

NUEVAS PERSPECTIVAS PARA LA CATALOGACIÓN: METADATOS VERSUS MARC

Virginia Ortiz-Repiso Jiménez*

Resumen: Se analizan los sistemas de metadatos que existen en diferentes colectivos para solucionar los problemas que plantea la búsqueda y recuperación efectiva de recursos en Internet. Se comparan, además, los nuevos sistemas con los registros bibliográficos utilizados en las bibliotecas. Por último, se estudia la posibilidad de que estos sistemas puedan reemplazar a las herramientas que se utilizan de forma generalizada en las bibliotecas: el formato MARC y las reglas de catalogación.

Palabras clave: metadatos; MARC; reglas de catalogación; catálogos hipertextuales.

Abstract: Metadata systems to solve searching and retrieval Internet problems are analysed. New metadata systems and traditional bibliographic records are compared. Finally, the possibility that those systems may replace the traditional library tools, such as MARC format and Cataloguing Rules, is examined.

Key words: metadata; MARC; cataloguing rules; hypertextual catalogues.

1 Introducción

Las bibliotecas y los centros de documentación han centrado sus esfuerzos desde la segunda guerra mundial en proporcionar acceso a las publicaciones. El conjunto de normas y estándares que se ha desarrollado en las últimas cuatro décadas ha permitido llevar a cabo un acceso a colecciones controladas de información. Las Anglo-American Cataloguing Rules, las reglas de catalogación españolas, los ISBN e ISSN, el formato MARC y el crecimiento de las bibliografías nacionales representan el gran esfuerzo profesional realizado para globalizar y normalizar el control bibliográfico. En el mismo sentido, la catalogación cooperativa, la conversión retrospectiva de catálogos y los servicios de indexación y resumen han dado como resultado grandes bases de datos que se han puesto a disposición de la comunidad internacional. Estos instrumentos contienen información acerca de información, lo que Terence R. Smith (1) denomina el entorno de meta-information de una biblioteca.

Sin embargo, en los últimos cinco años ha surgido un nuevo reto para las bibliotecas provocado por el crecimiento vertiginoso de la información electrónica y su consecuencia lógica, el nacimiento de la biblioteca digital. La diferencia fundamental entre las bibliotecas tradicionales y las digitales estriba en que en estas últimas el control de las colecciones se realiza de forma digital y el acceso a la información en esas colecciones se basa casi en su totalidad en tecnología digital.

* Departamento de Biblioteconomía y Documentación. Universidad Carlos III de Madrid. Correo electrónico: virginia@bib.uc3m.es.

Recibido: 1-6-98; 2.ª versión 30-4-99.

La tecnología digital cambia las características básicas de una biblioteca tradicional. En primer lugar, la organización de la información en objetos físicos se ve reemplazada por una organización más flexible en objetos lógicos. En segundo lugar, el acceso físico y único a la colección se reemplaza por un acceso lógico y múltiple a los objetos de información. Y por último, el entorno de la meta-information de la biblioteca se transforma: un objeto de información en formato digital permite el uso de tecnología también digital para extraer información del propio recurso informativo.

Este trabajo se centra principalmente en el impacto que, en el mundo bibliotecario en general y en el proceso catalográfico en particular, están teniendo los nuevos sistemas y estándares que se utilizan para describir, recuperar y acceder a los objetos de información en un entorno de red.

Tres son los objetivos básicos de este estudio. Por un lado, analizar los sistemas de metadatos que existen y/o están emergiendo en diferentes colectivos para solucionar los problemas que plantea la búsqueda y recuperación efectiva de recursos en Internet. Por otra, comparar estos sistemas de metadatos con los registros bibliográficos tradicionalmente utilizados en las bibliotecas. Y, por último, analizar si estos sistemas pueden reemplazar y/o mejorar las herramientas que se utilizan de forma generalizada en las bibliotecas: el formato MARC, las Reglas de Catalogación y los sistemas de recuperación.

Voy a comenzar describiendo de forma somera cada uno de estos conceptos para más tarde establecer relaciones entre ellos.

2 Internet y las herramientas de búsqueda

Encontrar información relevante en el World Wide Web se ha convertido en un problema que se incrementa día a día debido al crecimiento explosivo de los recursos de red. En efecto, el rápido desarrollo del WWW y el incremento de los recursos disponibles en Internet ha hecho necesaria la creación de herramientas que proporcionen acceso a los millones de documentos que existen en formato electrónico.

La Red, a través de los motores de búsqueda, utiliza métodos automáticos para identificar los recursos de Internet. Existe una gran cantidad de programas que navegan automáticamente a través de los espacios Web, buscando enlaces, recuperando documentos, indizándolos y creando bases de datos con ellos. Ahora bien, estos sistemas de gran potencia, recuperan gran cantidad de documentos pero con muy poca precisión. La causa principal no es que los métodos automáticos que utilizan describan de forma poco adecuada los recursos en la Red, sino porque los propios documentos de la red carecen de datos suficientes de descripción. Como resultado, hay mucho ruido a la hora de la recuperación y muchas veces se convierten en inaccesibles para el más paciente de los buscadores en la red. Además, mientras que el tamaño de Internet crezca exponencialmente, será cada vez más difícil moverse por esta «masa» indiscriminada de resultados de búsqueda. Al mismo tiempo, se corre el peligro de que las bases de datos creadas por estos mecanismos puedan llegar a ser mayores, incluso, que los propios recursos de Internet. El problema reside en que la indexación que realizan, aunque útil, es un sustituto muy pobre de la riqueza que proporciona la descripción del recurso.

A medida que la Red ha ido creciendo y que estos métodos de descripción se han mostrado cada vez más inadecuados, un amplio colectivo, que incluye profesionales de la información, informáticos de redes, diseñadores de *software*, investigadores y un amplio etcétera, ha empezado a plantearse la necesidad de crear descripciones y catálogos que identifiquen los recursos electrónicos en Internet de una manera más eficaz y permitan una búsqueda y recuperación más efectiva.

El primer intento tuvo lugar en 1995 cuando se celebró el primer seminario del Dublin Core Metadata al que fueron invitados bibliotecarios, especialistas en lenguajes de marcas, expertos en análisis documental e investigadores de bibliotecas digitales con el fin de promover un estándar que permitiera una recuperación más efectiva de los recursos en las redes. El objetivo primordial era llegar a un consenso interdisciplinar e internacional para diseñar un código de descripción de recursos.

Desde entonces, encontrar una solución a este problema se ha convertido en materia de estudio para muchas instituciones y asociaciones como son: la Biblioteca del Congreso, OCLC, la American Library Association, la National Science Foundation, etc.

Todas las personas y colectivos involucrados en los problemas que plantea la búsqueda y recuperación de recursos en las redes coinciden en que la mejor manera de solucionar el problema es la creación de metadatos que describan los recursos.

3 Metadatos y meta-information en las bibliotecas

Pero ¿qué son los metadatos?

Aunque el término *metadata* o metadatos no significa exactamente lo mismo en todos los colectivos (informáticos, bibliotecarios, gestores de bases de datos o físicos) en el entorno de las bibliotecas digitales se usa cada vez más para referirse a la disponibilidad de datos sobre los recursos de información. El término hace referencia a cualquier dato usado para ayudar a identificar, describir y localizar los recursos electrónicos en un entorno de red.

Los metadatos son, en otras palabras, datos acerca de datos. Los metadatos en sí no son algo nuevo. Lo que es nuevo hoy en día es la multitud de formatos (algunos muy simples en su descripción, otros más complejos) que se están diseñando y la forma de usarlos.

En el mundo bibliotecario el concepto de metadatos nos es familiar. Los registros bibliográficos que se han creado a lo largo de muchos años son esencialmente metadatos. Proporcionan información descriptiva y analítica sobre un objeto de información. Los catalogadores los han empleado como método descriptivo desde hace décadas, bien como registros MARC en los OPACs, bien como fichas catalográficas en los catálogos manuales. Por tanto, un registro catalográfico no es otra cosa que un conjunto de metadatos. Nosotros podríamos llamarlo catalogación pero para algunas personas este término conlleva una carga excesiva: formatos MARC, Reglas de Catalogación, etc. El término metadatos se utiliza pues como un término neutral (2).

El término metadatos, en el contexto de las bibliotecas, sean éstas tradicionales o digitales, se usa para referirse a información que:

- proporciona (normalmente de forma breve) una caracterización de un recurso de información específico en las colecciones de la biblioteca;
- se almacena en los catálogos; y
- se usa principalmente para ayudar al usuario a acceder a los recursos de información de su interés.

Los metadatos se usan para describir recursos digitales y no digitales localizados en un sistema distribuido en un entorno de red. Para que su uso sea eficaz deben estandarizarse. Tradicionalmente los metadatos incluían las reglas de catalogación y los formatos. Sin embargo, debido al gran incremento de la información electrónica, se han extendido, reflejando las necesidades de recuperación y uso de la información en entornos de red. Y, en lugar de usar el término «reglas de catalogación electrónicas», se utiliza el de metadatos.

La definición más exacta del término en el contexto Internet es la realizada por Tim Berners-Lee: *información legible por ordenador sobre recursos web* (3). La frase legible por ordenador es clave. Hablamos de información que los diferentes *softwares* pueden utilizar para hacernos más fácil la tarea de búsqueda y recuperación de recursos en la red. El mundo del World Wide Web es un mundo de información y parte de esa información es información sobre información. Ésta puede localizarse en el propio documento o puede formar una entidad separada, o puede transferirse acompañando al documento.

En efecto, los metadatos son importantes en la recuperación de la información global en Internet por distintas razones:

- Permiten indizar gran cantidad de datos de diferentes tipos sin necesidad de utilizar un gran ancho de banda ya que se indiza la representación del objeto y no el objeto en sí.
- Ayudan a descubrir y recuperar recursos en la red ya que analizan el contenido del objeto en profundidad.
- Comparten e integran recursos de información heterogéneos y localizados en sitios muy diversos.
- Pueden controlar el acceso a información restringida.

El uso de metadatos para organizar el contenido de la información en Internet se está extendiendo cada vez más. Los metadatos se usan de tres formas distintas: pueden ir acompañando al propio documento o recurso (en la cabecera de un documento HTML o cabeceras SGML en general); pueden formar un fichero separado de meta-information para describir recursos distintos a los de HTML (sonido, imagen, programa de ordenador, etc.); o también se puede crear con ellos una base de datos central o distribuida con punteros a los recursos que describen (4).

Los metadatos han evolucionado desde formatos de estructura muy simple a formatos más complejos. Y se han movido desde estándares emergentes propietarios a estándares internacionales. Y, muy importante, los metadatos que se crean se pueden compartir con otros.

La información que contienen es variada: desde información descriptiva similar a la que estamos acostumbrados a ver en las bibliotecas, hasta información que ayuda a la aplicación cliente a tomar una decisión sobre el formato o sobre la localización. Sus usuarios pertenecen también a distintas categorías, desde aquéllos que de-

sean conocer sólo los términos de disponibilidad de un recurso a aquéllos que desean tener más información sobre el contenido del objeto informativo. Los recursos son, además, de distinta tipología: algunos tienen una existencia efímera y sólo necesitan una descripción somera; unos pueden ser simples, otros más complejos. En este sentido nos encontramos con una tipología parecida a la que existe en el mundo impreso.

Bearnan y Sochats (5) definen seis tipos de información básica que deben contener los metadatos relacionados con diferentes propósitos:

1. Identificación y recuperación.
2. Condiciones de acceso y requisitos de uso.
3. Aspectos estructurales.
4. Aspectos contextuales.
5. Aspectos de contenido.
6. Uso de ese recurso (historia del recurso).

Estos aspectos se pueden aplicar a la mayor parte de los metadatos y se pueden utilizar tanto en recursos digitalizados como en los no digitalizados.

La idea que subyace en los estándares de metadatos que se están definiendo, especialmente en el Dublin Core, como veremos a continuación, es crear una estructura simple que permita al propio creador del documento asignar sus metadatos. En este sentido, los metadatos pueden crearse, además de por el propio autor/creador del recurso, con la ayuda de herramientas, cada vez más numerosas, creadas *ex profeso* para la generación automática de metadatos; por la organización que gestiona los recursos informativos; o por un servicio de información, es decir creadores externos que puedan realizarlos bien con medios humanos (bibliotecarios, por ejemplo) o automáticos (SOIF-Summary Object Interchange Format- en Harvest).

4 ¿Qué estándares siguen los metadatos?

Las bibliotecas, hasta el momento, parecían ser las únicas instituciones que contaban con una sintaxis ampliamente aceptada y bien regulada de creación de metadatos (MARC, reglas de catalogación, etc.). Pero sin embargo, en los últimos años, con la extensión y expansión de recursos electrónicos en Internet, están emergiendo otras sintaxis, que proporcionan también metadatos, y que prometen una mayor funcionalidad.

El estándar de metadatos ideal será aquel que:

- sea fácil de crear y mantener;
- utilice una semántica que pueda entenderse de forma común;
- pueda crearse de forma automática;
- pueda describir la forma, el contenido y la localización de la información;
- su estructura permita contenerlos en otros objetos;
- se pueda usar para construir múltiples índices (www, texto, etc.);
- pueda interoperar en los sistemas de indexación que existen; y
- pueda ampliarse según las necesidades.

Todos los conjuntos de metadatos que se están empleando para describir los recursos en las redes, bien sea como parte de los documentos, bien como entidades individuales pero enlazadas a los documentos, siguen la norma SGML *Standard Generalized Markup Language* (Lenguaje de marcas estándar generalizado). Estándar internacional desde 1986 (ISO 8879).

Aunque SGML es un estándar generalizado, no es, realmente, un lenguaje de marcas como tal. No proporciona por sí sólo un lenguaje de marcas que uno puede simplemente llevarse a casa y aplicarlo a una carta, una novela, un artículo o un registro catalográfico. SGML es conocido como un metalenguaje, esto significa que no es un único lenguaje sino una norma amplia para construir lenguajes de marcas. Proporciona una sintaxis para definir y expresar la estructura lógica de los documentos, así como las convenciones para nombrar los componentes o elementos de los documentos.

Se puede decir que SGML es un conjunto de reglas formales para definir lenguajes de marcas específicos para tipos específicos de documentos. Este lenguaje de marcas específico se denomina Definición del tipo de documento (DTD). Por ejemplo, la Asociación de Editores Americanos junto con OCLC ha desarrollado un juego de 3 DTDs: uno para libros, otro para publicaciones periódicas y otro para artículos de revista. HTML, tan conocido por todos, es un DTD específico de SGML. Puede haber, y de hecho ya existen, DTDs para bibliotecas, para museos y para archivos. No hace mucho el profesor Larson ha creado en Berkeley un DTD para el USMARC con el propósito de utilizarlo en un prototipo de catálogo bibliográfico que emplea tecnología de recuperación avanzada (6).

Además, los DTDs que se comparten y se siguen en una comunidad específica se pueden convertir y de hecho se convierten en estándares internacionales. Es el caso, por ejemplo, de la Asociación Americana de editores con la norma ISO 12083.

La combinación de marcas descriptivas y DTDs permite que múltiples tipos de software puedan procesar documentos codificados en SGML. Es decir, SGML es independiente del hardware y software que se utilice. La información que se crea utilizando esta norma no se vuelve obsoleta si el software se queda anticuado o se cambia de programa. En esto se parece al formato MARC pero con una diferencia muy clara. La sintaxis SGML y sus reglas son precisas y es posible diseñar software que pueda ajustarse a cualquier DTD. Normalmente el software tiene un conjunto de herramientas que permite al usuario adaptar su funcionalidad a su DTD. Como resultado, el mercado de software SGML puede ser en principio cualquiera. En este aspecto es muy diferente a lo que ocurre en el mercado de software para el formato MARC, que va dirigido fundamentalmente a bibliotecas, algunos archivos y algunos museos. Estas comunidades son históricamente comunidades pequeñas y pobres. Y, aunque no es mi intención criticar en este momento el software disponible para bibliotecas, los recursos que están a nuestro alcance reflejan recursos limitados. Sin embargo, el mercado SGML incluye virtualmente a cualquiera. De esta forma SGML está siendo utilizado por muchos productores de software de muy distintos tipos. Desde empresas que diseñan software específico para tratamiento de texto, como Word Perfect o Word de Microsoft, hasta empresas que se dedican a diseñar software para el sector de automóviles.

Para hacernos una idea de la amplitud y variedad de los usuarios potenciales de SGML no hace falta más que revisar, por ejemplo, las personas e instituciones que

están suscritas al Listserv dedicado a SGML en Carolina del Norte: el Research Library Group, Silicon Graphics, la Universidad de Berkeley, Dialog y un largo etcétera. La lista de fabricantes de *software* relacionado con SGML refleja, también, la confianza en el potencial de este mercado: Word Perfect, Microsoft, Xerox, Open Text, Frame, etc. Firmas que como todos sabemos no están interesadas en mercados pequeños.

5 MARC versus SGML

He señalado anteriormente que un registro bibliográfico es un tipo de metadatos. Difiere de otros en que desde hace más de tres décadas usa las reglas de catalogación para modelar los datos, el formato MARC como esquema de codificación y, además, utiliza sistemas en línea propietarios para la recuperación de la información. La mayoría de los sistemas de metadatos en Internet no tienen, evidentemente, la misma estabilidad, ni unas normas tan bien reguladas, ni un esquema de codificación, ni sistemas específicos para recuperar información. Pero esto está cambiando y tenemos que examinar el impacto que está teniendo en el mundo bibliotecario.

El formato MARC es hoy en día un conjunto complejo de estándares para describir, almacenar, manipular y recuperar datos bibliográficos legibles por ordenador. Es un estándar altamente desarrollado que se diseñó originalmente en los 60 para la descripción de libros impresos y que ha seguido adaptándose para proporcionar descripción, acceso y localización de la información de los recursos en la Red. En este sentido, el campo 856, recientemente incorporado (aunque no se utiliza en nuestro país) permite una descripción y enlace entre el registro MARC y el recurso electrónico que describe.

El formato MARC que ha sentado las bases de cooperación y comunicación de información bibliográfica en el mundo bibliotecario tiene, sin embargo, una serie de características negativas que podrían sintetizarse de la siguiente forma:

- Está estrictamente controlado, cualquier cambio o adiciones al formato tarda años en realizarse. Por ejemplo, la tilde de la ñ que tanto se utiliza en las URLs de las páginas personales ha tardado más de dos años en incorporarse al juego de caracteres del USMARC. El proyecto Intercat de OCLC no podía operar con URLs que tuvieran ese carácter.
- Es laborioso, lento y costoso de realizar ya que debido a su complejidad debe realizarse por profesionales cualificados.
- Aunque está compuesto por campos de longitud variable, está limitado a una longitud máxima de 100.000 caracteres que, si bien es más que suficiente para los registros bibliográficos tradicionales de las bibliotecas, no lo es para otro tipo de registros de otros ámbitos (archivos, por ejemplo).
- Se adapta muy mal a información estructurada jerárquicamente, ya que se basa en una estructura plana. Se diseñó para describir y acceder a la información de un registro bibliográfico y no para establecer relaciones entre registros. Se diseñó para describir y proporcionar acceso a la información contenida en los registros bibliográficos, pero no para establecer enlaces entre los campos de

un mismo registro o entre diferentes registros de una misma base de datos (7). Incluso cuando en los OPACs existen enlaces entre autores, materias y títulos, no va más allá de una búsqueda exacta de, por ejemplo, libros del mismo autor, materia o palabras similares en el título. El MARC no está pensado para describir y proporcionar acceso a colecciones más complejas a través de niveles que descienden en el análisis. Una solución posible es emplear registros múltiples, jerárquicamente interrelacionados, de varios niveles de análisis: nivel de colección, subunidad e ítem, indicándolo en la cabecera del formato. Pero el uso de múltiples registros supone, sin embargo, muchos problemas de control que no han sido adecuadamente solucionados ni en el formato ni en el software disponible para gestionarlo. La comunidad de usuarios MARC ha sido incapaz de dirigir el desarrollo del software apropiado. Esto no es extraño, ya que todos sabemos que la comunidad de bibliotecas, archivos y museos nunca cuenta con brillantes presupuestos.

Ante estas características, SGML proporciona un marco de trabajo prometedor por diversas razones:

- Puede tratar información jerárquicamente interrelacionada en tantos niveles como se necesiten.
- No tiene una limitación en el tamaño de los documentos.
- Es un estándar internacional adoptado por un número creciente de instituciones gubernamentales, de investigación y de la industria.
- Tanto SGML como MARC son lenguajes de marcas descriptivos que crean textos estructurados. Pero SGML permite una flexibilidad máxima en el uso del texto. El usuario puede controlar los formatos de indización, presentación e impresión. Su estructura facilita la creación de bases de datos más sofisticadas que las basadas en MARC y permite la indización y recuperación del documento como documento y componentes del documento mediante búsqueda booleana, adyacencia, proximidad, ranking de relevancia, etc. La estructura del texto SGML soporta además navegación en línea avanzada.
- Los registros pueden estar interrelacionados en distintos ficheros: el guión de una película y el vídeo; un mapa y el CD-ROM que lo contiene, por ejemplo.

En definitiva, la flexibilidad del texto estructurado en SGML es incuestionablemente superior a la jerarquía plana del texto estructurado en MARC. Ahora bien, este tema puede abordarse desde dos perspectivas bien distintas. Por una parte, se puede pensar que el uso de los estándares tradicionales es suficiente para analizar y describir los recursos en las redes. Es decir, continuar con el uso de los formatos MARC (el InterCat Project o CATTRIONA son ejemplos de ello). Por otra, se puede pensar que la información electrónica requiere el desarrollo de nuevos métodos y técnicas para describir y facilitar la recuperación de los recursos de información. En este sentido, hay que pensar que las bibliotecas nunca han sido innovadoras, en cuanto a tecnología se refiere, sino más bien adaptadoras de las innovaciones tecnológicas que se producen en y para otros ámbitos. En aras de adaptarse al entorno es necesario usar otros formatos que se ajusten mejor a la estructura hipertextual de la red. La adopción de estos formatos facilitará la creación de bases de datos que se adapten mejor al nuevo entorno que las posibilidades que ofrece el formato MARC.

6 Algunos proyectos de metadatos

Son numerosos los proyectos de diseño de metadatos que se están llevando a cabo. Los más importantes, atendiendo a su uso y a la organización que los ha creado, son:

- DC: Dublin Core Metadata.
- RDF: Resource Description Framework.
- TEI: Text Encoding Initiative.
- URC: Uniform Resource Characteristic/Citations.
- MARC DTD: MAchine Readable Cataloging Document Type Definition.
- GILS: Government Information Locator Service. Usado por los gobiernos de USA y Canadá.
- EAD: Encode Archive Description que se usa en el ámbito de archivos en USA.
- DOI: Digital Object Identifier que incluye el DC para la descripción e información propia del mundo de la edición, como condiciones de adquisición, precio, etc.
- WHO IS++Templates: uno de los primeros formatos de metadatos creado para describir los recursos en las redes, se utilizó originalmente para las listas de discusión y más tarde se extendió a otros tipos de recursos.

A continuación voy a describir los cinco primeros por ser los que mejor se ajustan al entorno de las bibliotecas. El análisis de estos proyectos servirá, también, para analizar la relación que existe entre SGML y un sistema de biblioteca en línea.

6.1 Dublin Core

El Dublin Core es la forma abreviada para el Dublin Metadata Core Element Set (Conjunto básico de elementos metadatos de Dublín). Es un proyecto cooperativo de ámbito internacional —promovido por OCLC y NCSA (National Center for Supercomputer Applications)— cuyo objetivo principal es crear un conjunto de elementos de datos que describan los documentos electrónicos de las redes con el fin de facilitar su búsqueda y recuperación.

Ese conjunto debe definirse de forma que los autores y editores de documentos en Internet puedan crear sus propios registros de metadatos sin necesidad de poseer una formación específica. Se pretende conseguir un nivel bibliográfico intermedio entre la estructura detallada del MARC y la cabecera del TEI, y la indización automática que realizan los motores de búsqueda de Internet. El Dublin Core está compuesto por un conjunto mínimo de elementos que facilitan la recuperación de información en la Red. Hay que decir que el formato USMARC está considerando incluir el conjunto de elementos del Dublin Core como parte central de su propio registro.

El Dublin Core está diseñado para facilitar la recuperación de recursos en las redes de una forma similar a un catálogo de biblioteca, pero con una estructura mucho más sencilla. Está formado por 15 elementos de datos y su punto más fuerte es que el diseño es tan intuitivo que los propios proveedores de información pueden codificar sus documentos al mismo tiempo que los crean. Una característica también

destacable es que estos elementos ponen más énfasis en el acceso y recuperación de la información que en la descripción, aspecto éste que siempre se ha criticado en el formato MARC.

Desde 1995 en que nació el proyecto, han tenido lugar cuatro seminarios para llegar a acuerdos sobre el Dublin Core: Dublin Metadata Workshop (1995), Warwick Metadata Workshop (abril 1996), Image Metadata Workshop (septiembre 1996) y el seminario celebrado en Canberra, Australia (1997). Los participantes en estos grupos de trabajo pertenecen a diferentes colectivos: bibliotecarios, archiveros, investigadores, informáticos, desarrolladores de *software*, editores y miembros del Engineering Task Force (IETF). Es fácil suponer que las aproximaciones al tema que hacían tan diversos colectivos eran muy distintas. Algunos estaban interesados en recursos electrónicos en general, otros sólo en un tipo de materiales específicos, como textos de humanidades o metadatos geoespaciales. Varios en los servicios y protocolos de red que podrían hacer uso de metadatos, mientras que otros adoptaban el punto de vista del autor, del editor o del usuario final.

Lo único que unía a todos los participantes era la creencia certera de que un conjunto de metadatos estándar era mejor que nada, ya que hasta el momento no había ningún acuerdo ni normalización alguna. En seguida se hizo evidente que un único conjunto de elementos de datos no podría satisfacer unas necesidades tan amplias y divergentes. El énfasis se puso entonces en definir un conjunto de elementos que pudiera utilizarse por los proveedores de información para describir sus propios recursos, fuera cual fuera su tipología. Se recomendó, también, que los autores y editores proporcionaran metadatos al mismo tiempo que los propios datos. Además, esta combinación de datos podría convertirse en un estándar oficial o en un estándar de hecho y tener diferentes usos.

El registro descriptivo fruto del primer seminario es el conocido como Dublin Core. No es un conjunto básico en el sentido de que requiera un número mínimo de elementos sino en el sentido de que está compuesto por un número pequeño de elementos que pueden aplicarse y entenderse de forma genérica por el mayor número de personas y colectivos. La atención se fijó, sobre todo, en describir propiedades intrínsecas al recurso: autor, materia, actualización, etc. y no en propiedades externas como coste, limitaciones de acceso, etc. Y siempre basándose en el principio de extensibilidad, es decir, cada grupo específico podría ampliar el conjunto de elementos básicos para adaptarlo a sus propias necesidades.

Los elementos que lo componen son 15:

1. Título
2. Autor o creador
3. Materias y palabras clave
4. Descripción
5. Editor
6. Otras contribuciones
7. Fecha
8. Tipo de recurso
9. Formato
10. Identificador del recurso
11. Fuente

12. Idioma
13. Relaciones
14. Alcance
15. Gestión de derechos

Los cuatro seminarios que han tenido lugar hasta la fecha han proporcionado los siguientes resultados y documentos (8):

- Una sintaxis concreta para el Dublin Core, expresada como definición de tipo de documento. Esta sintaxis se mapeó a etiquetas HTML para que los metadatos se pudieran insertar en documentos web. Los participantes del Consorcio WWW que se dedican a indexación y recuperación decidieron insertarlos en HTML. Esto ha supuesto un gran impulso para que se extienda su implantación. Los metadatos que se crean con el Dublin Core pueden ir insertados en el documento o pueden tratarse por separado e incorporarlos a otras bases de datos.
- Warwick Framework, arquitectura que permite reunir los diferentes paquetes de metadatos pudiendo acceder a ellos y mantenerlos. Esto permite agrupar las distintas extensiones que contengan datos que no formen parte del Dublin Core.
- Una guía para el usuario que permita a los autores y a los administradores de las colecciones realizar la descripción de los recursos. Incluye una guía básica y otra más compleja según el nivel del usuario.
- Extensión del Dublin Core para recursos visuales, imágenes digitales, como fotografías, diapositivas y ficheros de imágenes.

Hay bastantes proyectos que están experimentando con el Dublin Core. Muchos de ellos incluyen la interoperabilidad entre diferentes bases de datos incluidas las de registros MARC, objetos digitales y otras formas de metadatos. Entre ellos cabe destacar los que se están llevando a cabo en:

- La Biblioteca Nacional de Australia. Esta institución junto con otros organismos oficiales está utilizando el Dublin Core para proporcionar recuperación y acceso a 70 bases de datos de gran variedad y construidas en diferentes formatos: bases de datos de derecho, de revistas, de bibliotecas y de investigación.
- En varias bibliotecas nórdicas (Nordic Metadata Project) y en el Distributed Systems Technology Centre de Australia se están realizando proyectos para convertir el MARC en registros del Dublin Core y facilitar, así, la recuperación de información en las redes.
- Library of Congress a través del Programa Nacional de la Biblioteca Digital. El objetivo es dar acceso global a colecciones digitales de archivos, bibliotecas, museos y otras instituciones. Como uno de los objetivos es proporcionar acceso MARC, está trabajando en la conversión de registros entre el Dublin Core y USMARC y viceversa.
- El estándar Z39.50 empleará también los elementos del Dublin Core para la búsqueda y recuperación de información.

La tabla que se inserta a continuación (Tabla I) muestra la conversión entre el Dublin Core y el formato IBERMARC de datos bibliográficos.

Tabla I

Esquema de conversión entre el Dublin Core y el formato IBERMARC de datos bibliográficos

<i>Elementos del Dublin Core</i>	<i>IBERMARC</i>
Título Autor o creador	245 \$a título 1XX Encabezamiento principal de autor personal, entidad, congreso o título uniforme o 7XX si son dos o tres coautores. Encabezamientos secundarios de autor personal, entidad, congreso o título uniforme.
Materias y palabras clave	653 Término de indización no controlado. 650 Encabezamiento secundario de materia, término materia. 6XX Encabezamientos secundarios de materia.
Descripción Editor Otras contribuciones	520 Nota de sumario, resumen, etc. 260 \$b Nombre del editor, distribuidor. 7XX Encabezamientos secundarios de autor personal, entidad, congreso o título uniforme.
Fecha	260 \$c Fecha de publicación, distribución, etc. 008/07-10 Fecha de publicación.
Tipo de recurso	655 Término de indización género/forma. 256 Características del archivo. 500 Nota general.
Formato	856 Localización y acceso electrónico (no existe en IBERMARC).
Identificador del recurso	886 Información de otros registros MARC no integrados. 856 \$u Uniform Resource Location (no en IBERMARC). 020 ISBN.
Fuente	776 Asiento de soporte físico adicional. 786 Fuente de datos (no existe en IBERMARC).
Idioma	500 6 533 Nota general o nota de reproducción. 008/35-37 Lengua. 041 Código de lengua. 546 Nota de lengua.
Relaciones	772 Asiento de registro de la publicación principal. 773 Asiento de documento fuente. 775 Asiento de otras ediciones. 776 Asiento de soporte físico adicional. 780 Asiento de título anterior. 785 Asiento de título posterior. 787 Asiento de relaciones no especificadas.
Alcance	500 Nota general. Espacial: 034 Datos matemáticos codificados. 255 Área de datos matemáticos. Temporal: 045 Periodo de tiempo del contenido.
Gestión de derechos	540 Nota de condiciones de uso y reproducción.

6.2 Resource Description Framework (RDF)

El Dublin Core no representa el único esfuerzo para estandarizar recursos informativos en la web. Al mismo tiempo que se desarrollaba el DC, otros grupos estaban desarrollando también sistemas de descripción de datos (Netscape's Meta-Content Framework y Platform for Internet Content Selection-PICS-).

A mediados de 1997, el W3C (World Wide Web Consortium), foro de normalización de la web, comenzó a trabajar en este ámbito. Desde entonces las especificaciones del RDF (9), contenidas en los diferentes borradores que se han ido publicando, prometen convertirse en un estándar apoyado por las organizaciones de normalización más importantes —NISO e ISO— y por las principales empresas de la red —Netscape y Microsoft—.

Como su propio nombre indica, no es tan sólo un formato de metadatos sino un marco genérico de descripción de recursos. Su objetivo principal es proporcionar la interoperabilidad entre aplicaciones que intercambian información en la web. Es decir, crear un lenguaje (modelo y sintaxis) para el intercambio de descripciones de recursos web. Puede utilizarse en una gran variedad de áreas. Por ejemplo, en un motor de búsqueda para incrementar y mejorar la recuperación; en catalogación para describir el contenido y las relaciones de contenido disponibles en una sede web, en una página o en una biblioteca digital; para describir los derechos de propiedad de las páginas web; y para expresar los permisos de accesibilidad de los usuarios y de la sede web.

El formato del Dublin Core es el que ha tenido mayor influencia en el desarrollo del RDF ya que permite no sólo descripciones simples sino que también proporciona la habilidad de cualificar las descripciones que hagan posible tanto la elaboración de un dominio específico como una precisión descriptiva. En otras palabras, las especificaciones del esquema RDF proporcionan un sistema, que puede entender el ordenador, para definir formatos específicos como el Dublin Core, permitiendo, además, especificar tipos de recursos y propiedades.

La originalidad y la mayor capacidad del RDF frente al Dublin Core reside en que:

- permite especificaciones semánticas y una sintaxis para múltiples formatos de metadatos;
- se integra en la estructura XML (Extensible Markup Language) que enriquece las posibilidades de uso —el W3C está trabajando junto a los diseñadores del Dublin Core para implantarlo en un entorno XML—;
- su definición en XML cuenta con el apoyo de Netscape y Microsoft.

Estas características lo convierten en el estándar más prometedor para la búsqueda y recuperación de información tanto en Internet como en las bibliotecas digitales.

6.3 Text Encoding Initiative (TEI)

Es un proyecto cooperativo de ámbito internacional, cuyo objetivo inicial era desarrollar un DTD para los textos que se utilizan en lingüística, literatura y estudios históricos. Sin embargo, hoy en día, el objetivo principal de este proyecto es permitir a los investigadores de cualquier disciplina intercambiar y reutilizar recursos, independientemente del *software* y del *hardware* que utilicen y sin tener en cuenta dónde estén localizados (10). Tres instituciones participan en el TEI: la Association for Computers Humanities, la Association for Computational Linguistics y la Association for Literary and Linguistic Computing. Cuenta, además, con diversos colaboradores de Estados Unidos y Europa.

Las directrices marcadas en este proyecto señalan que todos los textos TEI deben ir precedidos de una cabecera (TEI header) que describa el texto. Un comité específico, compuesto por bibliotecarios y archiveros de Europa y Estados Unidos con formación específica en catalogación, se encarga de diseñar las especificaciones de esta cabecera. El propósito de ésta es usarla como medio de control bibliográfico. Puede almacenarse como parte separada del documento al que se refiere o ir unida intrínsecamente a él. Se puede utilizar, además, para describir otros recursos en las redes cuando sea necesario hacerlo.

El nivel de dificultad de creación de la cabecera TEI depende de lo detallada que sea la información que se introduzca y el nivel de adecuación a las AACR2 y el formato MARC que se quiera alcanzar. Es decir, puede ser tan complicada como el nivel más completo de descripción de estas herramientas o mucho más sencilla. Una de sus características principales es su gran flexibilidad que permite adaptarla fácilmente a cualquier tipo de usuario.

La cabecera TEI cuenta con cuatro tipos de información:

1. Descripción del fichero: características bibliográficas del documento y su fuente.
2. Descripción codificada que muestra cómo se codificó el texto.
3. Información descriptiva: información adicional contextual y no bibliográfica: idioma, colaboradores, materias, descriptores, etc.
4. Descripción histórica sobre los cambios que ha sufrido el texto electrónico: detalles sobre revisiones, adiciones al texto, etc.

Las directrices que se proporcionan para cumplimentar los elementos más estructurados siguen las AACR2 y las ISBD. Los elementos que no están estructurados contienen texto libre.

En cuanto a los requisitos de *software*, las cabeceras del TEI pueden manipularse, buscarse y recuperarse por cualquier software que trabaje con documentos SGML. Por ejemplo, el programa Panorama. Pero todavía no existe, aunque se está trabajando en ello, una relación entre la búsqueda en Internet y los protocolos TEI (se trata de incorporar los DTDs de SGML en los URCs —Uniform Resource Characteristic/Citations—).

La mayor parte de las implementaciones se llevan a cabo en el campo de las Humanidades. El Oxford Text Archive (<http://sable.ox.ac.uk/ota>) y el Electronic Text Centre de la Universidad de Virginia (<http://www.lib.virginia.edu/etext/ETC.html>) son los ejemplos más representativos.

Respecto a su compatibilidad con otros formatos, hay que mencionar que la biblioteca de la Universidad de Virginia ha desarrollado un programa de conversión de las cabeceras TEI a USMARC y viceversa para poder incluirlas en su catálogo.

6.4 Uniform Resource Characteristic/Citations (URC)

Este proyecto (11) está promovido por el Grupo de Trabajo de Internet Task Force's Uniform Resources Identifiers. El término URC se ha utilizado generalmente para identificar:

- Información catalográfica de los recursos en línea.
- Una normalización de metadatos pero no necesariamente con propósitos catalográficos.
- Información que pueda utilizarse en el proceso de asignar un Nombre de Recurso Uniforme (URN- Uniform Resource Name) que es una propuesta para asignar localizaciones únicas independientes a los recursos en la Red. Sería el equivalente al ISBN en el mundo del libro. Información que utilizan las aplicaciones para seleccionar recursos de red.

El URC se ha concebido como un servicio general al que las personas autorizadas pueden acceder para modificar el URL asociado al URN.

Puede contener metadatos como autor, título, publicación, materias, etc., que puedan ayudar a descubrir el recurso y recuperarlo. Puede incluir, además, otro tipo de elementos como firma electrónica e información revisada para averiguar la veracidad del recurso, un elemento de acceso para restricciones de uso, y un elemento que incluya la versión e historia del documento.

El servicio URC podría incluir posibilidades de búsqueda bibliográfica. En el futuro podría ser posible conectarse a varios servidores URC, junto con algunas sedes dedicadas como pueden ser OCLC o la Library of Congress.

Todavía no existe una versión definitiva de los elementos que formarían los metadatos, pero de todos los proyectos que se han reseñado es el que contiene el menor número de elementos bibliográficos (10). Su razón principal estriba en asegurar la recuperación automática, autenticar los recursos y restringir las capacidades de acceso. No contempla, tampoco, enlaces entre registros.

6.5 MARC DTD Project (Machine Readable Cataloging Document Type Definition)

El objetivo principal de este proyecto (12) de la Biblioteca del Congreso es definir un DTD que permita la conversión de los registros MARC a una estructura SGML y viceversa sin pérdida de datos. El proyecto incluye dos partes diferenciadas. Por un lado, el desarrollo del DTD correspondiente a los tres formatos USMARC: de datos bibliográficos, de datos de autoridad, de fondos y localizaciones. Y, por otro, el desarrollo de un *software* que permita la conversión entre los dos formatos.

El proyecto comenzó en 1995 con el desarrollo del DTD y la versión *alpha* estuvo ya disponible en mayo de 1996. La versión definitiva del software de conversión estuvo lista en enero de 1998 y desde entonces la Biblioteca del Congreso proporciona a través de sus páginas www la posibilidad de descargarlo de forma gratuita.

7 Conclusiones

Las estructuras de metadatos están adquiriendo un lugar central en la descripción de documentos electrónicos, de cualquier tipo y naturaleza, como medio de dotarlos de formas eficaces de recuperación.

El estándar que más se utiliza para la estructuración de estos datos es SGML que proporciona acceso a información jerárquica, bibliográfica y analítica compleja.

El formato MARC se ha mantenido fundamentalmente como una versión electrónica del catálogo en fichas con las limitaciones propias de este modelo. La función principal de este formato, cuando se diseñó por primera vez en los años 60, fue permitir la distribución electrónica de registros bibliográficos para la producción de fichas en papel. La ficha catalográfica se creó para objetos de información estáticos con una estructura plana de los datos. El catálogo en fichas apunta al objeto físico. Sin embargo, la cabecera del TEI, por ejemplo, proporciona, por una parte, información bibliográfica completa. Y, por otra, información no bibliográfica que permite el análisis, por medios humanos e informáticos, del texto electrónico representado en la cabecera. La mayoría de esta información analítica que se encuentra en la cabecera, en la descripción codificada y en la revisión histórica, tiene muy poco que ver con el registro MARC. Y, si está explícita en éste es en forma de notas en los campos 5XX. Notas que normalmente no son campos de búsqueda. Al formato MARC le falta, en definitiva, habilidad suficiente para tratar información estructurada jerárquicamente y proporcionar acceso a colecciones complejas descendiendo por los niveles de análisis. Por tanto, hay que reexaminar los formatos MARC y analizar cómo pueden evolucionar o transformarse en el futuro.

Hay que resaltar, además, que tanto la cabecera del TEI como el Dublin Core ponen más énfasis en el acceso que en la descripción. Esto es importante, ya que ha sido un fallo tradicional en los catálogos bibliotecarios en los que, por el contrario, se hace más hincapié en la descripción que en dotar a los registros de más y mejores elementos de recuperación.

SGML se puede utilizar para crear una gran variedad de documentos y además usarlo para proporcionar acceso y control a multitud de formatos de información en línea. No es un lenguaje de marcas propietario como es el formato MARC y no depende tampoco de un *software* propietario o de una oferta limitada de *software*. En esencia podría ser posible usar SGML como el estándar general para las herramientas informativas que se usan para catalogar (reglas, formato MARC, CDU, etc.), para crear los registros catalográficos y para crear los textos electrónicos de nuestro catálogo. Y, además, podríamos utilizar el texto para proporcionar acceso y control a los objetos digitales que no son texto.

Claramente hay ahora una oportunidad de crear y construir un entorno de información integrada en el que el catálogo proporcione acceso tanto a los documentos tradicionales como a información electrónica. Este entorno civilizado emergiría de los desarrollos presentes en Internet y del tipo de proyectos que se están llevando a cabo.

Un OPAC podría servir de pasarela para acceder a una base de datos de metadatos en Internet, así como a las bases de datos de la propia biblioteca. A través de las capacidades de búsqueda del Z39.50 y las conexiones TCP/IP se podría, con una única interfaz, proporcionar acceso a los usuarios a bases de datos catalográficas, bases de datos en CD-ROM, bases de datos en texto completo, servicios en línea y otras bases de datos compatibles con el Z39.50. Los navegadores Web se convertirían así en clientes de los servidores Z39.50. Como resultado, las bases de datos de los OPACs estarían integradas con una gran variedad de recursos internos de Internet, así como recursos externos (el Web de la Universidad de Yale es un ejemplo de esto). Si así fuera se tendrían que establecer perfiles Z39.50 para cada uno de los formatos de metadatos. Por el momento, sólo cuentan con él el MARC y el Dublin Core.

Algunos autores (13) mantienen que el formato MARC sigue siendo el más apropiado por su estructura y, sobre todo, porque su uso está muy extendido. Proponen crear programas de conversión entre los distintos formatos de metadatos con el objetivo de poder incluir los registros en los catálogos bibliotecarios. Yo no considero que el MARC sea la mejor forma de describir metadatos o que sea interesante convertir los metadatos en otros formatos a MARC. Pienso que podemos encontrarnos en la misma situación que cuando se instalaba un sistema integrado de gestión bibliotecaria y se seguía manteniendo el catálogo manual. Llegará un momento en que sólo uno de los dos sistemas prevalezca.

Ahora bien, en lo que sí coinciden la mayoría de los autores es en señalar que tanto las reglas de catalogación como los formatos MARC deben evolucionar para adaptarse al nuevo entorno.

El papel que jueguen los bibliotecarios en proporcionar descripciones y acceso a los recursos en las redes tendrá cada vez más relación con la creación, mantenimiento y mejora de descripciones de metadatos. En definitiva, buscar y estudiar metadatos y estándares para el futuro digital.

Los catalogadores deben adoptar nuevas perspectivas, nuevos métodos de organizar y recuperar la información. Deben, desde mi punto de vista, revisar los conceptos que van unidos a la forma de entender las relaciones que se establecen entre acceso, descripción y recuperación

El catálogo futuro debe ser flexible y multitarea y el MARC no puede proporcionarnos esa flexibilidad. Esto no quiere decir que el formato MARC ya no sirva. Hay billones y billones de registros MARC, el coste y el tiempo de convertirlos a SGML sería abrumador. Pero yo creo que el formato MARC no será por mucho más tiempo el único formato de codificación de datos bibliográficos en los sistemas bibliotecarios. El formato MARC no es suficientemente flexible para permitir a las bibliotecas aprovechar la tecnología de nuevo desarrollo de recuperación y acceso a la información.

Los catalogadores son organizadores expertos y tienen un papel muy importante que representar proporcionando, en la medida de lo posible y de acuerdo con la comunidad a la que van dirigidos, control bibliográfico a los recursos informativos en las redes, y deben utilizar este conocimiento para buscar nuevas soluciones en un medio nuevo.

Es ahora el momento de trasladar el éxito obtenido en el mundo impreso a la biblioteca digital del futuro. En este sentido, como ha demostrado el transcurrir de los años, las bibliotecas deberán enfrentarse a nuevas herramientas que hagan posible llevar a cabo con éxito la labor.

8 Anexo: Ejemplos de diferentes estructuras de metadatos¹

Dublin Core

Subject:
scheme=keywords:Electronically mediated environments

¹ Ejemplos tomados de: <http://www.oclc.org/oclc/man/colloq/xu.htm#conc>.

Cyberspace
Urbanism
Architecture

scheme=LCSH:Computer networks
Information technology
Virtual reality
Computers and civilization

Title: City of Bits: Space, Place, and the Infobahn

Author: Mitchell, William J.

Publisher: MIT Press

OtherAgents:

otherAgent role=WWW team member: Stevenson, Daniel C.
otherAgent role=WWW team member: Ehling, Teresa
otherAgent role=WWW team member: Kalin, Jeffrey T.
otherAgent role=WWW team member: Schoonover, Regina
otherAgent role=WWW team member: Beamish, Anne
otherAgent role=WWW team member: Ishizake, Suguru
otherAgent role=WWW team member: Urbanowski, Frank

Date: 1995

Identifiers:

scheme=ISBN: 0262133091

scheme=URL:

http://www-mitpress.mit.edu:80/City_of_Bits/WWW_Preamble.html

Object type: book

Form: Text/HTML, Video/(MPEG, Quicktime), Image/GIF

Language: English

Source: type=print ed.: City of Bits: Space, Place, and Infobahn

RDF con los elementos del Dublin Core

```
<rdf:RDF
  xmlns:rdf=>http://w3.org/TR/1999/PR-rdf-syntax-19990105#</>
  xmlns:dc=>http://purl.org/metadata/dublin\_core#</>
  xmlns:dcq=>http://purl.org/metadata/dublin\_core\_qualifiers#<>
  <rdf:Description about=>http://www.dlib.org/dlib/may98/05contents.html</>
    <dc:Title>D-LIB Magazine - The Magazine for Digital Library Research
      - May 1998</dc:Title>
    <dc:Description>D-LIB magazine is a monthly compilation of
      contributed stories, commentary, and briefings.</dc:Description>
    <dc:Contributor rdf:parseType=>Resource<>
      <dcq:AgentType
        rdf:resource=>http://purl.org/metadata/dublin\_core\_qualifiers#Editor</>
        <rdf:value>Amy Friedlander</rdf:value>
      </dc:Contributor>
    <dc:Publisher>Corporation for National Research Initiatives</dc:Publisher>
    <dc:Date>1998-01-05</dc:Date>
    <dc>Type>electronic journal</dc>Type>
```

```
<dc:Subject>
  <rdf:Bag>
    <rdf:li>library use studies</rdf:li>
    <rdf:li>magazines and newspapers</rdf:li>
  </rdf:Bag>
</dc:Subject>
<dc:Format>text/html</dc:Format>
<dc:Identifier>urn:issn:1082-9873</dc:Identifier>
<dc:Relation rdf:parseType=>Resource>>
  <dcq:RelationType
    rdf:resource=>http://purl.org/metadata/dublin\_core\_qualifiers#IsPartOf>/>
    <rdf:value resource=>http://www.dlib.org>/>
  </dc:Relation>
  </rdf:Description>
</rdf:RDF>
```

Cabecera TEI

```
<teiheader>
  <filedesc>
    <titlestmt>
      <title>City of bits: Space, Place, and the Infobahn</title>
      <author>William J. Mitchell</author>
    </titlestmt>
    <editionstmt>World-Wide Web ed.</editionstmt>
    <publicationstmt>
      <publisher>MIT Press</publisher>
      <pubplace>Cambridge, Mass.</pubplace>
      <idno type=oclc>32437789</idno>
      <date>1995</date>
    </publicationstmt>
    <notesstmt>
      <note>856 7 $u
URL:http://mitpress.mit.edu/City\_of\_Bits/WWW\_Preamble.html</note>
    </notesstmt>
    <sourcedesc>
      <bibfull>
        <titlestmt>
          <author>Mitchell, William J.</author>
          <title>City of bits</title>
          <title type=sub>space, place, and the infobahn</title>
        </titlestmt>
        <extent>225 p. : ill., maps, plans ; 24 cm.</extent>
        <imprintstmt>
          <pubplace>Cambridge, Mass.</pubplace>
          <publisher>MIT Press</publisher>
          <idno type=isbn>0262133091</idno>
          <idno type=oclc>33278259</idno>
          <date>1995</date>
        </imprintstmt>
      </bibfull>
    </sourcedesc>
  </filedesc>
</teiheader>
```

```
</bibfull>
</sourcedesc>
</filedesc>
<encodingdesc>NA</encodingdesc>
<profiledesc>
  <textclass>
    <keywords scheme=lcsh>
      <list>
        <item>Computer networks</item>
        <item>Information technology</item>
        <item>Virtual reality</item>
        <item>Computers and civilization</item>
      </list>
    </keywords>
    <classcode scheme=lc>TK5105.5.M57</classcode scheme>
  </textclass>
</profiledesc>
<revisiondesc>NA</revisiondesc>
```

Registro URC

```
URC.0 {
// This is a hypothetical record
URN: Universal Resource Name (not yet available)
Title: City of Bits: Space, Place, and the Infobahn
Author {
  Name: Mitchell, William J.
  Email:wjm@mit.edu
  Phone:617-253-4402
  Facsimile:617-253-9417
}
Subject: electronically mediated environments; cyberspace; urbanism; architecture

Abstract {
  Textual: A hyperlinked exploration of the «virtual
  city» which is now emerging through our burgeoning use of
  the information superhighway.
}
Location {
  URL:
  http://mitpress.mit.edu/City_of_Bits/WWW_Preamble.html
  Content type: text/html, image/GIF, video/MPEG,
  Quicktime
  Content-length:
  Signature (Not yet available)
}

Review {
  //There is no URN available for the media reviews.
```

```
//This URL will lead to the content of them.  
URL: http://www-mitpress.mit.edu/City_of_Bits/  
      reviews.html  
}  
Version: World-Wide Web ed.  
}
```

Registro USMARC

000 cmm Ia
001 32437789
003 OCoLC
005 19000000003748.0
008 950508s1995 maun d eng d
040 MYG \$c MYG
090 TK5105.5 \$b .M57 1995b
100 1 Mitchell, William J.
245 10 City of bits \$h [interactive multimedia] : \$b space, place, and the infobahn / \$c by William J. Mitchell.
250 World-Wide Web ed.
256 Computer data.
260 [Cambridge, Mass.] : \$b MIT Press, \$c 1995.
516 Text (HTML), images (GIF), and video (MPEG, QuickTime).
538 System requirements: Web browser; video viewer such as QuickTime or MPEGPlay required for video applications.
538 Mode of access: Internet. Address:
<http://mitpress.mit.edu/CityofBits/>.
500 Title from title screen.
530 Also available in printed ed.
520 A hyperlinked exploration of the «virtual city» which is now emerging through our burgeoning use of the information superhighway. Re-examines architecture and urbanism in light of our increasingly digital means of communication.
505 0 1. Pulling glass - 2. Electronic agoras - 3. Cyborg citizens - 4. Recombinant architecture - 5. Soft cities - 6. Bit biz - 7. Getting to the good bits.
504 Includes bibliographical references.
650 0 Computer networks.
650 0 Information technology.
650 0 Virtual reality.
650 0 Computers and civilization.
856 7 \$2 http \$z http://mitpress.mit.edu/City_of_Bits/
\$u <http://purl.oclc.org/OCLC/OLUC/32437789/1>

9 Referencias

1. SMITH, T. R. The Meta-Information Environment of Digital Libraries. *D-Lib Magazine* 1996, julio-agosto. Disponible en: <http://www.dlib.org/july96/new/07smith.html>.

2. CAPLAN, P. You Call It Corn, We Call It Syntax-Independent Metadata For Document-Like Objects. *The Public Access Computer Systems Review*, 1995, vol. 6, n.^o 4.
3. BERNERS-LEE, T. *Metadata Architecture*. <http://www.w3.org/DesignIssues/Metada-ta.html>.
4. DEMPSEY, L.; HEERY, R. A review of metadata: a survey of current resource description formats. *Journal of Documentation*, 1998, vol. 54, n.^o 2, pp. 145-172.
5. BEARMAN; SOCHATS. *Metadata requirements for evidence*. <http://www.oclc.org/oclc/research/projects/core/workshops/dc1conference/resources-bearman.htm>.
6. LARSON, R. *A few words about the USMARC.DTD*. <ftp://library.berkeley.edu/pub/sgml/marcDTD/README>.
7. GORMAN, M. After AACR2R: The Future of the Anglo-American Cataloging Rules. En *Origins, Content and Future of AACR2 Revised*, 1992, Chicago: American Library Association, pp. 89-94.
8. Para una descripción más detallada véase: http://purl.oclc.org/metadata/dublin_core.
9. *Resource Description Framework (RDF) Model and Syntax Specification: W3C Proposed Recommendation 05 January 1999*. <http://www.w3.org/TR/PR-rdf-syntax>.
10. BURNARD, L. *The Text Encoding Initiative Guidelines*. <ftp://infor.ox.ac.uk/pub/ota/TEI/doc/teij31.sgml>.
11. La documentación completa sobre el URC está disponible en: <http://www.acl.lanl.gov/URI>.
12. LIBRARY OF CONGRESS. NETWORK DEVELOPMENT AND MARC STANDARDS OFFICE. *MARC DTDs Document Type Definitions: Background and Development*. <http://lcweb.loc.gov/marc/marcDTback.htm>.
13. XU, A. Metadata conversion and Library Opac. *Serial Librarian*, 1997, vol. 33, n.^o 1-2, pp. 179-192.