. Wayne Jones, Judith R. Ahronheim, and Josephine Crawford, eds. Lanham: Scarecrow Press, 2002.

- DC-2001. *Proceedings of the International Conference on Dublin Core and Metadata Applications* [documento PDF]. Keizo Oyama and Hironobu Gotoda, eds. Tokyo: National Institute of Informatics, 2001. Disponible en: <http://www.nii.ac.jp/dc2001/proceedings/product/book.pdf>
- DEKKERS, Makx and Stuart L. Weibel. Dublin Core Metadata Initiative Progress Report and Workplan for 2002 [documento HTML]. *D-Lib Magazine*, February 2002, vol. 8, nº 2, Disponible en: [www.dlib.org/dlib/february02/weibel/02weibel.html](http://www.dlib.org/dlib/february02/weibel/02weibel.html).
- DÍAZ ORTUÑO, Pedro Manuel. Problemática y tendencias en la arquitectura de metadatos Web. *Anales de Documentación*, 2003, nº 6, p. 35-58. Disponible en línea en: <http://www.um.es/fccd/anales/ad06/ad0603.pdf>
- DOVEY, Matthew J. "Stuff" about "Stuff"- the Differing Meanings of "Metadata". *Vine (Theme issue: Metadata. Part 1)*, 1999, nº 116, p. 6-13.
- ESTIVILL RIUS, Assumpció. Organizació dels recursos Internet. *Ítem*, 1996, nº 18, p. 42-74.
- ESTIVILL RIUS, Assumpció. El processament dels recursos electrònics [documento HTML]. *BiD: Biblioteconomia i Documentació*, juny 1998, nº 1. Disponible en: <http://www.ub.es/biblio/bid/01estiv1.htm>
- ESTIVILL RIUS, Assumpció, et al. Recursos web i metadades: informe del projecte [documento HTML]. *BiD: Biblioteconomia i Documentació*, desembre 2001, nº 7. Disponible en: <http://www.ub.es/biblio/bid/07estiv1.htm>
- GORMAN, Michael. Metadata or Cataloging? A False Choice. *Journal of Internet Cataloging*, 1999, vol. 2, nº 1, p. 5-22.
- GUENTHER, Rebecca and Sally McCallum. New Metadata Standards for Digital Resources: MODS and METS. *Bulletin of the American Society for Information Science and Technology*, 2002/2003, vol. 29, nº 2, p. 12-15.
- HANNAH, Stan A. and Michael H. Harris. *Inventing the Future: Information Services for a New Millennium*. Stamford: Ablex Publishing Corporation, 1999.
- HEERY, Rachel, Harry Wagner. A Metadata Registry For The Semantic Web [documento HTML]. *D-Lib Magazine*, May 2002, vol 8, nº 5. Disponible en: <http://www.dlib.org/dlib/may02/wagner/05wagner.html>
- HEERY, Rachel and Manjula Patel. Application Profiles: Mixing and Matching Metadata Schemas [documento HTML]. *Ariadne*, 9-October-2000, issue 25. Disponible en: <http://www.ariadne.ac.uk/issue25/app-profiles>
- HUDGINS, Jean, Grace Agnew, Elizabeth Brown. *Getting Mileage out of Metadata: Applications for the Library*. Chicago: ALA, 1999.
- HUTHWAITE, Ann. AACR2 and Other Metadata Standards: The Way Forward. *Cataloging & Classification Quarterly*, 2003, vol. 36, nº 3/4, p.87-100.
- IFLA. *ISBD (ER) International Standard Bibliographic Description for Electronic Resources: revised from the ISBD (CF) International Standard Bibliographic Description for Computer Files*. München: Saur, 1997. Disponible en línea en: <http://www.ifla.org/VII/s13/pubs/isbd.htm>
- INTRODUCCIÓN a los Metadatos: vías a la información digital*. Murtha Baca, ed. Los Angeles: J. Paul Getty Trust, 1999.
- JUL, Eric. Cataloging Internet Resources: Survey and Prospectus. *Bulletin of the American Society for Information Science*, October/November 1997, vol. 24, nº 1, p. 6-9.
- KELLY, Brian. WebWatch: Conclusions From The WebWatch Project [documento HTML]. *Ariadne*, March 1999, issue 19. Disponible en: <http://www.ariadne.ac.uk/issue19/webwatch>
- KIM, Hyun-Hee and Chang-Seok Choi. XML: how it will be applied to digital library systems. *The Electronic Library*, 2000, vol. 18, nº 3, p. 183-189.

- LYNCH, Clifford A. Metadata Harvesting and the Open Archives Initiative [documento HTML]. *ARL Bimonthly Report*, August 2001, 217. Disponible en: <http://www.arl.org/newsltr/217/mhp.html>
- LYNCH, Clifford A. Searching the Internet: combining the skills of the librarian and the computer scientist may help organize the anarchy of the internet. *Scientific American*, March 1997, nº 276, p. 52-56.
- MÉNDEZ RODRÍGUEZ, Eva M<sup>a</sup>. El documento como objeto de información: Metadatos y esquemas de organización de la información en la Web. En: *Recursos informativos: creación, descripción y evaluación*. Felipe Zapico Alonso (coordinador). Badajoz: Junta de Extremadura, Consejería de Educación, Ciencia y Tecnología, 2003 (Serie Sociedad de la Información, nº 8), p. 59-76.
- MÉNDEZ RODRÍGUEZ, Eva M<sup>a</sup>. *Metadatos y recuperación de información: estándares, problemas y aplicabilidad en bibliotecas digitales*. Gijón: Trea, 2002.
- MILSTEAD, Jessica and Susan Feldman. Metadata: Cataloging by any Other Name. *Online*, January/February 1999, vol. 23, nº 1, p. 24-31.
- MOEN, William E., Penelope Bernardino. Assessing Metadata Utilization: An Analysis of MARC Content Designation Use. En: *DC-2003 Proceedings of the International DCM/ Metadata Conference and Workshop (2003. Seattle)*. [Seattle]: Information School of the University of Washington, etc., 2003, p. 171-180. Resumen en línea en: <http://www.siderean.com/dc2003/Paper58-abstract.pdf>
- MORGAN, Eric Lease. Getting Started with XML: A Manual and Workshop [documento PDF]. 2<sup>nd</sup> ed. *Musings on Information and Librarianship*. South Bend: Infomotions, 21 de abril de 2003, rev. 17 de septiembre de 2003. Disponible en: <http://www.infomotions.com/musings/getting-started/getting-started.pdf>
- MUXACH, Santi i Ana Lopo. Metadades a peu pla. *Item*, 1999, nº 24, p. 99-134.
- NÚÑEZ, Lluïsa. *Metadades: informe per al CBUC* [documento PDF]. Barcelona: Biblioteca Digital de Catalunya, 2000. Disponible en: <http://www.cbuc.es/5digital/meta0007.pdf>
- O'NEILL, Edward T. FRBR: Application of the entity-relationship model to Humphry Clinker. *Library Resources and Technical Services* vol. 46, nº 4, October 2002. Disponible en línea en: [http://www.oclc.org/research/publications/archive/2002/oneill\\_frbr22.pdf](http://www.oclc.org/research/publications/archive/2002/oneill_frbr22.pdf)
- PHILLIPS, Lee Anne. *Using XML: Special Edition*. Bestseller ed. Indianapolis: Que, 2000.
- SCHESCHY, Virginia M. Cataloging Procedures on the Web: the Greatest Thing since MARC. *Cataloging & Classification Quarterly*, 1998, vol. 26, nº 2, p. 11-23.
- SENSO, José A., Antonio de la Rosa Piñero. El concepto de metadato: algo más que la descripción de recursos electrónicos. *Ciencias da Informação*, Maio/Ago, 2003, vol. 3, nº 2, p. 95-106. Disponible en línea en: <http://www.ibict.br/cionline/320203/3220311.pdf>
- SPERBERG-MCQUEEN, C. M. *XML and what it will mean for libraries* [documento HTML]. Chicago: University of Illinois, 30 de junio de 1998. Disponible en: <http://tigger.uic.edu/~cmsmcq/talks/teidlf1.html>
- SULLIVAN, Danny. Death of Meta Tags [documento HTML]. *Search Engine Watch*, 1 de octubre de 2002. Disponible en: <http://searchenginewatch.com/sereport/article.php/2165061>
- TILLET, Barbara B. AACR2 and Metadata: Library Opportunities in the Global Semantic Web. *Cataloging & Classification Quarterly*, 2003, vol. 36, nº 3/4, p.101-119.
- WEBER, Mary Beth. Factors to be Considered in the Selection and Cataloging of Internet Resources. *Library Hi Tech*, 1999, vol. 17, nº 3, p. 298-303.
- XML in Libraries*. Roy Tennant (ed.). New York: Neal-Schuman Publishers, Inc, 2002.

### Sitios Web:

[ADAM] Art, Design, Architecture & Media  
Information Gateway

<http://www.adam.ac.uk>



<b>[AER]</b> Archivos Españoles en Red	<a href="http://aer.mcu.es">http://aer.mcu.es</a>
<b>[ASGF]</b> Australian Subject Gateways Forum	<a href="http://www.nla.gov.au/initiatives/sg/gateways.html">http://www.nla.gov.au/initiatives/sg/gateways.html</a>
<b>[AGLS]</b> Australian Government Locator Service	<a href="http://www.naa.gov.au/recordkeeping/gov_online/agls/summary.html">http://www.naa.gov.au/recordkeeping/gov_online/agls/summary.html</a>
<b>[ATUA]</b> Australian Trade Union Archives	<a href="http://www.atua.org.au/atua.htm">http://www.atua.org.au/atua.htm</a>
<b>[BDC]</b> Biblioteca Digital de Cataluña	<a href="http://www.cbuc.es/5digital/5mcdigital.htm">http://www.cbuc.es/5digital/5mcdigital.htm</a>
<b>[BiblioML]</b>	<a href="http://www.biblioml.org">http://www.biblioml.org</a>
<b>[CBUC]</b> Consorci de Biblioteques Universitàries de Catalunya	<a href="http://www.cbuc.es">http://www.cbuc.es</a>
<b>[CERIF]</b> Common European Research Information Format	<a href="http://www.cordis.lu/cerif">http://www.cordis.lu/cerif</a>
<b>[CONSER]</b> Cooperative Online SERials	<a href="http://www.loc.gov/acq/conser">http://www.loc.gov/acq/conser</a>
<b>[CORC]</b> Cooperative Online Resource Catalogue	<a href="http://corc.oclc.org">http://corc.oclc.org</a> <a href="http://www.oclc.org/connexion">http://www.oclc.org/connexion</a>
<b>[COVAX]</b> Contemporary Culture Virtual Archives in XML	<a href="http://www.covax.org">http://www.covax.org</a>
<b>[DCMI]</b> Dublin Core Metadata Initiative	<a href="http://www.dublincore.org">http://www.dublincore.org</a>
<b>[DCMI-ES]</b> Mirror en español de la DCMI	<a href="http://es.dublincore.org">http://es.dublincore.org</a>
<b>[DCMES]</b> Dublin Core Metadata Element Set (ISO 15836)	<a href="http://www.niso.org/international/SC4/n515.pdf">http://www.niso.org/international/SC4/n515.pdf</a>
<b>[DOI]</b> Digital Object Identifier	<a href="http://www.doi.org">http://www.doi.org</a>
<b>[e-GMS]</b> e-Government Metadata Standard	<a href="http://www.govtalk.gov.uk/schemasstandards/metadata.asp">http://www.govtalk.gov.uk/schemasstandards/metadata.asp</a>
<b>[E-LIS]</b> E-Prints in Library and Information Science	<a href="http://eprints.rclis.org">http://eprints.rclis.org</a>
<b>[EAD]</b> Encoded Archival Description	<a href="http://www.loc.gov/ead">http://www.loc.gov/ead</a>
<b>[EdNA]</b> Educational Network Australia	<a href="http://www.edna.edu.au/metadata">http://www.edna.edu.au/metadata</a>
<b>[FRBR]</b> Functional Requirements for Bibliographic Records	<a href="http://www.ifa.org/VII/s13/frbr/frbr.htm">http://www.ifa.org/VII/s13/frbr/frbr.htm</a>
<b>[FGDC]</b> Federal Geographic Data Committee	<a href="http://www.fgdc.gov">http://www.fgdc.gov</a>
<b>[GEM]</b> Gateway to Educational Materials	<a href="http://www.geminfo.org">http://www.geminfo.org</a>
<b>[GILS]</b> Global Information Locator Service	<a href="http://www.gils.net">http://www.gils.net</a>
<b>[ICPSR]</b> Inter-University Consortium for Political and Social Research	<a href="http://www.icpsr.umich.edu">http://www.icpsr.umich.edu</a>
<b>[JOSTOR]</b> Journal Storage Project	<a href="http://www.jstor.org">http://www.jstor.org</a>
<b>[IEEE/LOM]</b> Learning Object Metadata	<a href="http://ltsc.ieee.org/wg12">http://ltsc.ieee.org/wg12</a>
<b>[MARC]</b> Standards Website	<a href="http://lcweb.loc.gov/marc">http://lcweb.loc.gov/marc</a>
<b>[MARC-ES]</b> MARC en español	<a href="http://www.loc.gov/marc/marcspa.html">http://www.loc.gov/marc/marcspa.html</a>
<b>[MARCXML]</b> MARC XML schema	<a href="http://www.loc.gov/standards/marcxml">http://www.loc.gov/standards/marcxml</a>
<b>[METS]</b> Metadata Encoding and Transmission Standard	<a href="http://www.loc.gov/standards/mets">http://www.loc.gov/standards/mets</a>
<b>[MOA]</b> Making Of America	<a href="http://sunsite.berkeley.edu/MOA2">http://sunsite.berkeley.edu/MOA2</a>
<b>[MODS]</b> Metadata Object Description Schema	<a href="http://www.loc.gov/standards/mods">http://www.loc.gov/standards/mods</a>
<b>[NDLTD]</b> Networked Digital Library of Theses and Dissertations	<a href="http://www.theses.org">http://www.theses.org</a>
<b>[NZGLS]</b> New Zealand Government Locator Service	<a href="http://www.e-government.govt.nz/nzglis">http://www.e-government.govt.nz/nzglis</a>
<b>[OAI]</b> Open Archives Initiative	<a href="http://www.openarchives.org">http://www.openarchives.org</a>
<b>[OMNI]</b> UK's Gateway to High Quality Internet Resources in Health and Medicine	<a href="http://omni.ac.uk">http://omni.ac.uk</a>
<b>[ONIX]</b> Online Information eXchange	<a href="http://www.editeur.org/onix.html">http://www.editeur.org/onix.html</a>
<b>[PADI]</b> Preserving Access to Digital Information	<a href="http://www.nla.gov.au/padi">http://www.nla.gov.au/padi</a>
<b>[RDF]</b> Resource Description Framework	<a href="http://www.w3.org/RDF">http://www.w3.org/RDF</a>
<b>[SW]</b> Semantic Web (W3C)	<a href="http://www.w3.org/2001/sw">http://www.w3.org/2001/sw</a>
<b>[TDX]</b> Tesis Doctorals en Xarxa	<a href="http://www.tdx.cesca.es">http://www.tdx.cesca.es</a>
<b>[TEI]</b> Text Encoding Initiative	<a href="http://www.tei-c.org">http://www.tei-c.org</a>
<b>[W3C]</b> World-Wide Web Consortium	<a href="http://www.w3.org">http://www.w3.org</a>
<b>[WebLaw]</b>	<a href="http://www.weblaw.edu.au">http://www.weblaw.edu.au</a>
<b>[XML]</b> eXtensible Markup Language	<a href="http://www.w3.org/XML">http://www.w3.org/XML</a>

**[Z39.50]** Information Retrieval (Z39.50):  
Application Service Definition and Protocol  
Specification (ISO23950)

<http://lcweb.loc.gov/z3950/agency>

**[ZING]** Z39.50 International Next Generation

<http://www.loc.gov/z3950/agency/zing>

**[Zinterop]** Z39.50 Interoperability Testbed  
Project

<http://www.unt.edu/zinterop>

## NOTAS:

---

<sup>1</sup> Los ciclos de Kondratieff, también conocidos como olas de la innovación tecnológica, y la *S de Foster* son las dos representaciones gráficas ya clásicas que correlacionan la innovación tecnológica con el tiempo. Concretamente, la *S de Foster* determina las etapas de evolución que sufre toda nueva tecnología: emergencia, crecimiento, madurez y saturación y obsolescencia. Richard N. Foster. *Innovation: The attacker's advantage*. New York: Summit Books, 1986.

Stan A. Hannah y Michael H. Harris utilizan el formato MARC como ejemplo de la evolución tecnológica que representa la *S de Foster* y para explicar por qué no es rentable la inversión en una tecnología madura. La opinión de Hannah y Harris es significativa para el contenido de nuestro artículo sobre la evolución de la catalogación a tenor de Internet y de la WWW, ya que estos autores afirman claramente que *no es posible revisar el registro MARC para soportar la búsqueda de hipertexto* [Hannah y Harris, 1999, p. 17-19]. Si bien esta afirmación es demasiado categórica, como veremos, justifica la necesidad de los cambios que se están produciendo en la organización de la información digital y la tendencia de los metadatos que también trataremos aquí.

<sup>2</sup> La concepción del documento digital de la Web como objeto de información (DLO, *Document Like Object*), promovida desde el estándar de metadatos de propósito general más común (el Dublin Core) es ya tradicional en el ámbito de la descripción de recursos electrónicos, siendo el DLO el elemento informativo, cualquiera que sea su formato digital, susceptible de asignarle metadatos descriptivos y/o de contenido. No profundizaremos ahora en el concepto e implicaciones del DLO, para una explicación más exhaustiva *vid.* [Méndez, 2003]. Senso y de la Rosa [2003], también recogen esta idea de los documentos como objetos, partiendo también de las reflexiones de Berners-Lee en el documento *Metadata Architecture* donde, en mi opinión, se sienta las bases, ya en 1997, de lo que ahora conocemos como Web semántica.

<sup>3</sup> El Control Web Universal, es también lo que persigue la *Web Semántica*, pero en su caso, desde los estándares de estructuración y marcado de documentos (XML/RDF) y de la representación del conocimiento a través de ontologías, que destacaremos al hablar de la escuela estructural. *Vid. Infr.*

<sup>4</sup> Recordemos que la versión 3.2 del HTML fue Recomendación del W3C en 1997. *Vid. HTML 3.2 Reference Specification: W3C Recommendation, 14-Jan-1997* [documento HTML]. Dave Raggett. W3C, 14 de enero de 1997. Disponible en: <http://www.w3.org/TR/REC-html32>

<sup>5</sup> Nos referimos a las metaetiquetas del HTML `<meta name="keywords" content="... , ... , ... , ">` y `<meta name="description" content="...">` que se habían llegado a denominar, en torno a 1998, *metadatos de Altavista* [Kelly, 1999]. Sobre el uso del elemento `<meta>` *Vid.* [Méndez, 2002, p. 145-147]. *Cfr.* [Sullivan, 2002]

<sup>6</sup> Técnica que consiste básicamente en repetir, dentro de una etiqueta `<META>` una palabra clave con fines promocionales, de tal forma que el enlace del documento en cuestión suba al inicio de una lista de resultados de un buscador de Internet. [*Cfr.* Phillips, 2000, p. 490]. El hecho de que las etiquetas fuera *un imán* para esta práctica de *spam* [Sullivan, 2002] hizo que, a partir de 1997, los motores de búsqueda dejaran de considerarlas paulatinamente. En la actualidad, Inktomi `<http://www.inktomi.com>` es el único motor de búsqueda que tiene en cuenta el elemento donde se consignan los descriptores en la cabecera de un documento HTML.

<sup>7</sup> El *OCLC Internet Cataloging Colloquium*, se celebró en S. Antonio (Texas) a principios de 1996 `<http://digitalarchive.oclc.org/da/ViewObject.jsp?objid=0000003889>`.

<sup>8</sup> Huthwaite [2003, p. 88] afirma, refiriéndose a la segunda edición de las Reglas de Catalogación Anglo-Americanas, que *son estándares de metadatos que continúan siendo tan viables hoy como cualquiera de los otros estándares de metadatos que se utilizan para organizar la información*. A pesar de los cambios introducidos en la última revisión de las AACR2 en relación a los recursos electrónicos y publicados a finales de 2002 (con cambios del 2003 recogidos en una edición aparte, *Vid.*

[http://www.alastore.ala.org/SiteSolution.taf?\\_sn=catalog2&\\_pn=product\\_detail&\\_op=1257](http://www.alastore.ala.org/SiteSolution.taf?_sn=catalog2&_pn=product_detail&_op=1257)) esta afirmación es excesiva para el sentido estricto del concepto de metadatos; máxime si se equipara el estándar de catalogación americano a cualquier otro estándar de metadatos. Las Reglas de Catalogación, en cualquiera de sus interpretaciones nacionales de las ISBD son metadatos en tanto que implican una semántica de los datos para su interpretación y recuperación, sin embargo, éstos no son legibles por máquina, que es la característica que completa la concepción de metadatos para la Web.

<sup>9</sup> De una forma más técnica, FRBR utiliza el modelo de metadatos entidad-relación para los objetos de información, en vez del único concepto de registro que subyace a los estándares de catalogación, e incluye además cuatro modelos de representación: trabajo, expresión, manifestación e ítem. Vid. [O'Neill, 2002].

<sup>10</sup> Este es el caso, por ejemplo del famoso manual de Olson, adoptado y revisado desde su primera publicación en 1997 por el equipo de Eric Jul de OCLC, en la actualidad es la guía de catalogación principal del nuevo proyecto WorldCat de OCLC, junto a *Bibliografic Formats and Standards*, la guía de registros legibles por máquina de este interesante proyecto de OCLC. Vid. *Cataloging Internet Resources: A Manual and Practical Guide* [documento HTML], Nancy B. Olson, ed. 2<sup>nd</sup> ed. Dublin, Ohio: OCLC, rev. Agosto de 2002. Disponible en: <http://www.oclc.org/support/documentation/worldcat/cataloging/internetguide> y, *Bibliografic Formats and Standards. 3th ed.* [documento HTML]. Ron Jones, comp. Dublin, Ohio: OCLC, rev. Mayo de 2003. Disponible en: <http://www.oclc.org/bibformats>

<sup>11</sup> Vid. *supr.*, nota 2.

<sup>12</sup> Así por ejemplo, se han desarrollado múltiples estándares de metadatos a partir del conjunto de elementos del Dublin Core. Tal es el caso, por ejemplo, del esquema del (*Educational Network Australia* (EdNA), o del *Australian Government Locator Service* [AGLS] para la gestión, organización y recuperación de información administrativa. En cuanto a la compatibilidad o mapeo de la DCMI con otros formatos, podemos destacar por ejemplo, todos los esfuerzos que se han realizado este año por mapear los elementos de la ISO 19115 para la descripción de la información geográfica (FGDC) con el Dublin Core.

<sup>13</sup> Vid. *Cataloging and Metadata* (OCLC Products and Services) <http://www.oclc.org/services/cataloging>

<sup>14</sup> La mayoría de los OPACs (*Online Public Access Catalogs*) se han convertido en WebPACs (*Web Public Access Catalog*). En realidad los WebPACs no son otra cosas que los catálogos de acceso público tradicionales, a los que se da acceso a través de la Web. Todos los creadores de Sistemas Integrados de Gestión de Bibliotecas (SIGB) incluyen, como uno de sus módulos, la versión Web del catálogo o base de datos bibliográfica. Los WebPACs son la evolución lógica de los catálogos bibliotecarios adaptados a la interfaz Web, insertándose así, de una forma coherente, como un servicio más de información Web..

<sup>15</sup> Hyperion (Sirsi corporation): <http://www.sirsi.com/Sirsiproducts/hyperion.html>

<sup>16</sup> Ex-Libris Products: <http://www.exlibris-usa.com/products1.html>

<sup>17</sup> A finales de 1996 es cuando el consorcio Web publica el primer borrador de trabajo de la especificación de XML. Vid. World Wide Web Consortium. *Extensible Markup Language (XML): W3C Working Draft 14-Nov-96* [documento HTML]. Tim Bray, C. M. Sperberg-McQueen, eds. W3C, 14 de noviembre de 1996. Disponible en: <http://www.w3.org/TR/WD-xml-961114.html>

<sup>18</sup> El *Resources Description Framework* está aún en fase de desarrollo, íntimamente vinculado al proyecto de la Web semántica del W3C. Toda su complejidad y uso se expresa en siete especificaciones distintas: una recomendación del 22 de febrero de 1999 sobre el modelo y la sintaxis (que está siendo revisada en la actualidad por Dave Beckett); y seis documentos de trabajo, recientemente publicados como "propuestas de recomendación" el 15 de diciembre de 2003. Vid. RDF <<http://www.w3.org/RDF/#specs>>. Asimismo RDF depende de otras especificaciones como los espacios de nombre (*namespaces*) <<http://www.w3.org/TR/REC-xml-names>> y la sintaxis genérica de los URIs (*Uniform Resource Identifier*) <<http://www.isi.edu/in-notes/rfc2396.txt>>.

<sup>19</sup> [Michael Sperberg-McQueen es uno de los padres de la Iniciativa de Codificación Textual \(TEI\) y miembro del W3C. El 10 de diciembre de 2003 recibió la "Copa XML" 2003, galardón que se otorga a visionarios industriales o a pioneros en contribuciones al Extensible Markup Language \(XML\). \(Más información sobre Sperberg-McQueen http://www.w3.org/People/cmsmcq/\).](http://www.w3.org/People/cmsmcq/)

<sup>20</sup> XML4Lib: <http://sunsite.berkeley.edu/XML4Lib>

<sup>21</sup> Anteriormente la Library of Congress había desarrollado DTDs, en principio en SGML y posteriormente en XML. La creación del esquema demuestra una adaptación tecnológica rápida y necesaria del formato MARC a los nuevos estándares de representación/recuperación de información: El esquema XML para MARC21, se puede consultar en: <http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd>

<sup>22</sup> Esta idea de la confluencia de ambas escuelas coincide con lo que he dicho muchas veces, de manera informal, e incluso elemental, que "la Web semántica no es más que una Web más bibliotecaria". Sin embargo la idea que subyace a RDF y a los lenguajes de esquema o deducción, explicada por Tim Berners-Lee en su obra *Tejiendo la Web* [2000, p. 167 y ss.] aproxima a la idea de organización bibliotecaria de la Web como si fuera una gran base de datos, con información estructurada previamente y legible por máquina.

<sup>23</sup> *DCMI type Vocabulary*: <http://dublincore.org/documents/dcmi-type-vocabulary>

<sup>24</sup> A este tipo especial de "portales académicos" de información Web, pertenecen, por ejemplo, en el Reino Unido: OMNI *The UK's Gateway to High Quality Internet Resources in Health and Medicine* o ADAM, *the Art, Design, Architecture & Media Information Gateway*; y en Australia, por ejemplo: ATUA, *Australian Trade Union Archives*; o WebLaw entre otras recogidas en el ASGF, *Australian Subject Gateways Forum* <<http://www.nla.gov.au/initiatives/sg/gateways.html>>.

<sup>25</sup> *Vid. Overview of Documentation for DCMI Metadata Terms* <<http://dublincore.org/documents/dcmes-qualifiers>>, documento que reemplaza a la recomendación obsoleta de la DCMI <<http://dublincore.org/documents/2000/07/11/dcmes-qualifiers>>.

<sup>26</sup> Estos tres sectores informativos recibieron un tratamiento especial en la novena reunión del DC en Japón (DCMI-2001) donde se celebraron sesiones específicas para cada uno de estos sectores. El reconocimiento de las distintas comunidades informativas es algo habitual en el ámbito de los metadatos, y así lo demuestran los lemas de las siguientes conferencias de la DCMI: DC-2002: *Metadata for e-Communities: Supporting Diversity and Convergence* (Florenca, 2002) <<http://www.bncf.net/dc2002>> y DC-2003: *Supporting Communities of Discourse and Practice-Metadata Research & Applications* (Seattle, 2003) <<http://www.ischool.washington.edu/dc2003>>.

<sup>27</sup> DC Library Application Profile (Rebecca Guenther): <http://dublincore.org/documents/library-application-profile>

<sup>28</sup> *Vid. Z39.50 Profiles*: <http://lcweb.loc.gov/z3950/agency/profiles/profiles.html>

<sup>29</sup> *Vid. OCLC Research Activities and the Open Archives Initiative* <http://www.oclc.org/research/projects/oai/index.shtm>

<sup>30</sup> *Vid. El Registro de Metadatos de la DCMI*: <http://www.dublincore.org/dcregistry/index.html>

<sup>31</sup> Las Bibliotecas Nacionales de Australia <<http://www.nla.gov.au/initiatives/standards.html>>, Nueva Zelanda <<http://www.natlib.govt.nz/en/whatsnew/4initiatives.html#meta>> o Finlandia son casos paradigmáticos en este sentido ya que no sólo se ocupan de generar una política de descripción basada en metadatos para las respectivas bibliotecas nacionales, desarrollando estándares a tal efecto, sino que lo hacen con la pretensión de arbitrar una política bibliotecaria digital para todos los servicios de información digital de su dominio geográfico de país. Asimismo estas bibliotecas nacionales han generado sus esquemas de metadatos *ad hoc* para la preservación del patrimonio digital —*Vid. el servicio australiano de información sobre preservación digital PADI (Preserving Access to Digital Information)* <<http://www.nla.gov.au/padi>>—. Merece mención especial, hablando de la implicación que

tienen las bibliotecas nacionales con el mundo de la gestión de la información digital, el hecho de que la Biblioteca Nacional de Finlandia, es la primera institución afiliada a la DCMI <<http://www.lib.helsinki.fi/english>> donde culminan muchos de los objetivos establecidos en el Nordic Metadata Project (1995-1998) <<http://www.lib.helsinki.fi/meta/nmfinal.htm>>.

<sup>32</sup> Me parece confusa la denominación de este servicio —*Sistema de Recuperación de Recursos electrónicos*— ya que con este nombre uno se espera una *subject gateway*, es decir una colección virtual organizada de recursos electrónicos, no el acceso a una colección real digital o biblioteca digital. Para más información sobre este sistema, *vid.* [http://193.146.129.47/opacweb/AtrilW3\\_html/AtrilW3\\_html/presentacion.html](http://193.146.129.47/opacweb/AtrilW3_html/AtrilW3_html/presentacion.html)

<sup>33</sup> Para más información sobre el catálogo unificado basado en el protocolo OAI, consultar: *Electronic Thesis/Dissertation OAI Union Catalog*, en: <http://rocky.dlib.vt.edu/~etdunion/cgi-bin/index.pl>

<sup>34</sup> Para ilustrar las conclusiones que apuntamos aquí, utilizamos como ejemplo paradigmático el código fuente de la página principal de la Biblioteca de la Universidad Autónoma de Barcelona, porque en ella confluyen muchas de las prácticas habituales en el mercado de los sitios de bibliotecas en España.

<sup>35</sup> Sobre las actividades de Iris-index: <http://www.rediris.es/si/iris-index/actividades>; Sobre Webber: <http://www.rediris.es/app/webber>; Sobre MetaWebber: <http://www.rediris.es/si/iris-index/herramientas/metawebber>

<sup>36</sup> El Consorcio TEIDE está formado por el departamento de Topografía e Ingeniería Cartográfica de la Universidad Politécnica de Madrid, el departamento de Sistemas de Información de la Universidad Jaume I, y el departamento de Informática e Ingeniería de Sistemas de la Universidad de Zaragoza, y tiene como origen el proyecto de IDEMetadatos. Más información en la comunidad virtual RedGeomatica en: <http://dityc.euittu.upm.es/redgeomatica/metadatos.php>

<sup>37</sup> A pesar del número de suscriptores, la poca participación en esta lista (85 mensajes desde el 25/06/2003 al 29/10/2003) denota justamente la falta de desarrollo en estos temas en nuestro país.

<sup>38</sup> Desde el año 2000, encontramos colegas con ese título o función, por ejemplo: Dan Foley. Metadata Librarian, <<http://www.dlib.org/dlib/january00/chandler/chandler-appendixA.html>>; Jeremy Frumkin, Metadata Librarian <<http://digital.library.arizona.edu/~jfrumkin>>. Incluso están empezando a aparecer convocatorias de trabajo para bibliotecarios —aún no en nuestro país— bajo ese perfil. Por ejemplo, la biblioteca de la Universidad de Tennessee anunciaba un puesto de “*metadata librarian*” en agosto de 2002 <<http://toltec.lib.utk.edu/~lss/search/metadata.html>>; más recientemente, la biblioteca Albert R. Mann de la Universidad de Cornell, también en Estados Unidos pedía, el 19 de agosto de 2003, un “*metadata librarian*” *Vid.* <<http://www.acrlny.org/jobs/j30820a.htm>> o <<http://nasig.org/jobs/cornellmetadat.htm>>.

<sup>39</sup> Todos los recursos electrónicos consignados en esta bibliografía se han consultado por última vez el 17 de octubre de 2003, por ello no se consigna la fecha de acceso en cada referencia bibliográfica de los artículos, publicaciones o sitios web.