

**ESCUELA DE BIBLIOTECOLOGÍA
FACULTAD DE FILOSOFÍA Y HUMANIDADES
UNIVERSIDAD NACIONAL DE CÓRDOBA**



**"Análisis y comparación de metadatos para la descripción
de recursos electrónicos en línea"**

Trabajo final para optar al título de Lic. en Bibliotecología y Documentación

Alumnas:

ANGELOZZI, Silvina Marcela

MARTIN, Sandra Gisela

**Directora:
Lic. ELSA E. BARBER
Universidad de Buenos Aires
Facultad de Filosofía y Letras
Dpto. de Bibliotecología y Documentación**

Córdoba, febrero 2009

RESUMEN

Este trabajo consiste en un estudio comparativo de metadatos para la descripción de documentos accesibles por Internet. Se analizan las particularidades y dificultades propias de los recursos electrónicos en línea en cuanto a su catalogación. Se definen y caracterizan los metadatos, se describen y analizan los diferentes esquemas. La comparación se realiza teniendo en cuenta características como origen y propósito, estructura y contenido de la descripción, complejidad, sintaxis, contribución a la normalización internacional, interoperabilidad, extensibilidad, flexibilidad, mantenimiento, documentación existente, actualización y resultados obtenidos hasta el presente.

PALABRAS CLAVE: CATALOGACIÓN; INTERNET; METADATOS; RECURSOS ELECTRÓNICOS ; RECUPERACIÓN DE INFORMACIÓN

KEYWORDS: CATALOGING; INTERNET; METADATA; ELECTRONIC RESOURCES; INFORMATION RETRIEVAL.

ÍNDICE

RESUMEN	2
ÍNDICE	3
INTRODUCCIÓN	4
OBJETIVOS	5
MARCO TEÓRICO	5
METODOLOGÍA	8
Tipología	11
CATALOGACIÓN DE LOS RECURSOS ELECTRÓNICOS	12
METADATOS	15
Funciones	16
Tipología	18
NORMATIVAS INTERNACIONALES PARA LA DESCRIPCIÓN DE RECURSOS ELECTRÓNICOS	19
Estándares de contenido	20
a) AACR2R	20
b) ISBD	21
ANÁLISIS DE LOS ESQUEMAS DE METADATOS	24
MARC21/MARCXML	25
DUBLIN CORE	33
MODS (Metadata Object Description Schema)	46
TEI (Text Encoding Initiative)	67
LENGUAJES DE MARCADO	77
SGML	77
HTML	77
XML	78
RDF	79
COMPARACIÓN DE LOS ESQUEMAS DE METADATOS	81
METADATOS Y RECUPERACIÓN DE LA INFORMACIÓN	86
1- Buscadores o motores de búsqueda:	86
2.- Metabuscaros (<i>metasearch engines</i>)	88
3.- Directorios (<i>Web directories</i>)	89
CONCLUSIONES	90
BIBLIOGRAFÍA CONSULTADA	92

INTRODUCCIÓN

La sociedad de la información y la globalización conllevan a un cambio de paradigma en el campo de la organización del conocimiento. La información electrónica presenta una arquitectura hipertextual, no lineal y distribuida que plantea nuevos desafíos al profesional bibliotecario.

Internet ha permitido el acceso remoto a los catálogos y a los instrumentos de descripción, y más aún los propios fondos documentales de las bibliotecas están siendo digitalizados y puestos a disposición de los usuarios a través de este poderoso medio de comunicación. Una eclosión de diversos tipos de documentos va poblando Internet a un ritmo de crecimiento exponencial. Una consecuencia de ello es que cualquier catálogo de biblioteca puede abastecerse de la gran cantidad de documentos que se distribuyen por Internet, superando los fondos locales y abriéndose al inmenso espacio digital. Esto está claro en la teoría, pero aún muchas bibliotecas no catalogan fondos de Internet, aunque sí ofrezcan un espacio de consulta y un servicio de referencia virtual.

Como lo plantean Pinto Molina, García Marco, y Agustín Lacruz (2002), si bien no es una experiencia nueva el hecho de que una parte muy importante de los documentos de interés para los usuarios estén fuera de la propia biblioteca, sí es una situación innovadora que comiencen a ser extraordinariamente accesibles. Más adelante, estos autores, que señalan la facilidad con la que se pueden digitalizar los documentos textuales y audiovisuales, la transmisión instantánea a distancia, su simultaneidad de uso, su utilización sin deterioro y la reducción del espacio de almacenamiento, ponen de manifiesto que con los documentos digitales han desaparecido las limitaciones espaciales y temporales intrínsecas al resto de los documentos, lo que dispara de un modo espectacular sus posibilidades informativas. Por ello, la integración de registros bibliográficos de materiales remotos se ha convertido en una creciente necesidad para la creación y el mantenimiento de los catálogos de las bibliotecas.

En este estudio se aborda el tema de los metadatos y las normativas internacionales para la descripción de recursos electrónicos de información disponibles en línea, analizando los problemas que comportan a nivel conceptual y de desarrollo de nuevas formas de estructurar la información. Desde una perspectiva bibliotecológica, se da preponderancia a la normalización y a las posibilidades de aplicación de los desarrollos estudiados para organizar la información accesible vía Web.

La información digital difiere de la documentación impresa en muchos aspectos, por lo que se ha generado un debate en el ámbito de los profesionales bibliotecarios con respecto a los métodos tradicionales de procesamiento de la información. En este trabajo se trata de responder preguntas tales como: ¿Qué son realmente los metadatos? ¿Es un concepto realmente nuevo?

¿Cuál es la diferencia con la catalogación tradicional? ¿Para qué sirven? ¿Qué normativas han surgido? ¿Qué diferencias y similitudes hay entre ellas?

OBJETIVOS

- analizar similitudes y diferencias entre la catalogación de materiales tradicionales y de recursos de información en Internet;
- analizar diferentes esquemas de metadatos para la descripción de recursos electrónicos accesibles por Internet;
- comparar los esquemas analizados en base a los criterios establecidos;
- elaborar conclusiones en cuanto a los resultados de la comparación y en relación a la recuperación de información en la Web.

MARCO TEÓRICO

Los metadatos tienen un rol importante en el contexto del surgimiento de las bibliotecas digitales ya que, la organización de la información disponible en las redes depende de la administración efectiva de aquellos.

En este trabajo se entiende por metadatos, siguiendo a Taylor (2004), a la descripción codificada de un paquete de información cuyo propósito es proporcionar un nivel intermedio a través del cual se puede acceder a la información por medio de lo que uno desea ver o buscar, sin tener que investigar enormes cantidades de texto completo irrelevante.

Burnett, Bor Ng y Park (1999) señalan que los metadatos han sido abordados desde dos ópticas:

- **la del control bibliográfico**, con origen y proponentes desde la bibliotecología y,
- **la del procesamiento de datos**, con origen y proponentes desde las ciencias de la computación.

A pesar de las diferencias, estos dos abordajes usan metadatos para localizar, identificar, recuperar y administrar la información; y a medida que avanza el desarrollo, estas dos corrientes se van acercando e integrando.

Eito Brum (2008) manifiesta que la proliferación de los documentos publicados en la Web hizo que la comunidad bibliotecaria aplicara técnicas similares a la catalogación tradicional para la identificación, localización y posterior recuperación de estos nuevos recursos.

Las bibliotecas tienen una larga tradición en el desarrollo de sistemas de información, y se concentran en organizar y proveer acceso a los documentos. Para ello utilizan la descripción bibliográfica, el análisis de contenido y la clasificación, de modo que cada entidad de información pueda ser eficientemente descrita, identificada, localizada y recuperada. Mediante

el establecimiento de una normativa específica, han logrado implementar con éxito complejos sistemas documentales. Así, la unidad básica del catálogo, el registro bibliográfico, representa unívocamente y describe de forma precisa un documento.

La necesidad de intercambiar información entre instituciones a nivel local, nacional e internacional ha hecho que se desarrollen normas y procedimientos que aseguren la interoperabilidad de los sistemas de información. Aunque no se resuelvan todos los requerimientos de la organización de la información digital, las reglas que se vienen usando sirven de base y tienen un impacto significativo.

Las ciencias de la computación también tienen un largo camino recorrido en cuanto al procesamiento de datos y a la organización de la información. Hay numerosos sistemas computarizados de almacenamiento y recuperación de información, tanto para datos textuales como relacionales, que proveen seguridad, integridad y posibilidad de intercambio de datos, además de las funciones tradicionales de localización, identificación y recuperación. Todo el avance logrado desde las ciencias de la computación también tiene una fuerte influencia en el desarrollo de metadatos.

Desde la perspectiva del control bibliográfico, se desarrollaron metadatos orientados a la catalogación de los recursos de información en Internet, como la utilización del formato MARC (*MACHine Readable Cataloguing*) como estructura de metadatos y la formulación de Dublin Core. Esta corriente se focaliza en la descripción e identificación de los objetos de información.

Desde la perspectiva de la computación, se dio preponderancia a la manipulación de los datos para su uso, más allá de su mera descripción, y se incluyeron datos de acceso, de autenticación, administrativos, etc. Se han desarrollado lenguajes de marcación de textos como HTML (*Hyper Text Mark-up Language*), XML (*eXtensible Mark-up Language*), SGML (*Standard Generalized Mark-up Language*), Warwick Framework, y formatos como IAFA (*Internet Anonymous FTP Archives*) y TEI (*Text Encoding Initiative*).

A pesar de que estas dos aproximaciones a los metadatos son diferentes, de ningún modo son excluyentes o incompatibles, más bien son complementarias, y la solución definitiva provendrá de un estudio interdisciplinario, que integre técnicas de una y otra.

Como señala Taylor (2004), entre los profesionales de la información el concepto de metadatos puede aparecer complejo y confuso. Esto se debe, en parte, a la naturaleza multifacética del tema. Es importante recordar que el término metadatos puede significar diferentes cosas para diferentes comunidades. Cuando un bibliotecario habla de metadatos, tiene una noción diferente de la que posee quien trabaja, por ejemplo, con metadatos geoespaciales.

Dovey (1999) identifica tres corrientes principales a las que denomina escuelas: "*la escuela de la catalogación*", "*la escuela estructuralista*" y "*la escuela de la estructura de datos*". La primera, incluye a los catalogadores que proceden del mundo bibliotecario y que han

estado generando metadatos prácticamente desde que existen las bibliotecas; también aquí ubica a los informáticos preocupados por generar estructuras que hagan la información más manejable, y dedicados a construir metadatos para catalogar. En esta línea se ubican los creadores de Dublin Core y RDF. La segunda, la escuela estructuralista, está constituida por los creadores de los lenguajes de marcado, cuyos estándares y modelos están basados en la estructura del texto, como por ejemplo HTML, SGML. La tercera de estas escuelas, llamada de estructura de datos, pone de relieve el problema de la separación del contenido de la presentación en los documentos electrónicos; esta corriente ve al XML como un lenguaje universal para describir estructuras de datos, adecuado para el intercambio tanto de la estructura de las bases de datos como del contenido.

Tal como lo señalan Pinto Molina, García Marco y Agustín Lacruz (2002), si algo caracteriza el momento actual es la superación de la ecuación tradicional "documento = texto verbal escrito". La aparición de diferentes soportes a lo largo del tiempo ha llevado a utilizar ciertos nuevos términos muy imprecisos como "materiales especiales", "nuevos documentos", "documentos no textuales". El problema se complica aún más con los documentos multimedia y virtuales. Una de las características de éstos es que integran los tipos anteriores, por ejemplo un multimedia interactivo contiene textos escritos, sonidos, imágenes y videos.

En la actualidad, cualquier documento se puede convertir a formato digital, así cuando se habla de documento digital o electrónico, se hace referencia al formato de codificación utilizado para su almacenamiento y procesamiento más que a la forma de presentación de la información contenida en él. Cuando se utiliza el término documento digital sólo para aquéllos distribuidos por Internet, se confunde formato y medio, ya que Internet es un medio de comunicación y dichos documentos preceden a Internet.

La utilización del término "recurso de información" es de procedencia anglosajona donde se ha popularizado el término "*information resources*". En este trabajo se lo utiliza, si bien puede objetársele el ser poco específico, porque supera, en cierto modo, la dificultad terminológica aún no resuelta en cuanto a la tipología documental. Es importante destacar que las ISBD(CF) (*International Standard Bibliographic Description: Computer Files*) se revisan en 1997, y se cambia la denominación de "archivos de ordenador" por "recursos electrónicos", dando origen a la ISBD(ER) (*International Standard Bibliographic Description: Electronic Resources*).

Siguiendo esta línea, otro concepto que se adopta en este trabajo es el de "objeto de información", DLO (*Document Like Object*) según su sigla en inglés. Tal como lo plantea Méndez Rodríguez (2002), una de las claves principales de la teoría de metadatos está en definir qué es un objeto de datos o qué es un documento entendido como objeto. Así, en un ambiente digital, se puede definir un DLO como un elemento o un grupo de elementos que constituyen una unidad informativa que una computadora puede manejar, independientemente del formato

del archivo y del tipo de información que reflejen. Entendido así, en las páginas subsiguientes se utilizan términos como recurso electrónico, documento electrónico, objeto digital u objeto de información, para referirse a esta entidad.

METODOLOGÍA

El presente trabajo responde a un diseño bibliográfico o investigación documental. En este sentido, Tamayo y Tamayo (2000) define la investigación documental como: “la que se realiza con base a la revisión de documentos, manuales, revistas, periódicos, actas científicas, conclusiones de simposios y seminarios y/o cualquier tipo de publicación considerado como fuente de información.” El presente estudio se limita a aquellos proyectos cuyos responsables son instituciones de reconocido prestigio en el campo de la bibliotecología, tales como la *Library of Congress (LC)* u *Online Computer Library Center (OCLC)*, a emprendimientos cooperativos internacionales y de instituciones pioneras en redes digitales. Así, los esquemas de metadatos seleccionados para la comparación son: MARC21/MARC XML, Dublin Core, MODS (*Metadata Object Description Schema*) y TEI (*Text Encoding Initiative*).

Se realizó una búsqueda sobre el tema en libros y publicaciones periódicas, y especialmente en Internet, ya que la mayor parte de la información relevante y actualizada sobre la Web está en la misma Web. Se analizó bibliografía publicada a partir de 1994, año en el que el tema de metadatos cobra fuerza en el ámbito bibliotecológico y llega hasta la fecha de cierre de este trabajo, en diciembre de 2009. Para la descripción de los elementos y características de los esquemas de metadatos se utilizaron las fuentes primarias de información, es decir los sitios Web oficiales de cada iniciativa.

La grilla de comparación se construyó sobre la base de los criterios propuestos por Heery (1996), Dempsey y Heery (1998) y Taylor (2004) y quedó conformada de la siguiente manera:

- **Disciplina/ámbito de aplicación:** considera quiénes están usando activamente este formato y si éste está asociado con una disciplina académica o interés profesional en particular.
- **Facilidad de creación de la descripción:** se refiere a la complejidad de la creación del registro, si se requieren habilidades especiales, si están diseñados para que asigne los metadatos el autor, el editor o los proveedores centralizados.
- **Contenido:** incluye estructura, sintaxis y contenido.
- **Protocolos de Internet asociados:** hace referencia a si los metadatos son soportados por los protocolos de Internet existentes y si una base de datos con dichos metadatos puede ser interrogada usando los protocolos existentes.
- **Contribución a la normalización internacional:** considera la ratificación del esquema de metadatos como norma internacional.

- **Interoperabilidad:** se refiere a la compatibilidad de dos o más sistemas de modo que puedan intercambiar datos e información y utilizar la información sin ninguna manipulación especial.
- **Extensibilidad:** tiene en cuenta la posibilidad de incluir mayor nivel de detalle, si fuera necesario, mediante campos opcionales, calificadores, etc.
- **Flexibilidad:** se puede aplicar a dos aspectos: a la flexibilidad para describir diferentes tipos de documentos y a la flexibilidad para evolucionar de acuerdo con los cambios en las reglas de catalogación y tecnología.

A través de este cotejo se pretende establecer cuáles son los metadatos que se perfilan con mayor posibilidad de éxito y difusión en el ámbito de la catalogación de recursos en red.

RECURSOS ELECTRÓNICOS

Lyons (1997) define un objeto digital o recurso electrónico como un conjunto de bits arreglados en una secuencia, que contiene datos que pueden ser interpretados, y que incluye un identificador único.

Según las Reglas de Catalogación Angloamericanas (segunda edición, actualización 2003, AACR2R), un recurso electrónico es un documento codificado para ser manipulado con la computadora. Se los divide, por un lado, en **datos** (información en la forma de números, letras, gráficos, imágenes y sonido, o combinación de cualesquiera de ellas) y **programas** (instrucciones o rutinas para ejecutar ciertas tareas); y por otro, se distinguen entre los de **acceso local** (con soporte físico, disco, casete, disquete) y **remoto** (disponibles a través de redes).

Según las normas de la Descripción Bibliográfica Internacional Normalizada (ISBD), (edición consolidada, 2007), los recursos electrónicos son materiales controlados por computadora, incluidos los que requieran el uso de periféricos (por ejemplo un lector de CD-ROM), y dichos recursos pueden ser o no utilizados en un modo interactivo. Los divide en: **datos** (información en la forma de números, letras, gráficos, imágenes y sonido, o combinación de cualesquiera de ellas) y **programas** (instrucciones o rutinas para ejecutar ciertas tareas), pueden ser también una **combinación de datos y programas** (por ejemplo, un software educativo que contenga texto, gráficos y programas).

Esta definición se aplica a los recursos que están disponibles e incluye también aquellos de acceso remoto via telecomunicaciones. Sin embargo, los juguetes programados, las calculadoras y otros objetos programados se consideran fuera del alcance de la ISBD.

Una característica interesante de destacar es la cantidad de actores que intervienen en la producción y conservación de los recursos electrónicos (Lazinger, 2001):

1. **Creadores:** individuos o instituciones generadores de la información contenida en ellos.
2. **Editores:** en el caso de recursos publicados que poseen los derechos y licencias.
3. **Distribuidores:** proveedores del acceso que no crean ni publican la información, sino que sólo la ponen a disposición; no se incluyen aquí las bibliotecas.
4. **Administradores de sistemas:** responsables del manejo informático de los archivos, de los programas de recuperación de la información, de los equipos de computación e interconexión de redes, servicios de respaldo y recuperación, migración de datos, establecimiento de códigos de seguridad.
5. **Bibliotecas:** aisladas o en consorcios, encargadas de poner los recursos de información a disposición de su comunidad de usuarios, se mantienen archivos digitales *in situ* o se conectan a proveedores comerciales.

6. **Archivos digitales:** mantienen copias de los recursos electrónicos, una a la que se permite el acceso y otra copia de resguardo, distinguiendo entre ambos roles: uso y preservación.

Tipología

Según Ede (1999) existen dos tipos de recursos electrónicos:

- **en mano:** soportes físicos como disquetes, CD-ROM o cintas magnéticas;
- **en línea:** cualquier objeto que se puede encontrar en la Web: bases de datos, revistas electrónicas, páginas, listas de correo, grupos de noticia, sitios ftp, etc.

En cambio, Hedstrom (1998), los divide en:

- **creados originalmente en forma digital** y adquiridos a través de compra, donación o depósito legal y,
- **digitalizados** a partir de documentos impresos, o manuscritos convertidos a formato electrónico.

Una diferencia entre ambas categorías es que los primeros son más amorfos, menos controlados bibliográficamente.

Por su parte, Waters (1995), establece cinco categorías:

1. Origen y modo de distribución:

- de propiedad personal (e-mail, notas, manuscritos, pre-impresos, etc.),
- de propiedad corporativa (reportes de empresas, registros financieros, documentos de planificación, etc.),
- emanados de editoriales (libros, revistas, grabaciones, filmes),
- pertenecientes a bibliotecas, museos e instituciones educacionales.

2. Contexto:

- fuentes secundarias (catálogos, índices, directorios, etc.),
- fuentes primarias (textos monográficos, imágenes, grabaciones, colecciones de datos).

3. Estructura de codificación, formato y contenido:

- codificados según formatos normalizados (ASCII; TIFF, etc.),
- con características de autodescripción (SGML con DTD),

- multimedia que incorporen varios formatos.

4. Cualidades dinámicas:

- pueden ser revisados y actualizados, de modo que existen versiones o ediciones,
- con cambios acumulativos o de manera interactiva, como por ejemplo las contribuciones a listas de discusión,
- variación de los enlaces entre un objeto y otros.

5. Uso:

- Objetos que por falta de consulta son borrados, sin tener en consideración una posible demanda futura.

A su vez, Law (1996) los agrupa en cuatro categorías:

- 1- **Publicados comercialmente** (libros y revistas electrónicos, bases de datos, etc.), este material es en general más fácil de describir.
- 2- **Documentos de trabajo** (literatura gris, pre-impresos, páginas Web personales, etc.), material difícil de controlar, no está asegurada su permanencia o acceso.
- 3- **Datos en bruto** (estadísticas, producidos por los gobiernos y organizaciones no gubernamentales, tomados de satélites), muchas veces es información encriptada y puesta a disposición sólo de un grupo selecto de usuarios bajo ciertos términos.
- 4- **Archivos científicos** (trabajos y correspondencia entre investigadores), considerados muy importantes para el avance de una disciplina.

CATALOGACIÓN DE LOS RECURSOS ELECTRÓNICOS

A continuación se presentan las características especiales de los recursos electrónicos, las cuales influyen a la hora de catalogarlos.

Wendler (2000) presenta las siguientes características:

- **Variabilidad en la presentación:** la apariencia y el orden del contenido no están normalizados, y la presentación depende muchas veces del navegador.
- **Formas no familiares de publicación:** se asemejan a las publicaciones de hojas sueltas removibles, la nueva versión sustituye a la anterior y frecuentemente sin una numeración o indicación de que la versión ha cambiado.
- **Relaciones mal definidas de los materiales con otros soportes:** un subconjunto de las publicaciones electrónicas se deriva o está vinculado con documentos en otros

soportes. No siempre es posible identificar el grado de coincidencia en cuanto al contenido o no existe consenso con respecto a cómo expresar estas relaciones.

- **Mutabilidad:** nada garantiza que un recurso en línea permanezca inalterado en el tiempo ni que su contenido no sea alterado sin previo aviso.
- **Movilidad:** no se puede asegurar la permanencia de un documento en el mismo sitio Web.
- **Ausencia de un receptor físico:** la falta de un usuario convencional elimina el punto clave sobre el cual la mayoría de las bibliotecas ha construido los procedimientos de mantenimiento del catálogo.

A todo lo anterior se puede agregar:

- **Gran variedad de recursos:** en Internet hay gran cantidad de textos, imágenes, sonidos, presentados en las más variadas formas como ser archivos, informes, listas de discusión, artículos de revistas, directorios, libros electrónicos, páginas Web, portales.

Heery (1996) presenta además dificultades relacionadas con las siguientes características:

- **Ubicación:** un registro de un catálogo en línea de acceso público (OPAC) de una biblioteca se refiere a ubicaciones dentro de dicha institución. Los metadatos se refieren a ubicaciones remotas, por lo tanto se requerirán datos asociados al tipo de acceso (ftp, http) y a las restricciones de acceso (clave requerida). Muchas veces un recurso puede estar alojado en varios servidores (sitios espejo), en este sentido, los metadatos se asemejan a un registro de catálogo colectivo.
- **Versiones diferentes:** el mismo documento puede existir en diferentes formatos a la vez (p. ej.: Postscript, ASCII). Podría asemejarse a las distintas ediciones de un impreso.
- **"Granularity":** o nivel de detalle, otra particularidad de los recursos electrónicos en línea es que bajo un mismo título se pueden agrupar múltiples recursos. Ejemplo de esto son las suscripciones a texto completo de un paquete de revistas, que plantea la alternativa de catalogar individualmente cada título o simplemente el conjunto de ellos. Esto es simple cuando la cantidad de títulos es pequeña, pero se complica mucho para aquellas bases de datos de texto completo con miles y miles de revistas. El catalogador debe decidir a qué nivel analizar un documento: ¿se describe una base

de datos o deben incluirse los registros individuales?, en el caso de las páginas Web en varios niveles, por ejemplo, ¿cada vínculo se describe por separado?, ¿es un documento o varios?

Beall (1997) agrega como característica especial la "**volatilidad**" y la distingue en tres tipos: **externa**, referida al cambio de URL o desaparición del sitio completo; **interna**, relacionada con el mantenimiento de los enlaces referidos en la página y la **intelectual**, en cuanto a la actualización de contenidos. Señala que los sitios mantenidos por individuos tienen mayor grado de volatilidad que aquellos asociados con instituciones.

Sintetizando lo expuesto, se presentan en el siguiente cuadro las principales diferencias de los recursos electrónicos en línea con los libros, en cuanto a catalogación se refiere:

Libros	Recursos electrónicos en línea
Presentación estática.	Variabilidad en la presentación.
Responsables se identifican fácilmente (portada).	Muchos responsables involucrados, no siempre detallados cuando se presenta el recurso.
Forma de publicación normalizada.	Formas no familiares de publicación.
Generalmente soporte original.	Relaciones mal definidas de los materiales con otros soportes.
Contenido inmutable.	Mutabilidad.
Permanece en su ubicación.	Movilidad.
Interacción bibliotecario-usuario.	Ausencia de un receptor físico identificable (hombre o programa).

El cuadro anterior resume algunas de las razones por las cuales las nuevas formas de publicación alteran nuestros modelos establecidos para el control bibliográfico.

Las formas básicas de las publicaciones electrónicas en línea aún no son estables ni normalizadas por lo cual describirlas resulta particularmente complejo. Los conceptos establecidos en las reglas de catalogación, tales como la fuente principal de información (portada), son difíciles de aplicar a estos documentos. Es necesario ampliar los modelos abstractos, las normativas y los procedimientos de la bibliotecología, para poder describir las publicaciones electrónicas.

Por estas razones es que han surgido opiniones diversas entre los bibliotecólogos con respecto a la catalogación de los recursos en línea; así, hay quienes los consideran demasiado inestables como para ser efectivamente catalogados de la manera tradicional (Wendler, 2000; Méndez Rodríguez, 2002); otros esperan que se desarrollen buscadores poderosos y

herramientas automáticas de indización como para obviar la necesidad de catálogos; y algunos, como Xu (1996), Jul (1997), llegaron a la conclusión de que, a pesar de las numerosas dificultades que se encuentran al tratar de catalogarlos, los recursos de información disponibles en Internet deben ser catalogados como se viene haciendo con otros tipos de documentos para que puedan ser recuperados y utilizados eficientemente.

METADATOS

Existen diferentes definiciones sobre el término metadatos, siendo la más sencilla y de aceptación general: *“datos sobre los datos”*.

A continuación se presentan una serie de definiciones de diferentes autores:

- **Caplan (1995)** adopta la definición “datos sobre datos” y considera las fichas de los catálogos tradicionales como metadatos, ya que poseen datos bibliográficos: autor, título, editorial, etc., que se refieren a otros documentos.
- **Xu (1997)** sostiene que es un conjunto de elementos que pueden ser usados para describir y representar objetos de información. En esta definición, Xu habla de objetos de información, independientemente del medio o soporte.
- **Younger (1997)** entiende que los metadatos describen recursos, indican dónde están ubicados y qué se requiere para utilizarlos exitosamente.
- En el marco del **World Wide Web Consortium (W3 Consortium)**, se los define como información sobre objetos de información en la Web, que puede ser leída por computadora (Swick, 2002).
- **Dempsey y Heery (1998)** los definen como datos que describen los atributos de un recurso.
- **Tennant (1998)** define los metadatos como "información estructurada sobre información", y destaca que la palabra clave es "estructurada", ya que una descripción en texto libre no es suficiente; es necesario contar con ciertos elementos identificados formalmente y con una codificación para la especificación de una sintaxis dada.
- **Milstead y Feldman (1999)** consideran el registro bibliográfico en sí como metadatos y no necesariamente necesitarían estar codificados.
- Para **Wendler (2000)**, la definición es: "información necesaria para identificar, localizar, manejar y acceder a un recurso electrónico".
- **Gorman (2000)** diferencia lo que son estándares de estructura de los que prescriben el contenido del registro bibliográfico. Para este autor, los metadatos han sido diseñados para responder a las necesidades de:

a) contar con una opción para la catalogación de los recursos electrónicos

b) encontrar una alternativa intermedia entre los altos costos de la catalogación "tradicional" y la simple búsqueda por palabra clave.

Para este autor no es necesario "inventar" algo nuevo ya que las técnicas de descripción bibliográfica de la catalogación tradicional cubren satisfactoriamente las necesidades de catalogación de los recursos electrónicos en línea, de acuerdo con las actualizaciones de las ISBD y de las Reglas de Catalogación Angloamericanas (AACR2R).

- **Taylor (2004)** introduce dos conceptos, el de contenido y el de codificación, presentes ambos en la acepción del uso común del término metadatos. Cuando sólo hay contenido, se lo denomina un registro bibliográfico; cuando hay sólo codificación, es identificado como una "estructura". Para esta autora, metadatos es una descripción codificada de un paquete de información.
- El documento *Understanding metadata* (NISO, 2004) define:

Esquema de metadatos: es un conjunto de elementos de metadatos diseñados para un objetivo específico. El esquema define los nombres de los elementos y la semántica. Opcionalmente pueden determinar reglas de cómo asignar el contenido.

Semántica: constituye la definición o el significado cada uno de los elementos.

Contenido: es el valor asignado a los elementos de metadatos.

Codificación: representa la sintaxis o estructura. Los metadatos pueden ser codificados en diferentes sintaxis. Un esquema de metadatos puede tener reglas de sintaxis que indican cómo deben ser codificados sus elementos y puede no tener reglas en cuyo caso se denomina esquema con sintaxis independiente.

En este trabajo se entiende por metadatos, siguiendo a Taylor (2004), la descripción codificada de un paquete de información como por ejemplo, un registro AACR2 codificado con MARC, un registro Dublin Core, etc. El propósito de los metadatos es proporcionar un nivel intermedio a través del cual se puede acceder a la información por medio de lo que uno desea ver o buscar, sin tener que investigar enormes cantidades de texto completo irrelevante.

Funciones

Según Dempsey y Heery (1998) a un usuario, persona o programa, le interesaría respecto de los recursos electrónicos:

- Descubrir o saber de la existencia de éstos.
- Además de una descripción básica (autor, título, tema), le interesa conocer el origen, la integridad, la autenticidad y las relaciones con otros recursos.
- Informarse acerca de los derechos de propiedad intelectual.
- Saber acerca de las posibilidades de manipulación, si se puede extraer parte de él, qué herramientas son necesarias para ello, qué equipamiento mínimo se requiere.

- Conocer bajo qué condiciones se puede utilizar, si es necesario pagar o es de libre acceso, si está restringido a personas de una institución, si se solicita clave de acceso.
- Saber qué interfaces técnicas, qué protocolos de acceso se utilizan, qué tipo de búsquedas soporta, cuáles son los formatos de salida.

Los metadatos pretenden que toda la información sobre los datos esté disponible en línea, así las funciones básicas son:

1. Describir un recurso de información (identificación).
2. Brindar puntos de acceso (recuperación).
3. Codificar la descripción para que sea recuperada automáticamente por los buscadores (localización y acceso).

Como puede apreciarse las funciones anteriormente citadas coinciden con las de la catalogación tradicional.

A estas funciones se agregan (Lagoze, 1996):

1. Proveer términos y condiciones para el uso de dicha información.
2. Presentar datos administrativos.
3. Establecer vínculos y relaciones, por ejemplo de un artículo con la revista que los contiene, entre un documento y su traducción, entre los componentes de un multimedia.
4. Dar datos estructurales y técnicos.
5. Facilitar la interoperabilidad de sistemas para el intercambio de información.
6. Brindar datos sobre la historia o procedencia, como su fuente original y transformaciones subsiguientes.

Las funciones enumeradas en el párrafo anterior están referidas a la administración y el uso de los recursos de información.

Para Ianella y Waugh (1997), los metadatos cumplen la función de:

- Resumir el significado (es decir, sobre qué tratan los datos).
- Permitir la búsqueda.
- Ayudar a los usuarios a determinar si la información es la que realmente buscan.
- Prevenir que ciertos usuarios (ej. niños) accedan a contenidos no permitidos.
- Recuperar y poder usar una copia del recurso.
- Instruir sobre cómo interpretar los datos (p. ej.: formato, codificación, encriptación).

- Ayudar a decidir qué formato escoger, en el caso que se presenten varios.
- Presentar información sobre el uso, como condiciones legales, tamaño, fecha original.
- Informar sobre la historia, como ser la fuente original y las transformaciones subsiguientes.
- Dar datos para contactarse con los propietarios.
- Indicar relaciones con otros recursos (como ser vínculos a versiones anteriores o posteriores, documentos derivados, programas que se requieren para su utilización).
- Proveer información sobre la administración de los datos (p. ej.: requerimientos para su archivo).

En síntesis, a partir de lo expuesto por Dempsey y Heery (1998), Lagoze (1996), Iannella y Waugh (1997), y como se desprende de la definición de Taylor (2004) anteriormente citada, es posible afirmar que las funciones primarias de los metadatos son facilitar la identificación, ubicación, recuperación, manipulación y el uso de recursos de información accesibles en línea.

Tipología

De acuerdo a la naturaleza de los datos que describen, existen tres tipos de metadatos (Lazinger, 2001):

1. **descriptivos:** información bibliográfica.
2. **estructurales:** referidos a formatos y estructuras.
3. **administrativos:** incluye permisos, copyright y toda información relacionada al acceso.

Weibel (1997) distingue tres modelos para la creación y el uso de los metadatos:

1. **modelo embebido** (*embedded model*), los metadatos son formulados por el creador del documento electrónico y la descripción reside incluida en éste. No se requiere de otro sistema adicional de organización. Esto es equivalente a la catalogación en la fuente o catalogación en publicación (CIP).
2. **generados por terceros** (*third-party metadata*), una agencia crea, reúne y administra los registros, pero éstos no están incluidos en el recurso en sí. Se generarían productos similares a las bases de datos referenciales. Si bien el registro está separado del recurso, se le puede agregar un vínculo para acceder directamente a él.

3. **visualización filtrada** (*the view-filter mode*): una agencia administra registros de metadatos, con diferentes estructuras y los traduce en una descripción común compartida por todos los usuarios. Tiene la gran ventaja de proveer una interfase única de búsqueda.

De acuerdo con la riqueza de la descripción y nivel de detalle, los metadatos pueden clasificarse en tres niveles (Dempsey & Heery, 1998):

Nivel 1: datos no estructurados, extraídos automáticamente de los documentos e indizados para la búsqueda.

Nivel 2: estructurados, son relativamente nuevos y muchos han sido diseñados *ad hoc* para su uso en Internet. En principio no estuvieron normalizados, pero actualmente algunos están surgiendo como normas, soportan búsquedas por campo y son relativamente fáciles de aplicar, están pensados para que los incluya el autor del documento. Por ejemplo: Dublin Core.

Nivel 3: ricos para la descripción de documentos, diseñados para ser aplicados por profesionales de las ciencias de la información, son complejos y basados en normas internacionales. Su nivel de detalle implica alta prestación en la recuperación y dificultad en su aplicación. Por ejemplo, el formato MARC.

NORMATIVAS INTERNACIONALES PARA LA DESCRIPCIÓN DE RECURSOS ELECTRÓNICOS

Las reglas de descripción bibliográfica que prescriben el contenido de la catalogación fueron desarrolladas mucho antes que los estándares de codificación y continúan existiendo como normativas separadas.

Gorman (2000) y Taylor (2004) diferencian entre los **estándares de contenido** y los de **estructura**. Los primeros son reglas que indican qué datos deben ser registrados para la identificación del documento y en qué forma deben registrarse. Los estándares de estructura son los que se conocen como formatos; éstos permiten el registro de los datos descriptivos de manera codificada que puede ser leída e interpretada por la computadora.

Cuando se habla de metadatos, como señala Taylor (2004), es difícil hacer esta división entre contenido y estructura, ya que la descripción a través de metadatos tiene intrínsecamente

asociada una codificación. Un registro de metadatos puede ser creado primero determinando el contenido descriptivo y luego codificándolo, o se puede comenzar con una estructura de códigos y rellenar luego con los datos cada campo. A su vez, un mismo contenido puede ser codificado con estructuras diferentes, y ciertos metadatos incluyen ambas especificaciones tanto para la codificación como para el contenido.

Estándares de contenido

Los estándares para registro de los datos descriptivos que existen actualmente, probados por largo tiempo, perfeccionados y utilizados ampliamente por las bibliotecas, son AACR2 e ISBD. Si bien no es objetivo de este trabajo analizar con exhaustividad las reglas de catalogación, a continuación se presenta brevemente el modo en que tratan los recursos electrónicos ambas normativas.

a) AACR2R

En la actualidad se mantiene la segunda edición de las AACR, que fue revisada en el año 2002, y actualizada en los años 2003, 2004 y 2005, con los cambios aprobados por el *Joint Steering Committee for Revision of AACR (JSC)* que, a su vez, cambió su propia denominación por *Joint Steering Committee for Development of RDA (JSC)* en abril de 2007. El JSC se encuentra trabajando en el nuevo código de catalogación “RDA: Resource Description and Access”. Según los anuncios del JSC se implementará en junio del 2010. Para que el organismo se pueda concentrar en este trabajo la última actualización para las AACR2 fue en 2005.

El capítulo 9 es el dedicado a recursos electrónicos. Las enmiendas del año 2001 de las AACR2 incluyen una revisión de las reglas para describirlos. Uno de los cambios es el reemplazo del término “archivo de computadora” por el término “recurso electrónico”; otro es la inclusión de la distinción entre recursos electrónicos de acceso directo y de acceso remoto. En ese momento se realizaron los cambios para ajustar la descripción a la ISBD (ER). En 2004 se eliminó el área 3 (Área de tipo y extensión del recurso). También se autorizó la aplicación de términos de uso común para describir el soporte de recursos electrónicos de acceso directo, como por ejemplo: CD-ROM, DVD-ROM. Se proveyó la opción de registrar la extensión de los recursos de acceso remoto (por ejemplo, 1 texto electrónico) en el área de descripción física.

Las áreas que deben consignarse son:

- Título y mención de responsabilidad (título propio, designación general del material, título paralelo, subtítulo, una o más menciones de responsabilidad).
- Edición (versión, tirada, actualización, nivel).

- Publicación, distribución, etc., (lugar, editor, productor o distribuidor, fecha).
- Descripción física (soporte, características técnicas, dimensiones).
- Serie (título de la serie, subserie, número de serie).
- Notas.
- Área de número normalizado y condiciones de disponibilidad.

En 2002, en las AACR2 el capítulo 12 cambia su denominación de “Publicaciones seriadas” a “Recursos Continuos” y se incorpora la categoría de *recurso integrado (integrating resource)* definido de la siguiente manera:

Un recurso integrado es aquel que se adiciona o se cambia mediante actualizaciones que no permanecen separadas y que se integran para formar un todo. Los recursos integrados pueden ser finitos o continuos. (AACR2, 2002)

Este concepto es aplicable a los recursos electrónicos dinámicos como son las bases de datos y los sitios Web, por lo que en estos casos deberá consultarse también el Capítulo 12 para su descripción.

Las RDA están pensadas para ayudar a las bibliotecas a progresar hacia el entorno digital, apoyan el intercambio de metadatos entre las comunidades que usan diferentes esquemas. Los usuarios podrán utilizar el contenido de RDA con muchos otros esquemas de codificación (por ejemplo, MODS (Metadata Object Description Standard), MARC 21 o Dublin Core). El código RDA es adaptable y flexible, con el potencial de poder ser utilizado por otras comunidades además de las bibliotecas. Provee un marco flexible para el contenido de la descripción de recursos digitales y a la vez sirve a las necesidades de organización de recursos tradicionales de las bibliotecas, permite la descripción de múltiples aspectos de una obra dado que se utiliza el modelo conceptual FRBR (*Functional Requirements for Bibliographic Records*). RDA ha identificado y agregado elementos no incluidos en RCAA2, que se utilizan comúnmente para la descripción de recursos digitales.

b) ISBD

Los recursos electrónicos en ISBD (2007) son tratados de dos formas, dependiendo de si su acceso es directo o remoto. Por **acceso directo** se entiende que hay un medio físico que se puede describir, por ejemplo, un disco, casete, cartucho, que debe ser insertado en una computadora o periférico. Por **acceso remoto** se entiende que no hay un medio físico que

manipular, sino que el acceso es sólo a través de un dispositivo de entrada-salida (por ejemplo, una terminal) conectado a un sistema o red de computadoras, o utilizando recursos almacenados en un disco duro.

De acuerdo a las ISBD, las áreas de la descripción que deben incluirse son:

- Título y mención de responsabilidad (título propio, designación general del material, título paralelo, subtítulo, una o más menciones de responsabilidad).
- Edición (versión, tirada, actualización, nivel).
- Publicación, producción y distribución (lugar, editor, productor o distribuidor, fecha).
- Descripción física (soporte, características técnicas, dimensiones) se aplica sólo a recursos de acceso local, para los de acceso remoto no se proveen estos datos.
- Serie (título de la serie, subserie, número de serie).
- Notas.
- Identificador del recurso y términos de disponibilidad.

El área de tipo y extensión de recurso (Área 3 presente en la ISBD(ER) 1997) ha sido eliminada para recursos electrónicos en la edición consolidada de 2007. En tanto que como designación general del material (DGM) se adopta [Recurso electrónico].

En el área de notas, incluye las siguientes notas específicas:

- Requerimientos del sistema: características del equipo en cuanto a hardware y software. Esta nota es obligatoria para recursos de acceso directo y precede a las demás.
- Modo de acceso: protocolo requerido y ubicación del recurso (URL, servidor FTP, etc.). Esta nota es obligatoria para recursos de acceso remoto.
- Ítem descripto: en el caso de acceso remoto debe registrarse siempre la fecha de consulta.
- Identificador, por ejemplo **DOI** (*Digital Object Identifier*).

A partir de 2002, siguiendo los cambios de la AACR2, se incorpora también la categoría de “*recurso integrado*”, ciertos recursos electrónicos, como por ejemplo las páginas Web, se tratan como recursos integrados.

ANÁLISIS DE LOS ESQUEMAS DE METADATOS

Los esquemas de metadatos que se analizan en este trabajo se han seleccionado en función de ciertos criterios: que estén actualmente en aplicación, que hayan sido generados en forma cooperativa o por instituciones con trayectoria reconocida en la descripción bibliográfica, que posean una base bibliotecológica, y con posibilidades de que se sigan actualizando. Por lo tanto, los esquemas de metadatos que se comparan en este trabajo son MARC21/MARC XML, Dublin Core, MODS y TEI.

Para elaborar la grilla de comparación se adoptaron los criterios que surgen de la adaptación de los ya propuestos por Heery (1996), Dempsey & Heery (1998) y Taylor (2004), como se expuso anteriormente. A continuación se define su alcance.

1. Año de origen.

4.2. Organismo creador.

4.3. **Mantenimiento:** se refiere especialmente al organismo encargado del desarrollo, actualización y mantenimiento del esquema de metadatos.

4.4. **Documentación existente:** existencia de manuales, tutoriales, listas de distribución, etc.

4.5. **Objetivos/misión:** la finalidad con que han sido creados.

4.6. **Disciplina/ámbito de aplicación:** área académica o grupo profesional que utiliza el formato de metadatos.

4.7. **Corriente teórica a la que responden:** si proviene de la comunidad bibliotecaria, informática, de editores o proveedores.

4.8. **Simplicidad/complejidad:** en cuanto a la facilidad de creación del registro, si se requieren habilidades especiales, si están diseñados para que asigne los metadatos el autor, el editor o los proveedores centralizados.

4.9. **Interoperabilidad:** entendida según la definición del documento *Understanding metadata*. (NISO, 2004) como la habilidad de sistemas múltiples con diferentes plataformas de hardware y software, estructuras de datos e interfaces, para intercambiar datos con una mínima pérdida de contenido y funcionalidad.

4.10. **Extensibilidad:** posibilidad de incluir mayor nivel de detalle, si es necesario, mediante campos opcionales, calificadores, etc.

4.11. **Flexibilidad:** posibilidad de elegir los campos/elementos a incluir.

4.12. **Estructura:** elementos o campos de la codificación y características de éstos.

a. **Opción:** si contiene campos/elementos que son opcionales y obligatorios.

a-b. **Repetición:** si admite campos y subcampos repetibles.

a-c. **Subcampos:** admisión o no de subcampos.

a-d. **Control:** posibilidad de controlar valores en algunos campos.

13. **Contenido:** si prescriben cómo debe ser elaborada la descripción en cada campo o si se valen de normativas externas, como por ejemplo AACR2.

~~43.14.~~ **Nivel de catalogación permitida:** se refiere al nivel de detalle de catalogación permitido, asociado a nivel de complejidad 1, 2 ó 3, que va desde los formatos simples, a los ricos o muy estructurados.

~~43.15.~~ **Actualización:** últimas versiones o modificaciones disponibles.

~~43.16.~~ **Contribución a la normalización internacional:** protocolos de Internet asociados y clases de normas que incluyen, y si en sí mismos se han consolidado como normas de hecho o de derecho. Por normas de hecho se entiende a las adoptadas por organismos o asociaciones profesionales (IFLA, ALA, etc.) y por normas de derecho, a las adoptadas por organismos de normalización (ISO, AENOR, etc.)

~~43.17.~~ **Sintaxis:** la sintaxis del esquema de metadatos que indica cómo deben ser codificados sus elementos.

~~43.18.~~ **Resultados obtenidos hasta el presente:** se refiere al grado de difusión del esquema de metadatos en cuanto a aplicaciones concretas.

A continuación se presenta el análisis de cada esquema de metadatos, según las categorías descriptas.

MARC21/MARCXML

AÑO DE ORIGEN:

MARC es un término genérico aplicado a una familia de formatos *MAchine Readable Cataloguing* (Catalogación legible por máquina). MARC21, surgió en 1998 de la armonización de USMARC y CANMARC. En 2004, la *British Library* discontinuó UKMARC y se sumó a MARC21.

La tarea de desarrollar un formato legible por computadora para los registros del catálogo se inició en la Biblioteca del Congreso de Estados Unidos a principios de 1965. La meta era desarrollar uno que empleara campos de longitud variable en un registro legible por máquina, que pudiera emplearse para representar e intercambiar datos bibliográficos entre sistemas automatizados. De esta forma se diseñó un sistema que fuera de "comunicación" y no sólo un formato de procesamiento o almacenamiento.

La versión inicial de MARC XML apareció en 2002, como un marco para trabajar con formato MARC en un ambiente XML. El objetivo era que fuese flexible y extensible como para permitir a los usuarios trabajar con MARC de acuerdo a sus propias necesidades.

ORGANISMO CREADOR: *Library of Congress*

MANTENIMIENTO: a cargo de la *Network Development and MARC Standards Office, Library of Congress* y la *Standards and the Support Office, Library and Archives Canada*, que trabajan en permanente consulta con varias comunidades de usuarios.

DOCUMENTACIÓN EXISTENTE: manuales en línea en varios idiomas accesibles a través de la Web de LC, tutoriales y listas de distribución.

OBJETIVOS/MISIÓN: proveer un mecanismo mediante el cual las computadoras intercambien, usen e interpreten información bibliográfica. Ha sido desarrollado inicialmente para registro de documentos impresos, específicamente obras monográficas, y luego fue incorporando campos para otro tipo de recursos: publicaciones seriadas, archivos de computadora, mapas, música, registros sonoros, manuscritos, material audiovisual y recursos electrónicos.

Se ha definido una familia de formatos para albergar cinco tipos de datos: bibliográficos, existencias, autoridades, clasificación e información comunitaria.

MARC ha sido diseñado también para soportar funciones de despliegue visual en pantalla y para la recuperación de información.

DISCIPLINA/ÁMBITO DE APLICACIÓN: se utiliza para describir documentos en distintos formatos y en todas las disciplinas, y es empleado especialmente por la comunidad bibliotecaria.

CORRIENTE TEÓRICA A LA QUE RESPONDE: proviene de la comunidad bibliotecaria; por lo tanto, es ampliamente utilizado en ese ámbito. Se puede decir que proviene de una corriente estructuralista.

SIMPLICIDAD/COMPLEJIDAD: MARC es un formato complejo; para la creación de un registro bibliográfico completo se requiere personal especializado y capacitado, con conocimiento profundo de las reglas de catalogación. Dada su complejidad, presenta un alto costo de la descripción.

INTEROPERABILIDAD: puede ser exportado a distintos formatos bibliográficos y de metadatos.

MARC ha sido mapeado a los siguientes formatos:

- MODS
- Dublin Core
- MARC Character Sets to UCS/Unicode

- Digital Geospatial Metadata

Los siguientes metadatos han sido mapeados a MARC21.

- MODS
- Dublin Core
- UNIMARC to MARC 21
- ONIX (Online Information Exchange)
- Digital Geospatial Metadata to MARC

EXTENSIBILIDAD: permite incluir mayor nivel de detalle, si es necesario, mediante campos opcionales, o el empleo de ciertos subcampos.

FLEXIBILIDAD: MARC tiene poca flexibilidad a la hora de elegir los campos para una descripción bibliográfica. Hay campos obligatorios establecidos.

ESTRUCTURA:

Un registro MARC incluye tres elementos: la estructura, la designación del contenido y los datos bibliográficos. A su vez, la estructura de un registro MARC tiene tres componentes principales: una cabecera, con extensión fija, la cual contiene elementos necesarios para procesar el registro; un directorio, que consiste de una serie de entradas que incluyen la etiqueta, su extensión y la posición de inicio de cada una de ellas; y los campos variables, los cuales conforman la mayor parte.

En MARC21 Bibliográfico, existen campos de datos diseñados para incluir los nombres de las personas y organismos que crearon el documento o contribuyeron a su creación, diferenciando entre autoría personal y corporativa, así como entre punto de acceso principal y secundario (campos 1XX y 7XX). Proporciona campos para mostrar y dar acceso a varios títulos relacionados con el documento (campos 240, 245, 246, 730, 740). Tiene un campo para la información sobre la edición (250), uno para la publicación (260) y otro para la descripción física (300). La información relacionada es codificada en los campos de serie (4XX y 8XX), en los campos de enlace (76X-78X), igual que en algunos campos de notas (5XX). Los campos 05X-09X contienen números de clasificación o información apropiada para la localización física de los documentos, mientras que el campo 856 permite la localización y acceso a los recursos electrónicos disponibles en la Web.

Características generales de los campos:

a) Opción: contiene campos que son opcionales y obligatorios

a)b) Repetición: en teoría todos los campos y subcampos son repetibles, pero la naturaleza de ciertos datos hace que determinados campos o subcampos sean no

repetibles como por ejemplo el campo 100 (campo no repetible) \$a (subcampo no repetible) \$c (subcampo repetible).

a)c) Subcampos: contiene numerosos subcampos.

a)d) Control: recomienda controlar valores en algunos campos, como por ejemplo en el uso de vocabularios controlados y en el control de autoridades.

En el siguiente cuadro se presenta un esquema de los grupos de campos de datos y luego se detalla el subconjunto específico para recursos electrónicos:

Etiquetas		Datos que se registran
00X		Campos de control
0XX		Campos de datos codificados
1XX	100	Punto acceso principal. Autor personal
	110	Punto acceso principal. Autor corporativo
	111	Punto acceso principal. Conferencia
	130	Punto acceso principal. Título uniforme
2XX	240	Título uniforme
	245	Título propiamente dicho y mención de responsabilidad
	246	Variante del título
	250	Mención de edición
	256	Características del archivo de computadora
	260	Lugar, editor, fecha
	270	Dirección postal, etc. del editor
3XX	300	Descripción física
4XX		Serie
5XX	Notas	
	506	Nota de restricción de acceso
	516	Nota de tipo de archivo o dato digital
	521	Nota de audiencia, a quién va dirigido el material
	530	Nota de otros soportes adicionales disponibles
	534	Nota de versión original
	538	Nota de requerimientos del sistema y modo de acceso
	540	Nota que determinan el uso y la reproducción
	545	Nota de datos biográficos o históricos
	555	Nota de índice acumulativo
562	Nota de identificación de la copia y versión	

6XX		Puntos de acceso por materia
7XX		Puntos de acceso secundarios, autores, títulos
8XX		Puntos de acceso secundarios de serie. Otros
	856	Acceso y localización electrónica
9XX		Campos locales

El formato MARC se ha actualizado para incluir características propias de los recursos electrónicos con los siguientes campos:

538 requerimientos del sistema y modo de acceso - Nota

256 características del archivo digital

516 tipo de archivo o dato digital - Nota

856 acceso y localización electrónica

Los tres primeros campos (256, 538, 516) se aplican tanto a archivos digitales locales como remotos.

El campo 856 “Acceso y localización electrónica” fue desarrollado en 1992 especialmente para los recursos electrónicos disponibles en la Web o a través de redes. En él se incluyen los siguientes datos (MARC21 Concise Bibliographic Format, 2009):

Indicadores:

Primero	Método de acceso
#	No definido
0	Correo electrónico
1	FTP
2	Acceso Remoto (Telnet)
3	Línea telefónica (Dial-up)
4	HTTP
7	Método de acceso provisto en subcampo \$2
Segundo	Relación
#	No definido
0	Recurso (indica que se provee la localización del recurso como un todo)
1	Versión del recurso (el recurso que se cataloga no es el electrónico, pero posee una versión electrónica indicada aquí)

2	Recurso relacionado (la localización que se provee no es el recurso en sí, sino uno relacionado)
8	No se despliega constante.

Códigos de subcampo:

\$a	Nombre del servidor
\$b	Número de acceso
\$c	Información sobre compresión
\$d	Ruta
\$f	Nombre del archivo electrónico
\$h	Clave del usuario solicitante
\$i	Instrucciones de procesamiento
\$j	Bits por segundo
\$k	Clave de acceso (<i>Password</i>)
\$l	Clave de conexión (<i>Logon, login</i>)
\$m	Contacto para obtener ayuda
\$n	Localización del servidor indicado en el subcampo \$a
\$o	Sistema operativo
\$p	Puerto
\$q	Tipo de formato electrónico
\$r	Configuración
\$s	Tamaño del archivo
\$t	Emulación de la terminal
\$u	Identificador Uniforme de Recursos (URI)
\$v	Horario de acceso
\$w	Número de control del registro relacionado
\$x	Nota sobre acceso restringido
\$y	Link al texto
\$z	Nota sobre acceso público
\$2	Otro método de acceso especificado en \$2
\$3	Material o parte específica
\$6	Relación
\$8	Campos relacionados y secuencia numérica

CONTENIDO: MARC es en sí un estándar de estructura. El contenido de la mayoría de los elementos está definido por normativas fuera del formato en sí mismo, por ejemplo: *Anglo-American Cataloguing Rules, Library of Congress Subject Headings, National Library of Medicine Classification*.

NIVEL DE CATALOGACIÓN PERMITIDA: No es necesario decir mucho acerca de la conveniencia de MARC para codificar datos descriptivos, ya que es bien conocido el nivel de detalle que permite, ubicándose en Nivel 3 (Dempsey & Heery, 1998). Están previstos dos niveles definidos para los Estados Unidos: un nivel mínimo y el nivel completo o nacional que contiene la suficiente información catalográfica para ser utilizada por agencias de catalogación tanto nacionales como internacionales.

ACTUALIZACIÓN: la última actualización publicada es de octubre del 2009. En 2008 se actualizó la versión MARC21 Lite, un subconjunto del formato completo, que incluye los datos esenciales necesarios para la descripción bibliográfica de los documentos.

CONTRIBUCIÓN A LA NORMALIZACIÓN INTERNACIONAL: es estándar de hecho para la representación de registros bibliográficos. La estructura de los registros MARC es una implementación de normativas nacionales e internacionales: *Information Interchange Format* (ANSI Z39.2) y *Format for Information Exchange* (ISO 2709) respectivamente.

MARC utiliza además, los siguientes estándares:

- ISO 2108: International Standard Book Numbering (ISBN) (ANSI/NISO Z39.21)
- ISO 3901: International Standard Recording Code (ISRC)
- ISO 3297: International Standard Serial Number (ISSN) (ANSI/NISO Z39.9)
- ISO 10324: Holdings Statements - Summary Level (ANSI/NISO Z39.71)

SINTAXIS: se puede codificar en distintas sintaxis como HTML y XML.

Ejemplo de un registro MARC codificado en XML tomado de:

<http://www.loc.gov/standards/marcxml/Sandburg/sandburg.xml>

```
<?xml version="1.0" encoding="UTF-8" ?>
<collection xmlns="http://www.loc.gov/MARC21/slim">
<record>
<leader>01142cam 2200301 a 4500</leader>
<controlfield tag="001">92005291</controlfield>
<controlfield tag="003">DLC</controlfield>
<controlfield tag="005">19930521155141.9</controlfield>
<controlfield tag="008">920219s1993 caua j 000 0 eng</controlfield>
<datafield tag="010" ind1="" ind2="">
```

```

<subfield code="a">92005291</subfield>
  </datafield>
: <datafield tag="020" ind1="" ind2="">
  <subfield code="a">0152038655 :</subfield>
  <subfield code="c">$15.95</subfield>
  </datafield>
: <datafield tag="040" ind1="" ind2="">
  <subfield code="a">DLC</subfield>
  <subfield code="c">DLC</subfield>
  <subfield code="d">DLC</subfield>
  </datafield>
: <datafield tag="042" ind1="" ind2="">
  <subfield code="a">lcac</subfield>
  </datafield>
: <datafield tag="050" ind1="0" ind2="0">
  <subfield code="a">PS3537.A618</subfield>
  <subfield code="b">A88 1993</subfield>
  </datafield>
: <datafield tag="082" ind1="0" ind2="0">
  <subfield code="a">811/.52</subfield>
  <subfield code="2">20</subfield>
  </datafield>
: <datafield tag="100" ind1="1" ind2="">
  <subfield code="a">Sandburg, Carl,</subfield>
  <subfield code="d">1878-1967.</subfield>
  </datafield>
: <datafield tag="245" ind1="1" ind2="0">
  <subfield code="a">Arithmetic /</subfield>
  <subfield code="c">Carl Sandburg ; illustrated as an anamorphic adventure by Ted
  Rand.</subfield>
  </datafield>
: <datafield tag="250" ind1="" ind2="">
  <subfield code="a">1st ed.</subfield>
  </datafield>
: <datafield tag="260" ind1="" ind2="">
  <subfield code="a">San Diego :</subfield>
  <subfield code="b">Harcourt Brace Jovanovich,</subfield>
  <subfield code="c">c1993.</subfield>
  </datafield>
: <datafield tag="300" ind1="" ind2="">
  <subfield code="a">1 v. (unpaged) :</subfield>
  <subfield code="b">ill. (some col.) ;</subfield>
  <subfield code="c">26 cm.</subfield>
  </datafield>
: <datafield tag="500" ind1="" ind2="">
  <subfield code="a">One Mylar sheet included in pocket.</subfield>
  </datafield>
: <datafield tag="520" ind1="" ind2="">
  <subfield code="a">A poem about numbers and their characteristics. Features
  anamorphic, or distorted, drawings which can be restored to normal by viewing
  from a particular angle or by viewing the image's reflection in the provided Mylar
  cone.</subfield>
  </datafield>
: <datafield tag="650" ind1="" ind2="0">
  <subfield code="a">Arithmetic</subfield>

```

```

<subfield code="x">Juvenile poetry.</subfield>
</datafield>
= <datafield tag="650" ind1="" ind2="0">
<subfield code="a">Children's poetry, American.</subfield>
</datafield>
= <datafield tag="650" ind1="" ind2="1">
<subfield code="a">Arithmetic</subfield>
<subfield code="x">Poetry.</subfield>
</datafield>
= <datafield tag="650" ind1="" ind2="1">
<subfield code="a">American poetry.</subfield>
</datafield>
= <datafield tag="650" ind1="" ind2="1">
<subfield code="a">Visual perception.</subfield>
</datafield>
= <datafield tag="700" ind1="1" ind2="">
<subfield code="a">Rand, Ted,</subfield>
<subfield code="e">ill.</subfield>
</datafield>
</record>
</collection>

```

RESULTADOS OBTENIDOS HASTA EL PRESENTE: el formato MARC, en cuanto a recursos electrónicos, es utilizado por las bibliotecas que participan o participaron de los diferentes proyectos de OCLC, como InterCat y CORC, actualmente absorbidos por el servicio *Connexion* que alimenta el catálogo WorlCat y por aquellas bibliotecas que incluyen recursos Web en sus catálogos basados en MARC.

DUBLIN CORE

AÑO DE ORIGEN: 1995.

Dublin Core es la forma abreviada para el *Dublin Core Metadata Element Set* (DCMES). Nace en 1994 por idea de Stuart Weibel, un investigador de OCLC. En 1995, se lleva a cabo una reunión de trabajo realizada en Dublin (Ohio, Estados Unidos), en la cual participaron bibliotecarios, investigadores en bibliotecas digitales, expertos en computación y lenguajes de marcado de texto electrónico, y acordaron un conjunto de elementos o datos que deberían proveer los creadores de documentos electrónicos sobre éstos.

ORGANISMO CREADOR: Online Computer Library Center (OCLC) y National Center for Supercomputing Applications (NCSA).

MANTENIMIENTO: The Dublin Core Metadata Initiative (DCMI)

Desde 1995, año en que nació el proyecto, han tenido lugar distintos seminarios de trabajo y conferencias internacionales para llegar a acuerdos sobre el Dublin Core.

DOCUMENTACIÓN EXISTENTE: manuales en línea en el sitio Web oficial de Dublin Core, que reúne toda la información sobre la iniciativa.

OBJETIVOS/MISIÓN: Dublin Core (DC) ha sido creado *ad hoc* para la catalogación de documentos electrónicos con el objetivo de definir un conjunto de elementos básicos para que los propios autores puedan describir sus recursos Web, de modo que puedan ser recuperados por los motores de búsqueda de Internet. Entre los propósitos también está definir el marco para la interoperabilidad entre conjuntos de metadatos y apoyar el desarrollo de elementos específicos para ciertas disciplinas. DC nació como una forma muy simple de generar metadatos. De esta manera la idea fue apartarse de las estrictas normas bibliotecológicas y permitir una catalogación mucho más sencilla para el público en general.

DISCIPLINA/ÁMBITO DE APLICACIÓN: se utiliza para describir documentos en cualquier ámbito o materia. DC es aplicado por diferentes comunidades fuera de la bibliotecología tales como la educativa, la gubernamental, la dedicada a los archivos, y la que se ocupa de los museos.

CORRIENTE TEÓRICA A LA QUE RESPONDEN: proviene de la interacción de la comunidad bibliotecaria e informática.

SIMPLICIDAD/COMPLEJIDAD: en su origen es simple; son sólo 15 elementos básicos y está pensado para que el autor pueda incluir los metadatos al momento de crear el documento. Existen dos corrientes: los *minimalistas*, que bogan por mantener el nivel de simplicidad y los *estructuralistas*, que proponen un mayor nivel de distinción y extensibilidad. De allí que surjan las posturas de utilizar DC con calificadores o sin calificadores. No requiere de catalogadores expertos; y debido a su sencillez presenta una alta productividad.

INTEROPERABILIDAD: puede ser exportado a distintos formatos bibliográficos y de metadatos.

EXTENSIBILIDAD: permite incluir mayor nivel de detalle, si es necesario, mediante los calificadores.

FLEXIBILIDAD: nada en DC es obligatorio, todos los elementos son opcionales; así el usuario elige ajustar la descripción de acuerdo a sus necesidades.

ESTRUCTURA: El estándar Dublin Core incluye dos niveles: simple y calificado. El nivel simple contiene los quince elementos básicos y el nivel calificado incluye tres elementos adicionales (Audiencia, Procedencia y Administrador de los derechos), así como un grupo de calificadores que refinan la semántica de los elementos. Utiliza etiquetas basadas en palabras y no numéricas.

Características generales de los elementos:

- a) Opción: todos los elementos son opcionales.
- b) Repetición: todos los elementos son repetibles.
- ~~b)c~~ Sucampos: no posee subcampos pero sí calificadores.
- ~~b)d~~ Control: no requiere control de valores en los elementos.

Los 15 elementos básicos pueden reunirse en 3 grupos relacionados con:

- el contenido.
- la propiedad intelectual.
- los datos referidos al uso, instalación y manipulación.

En el siguiente cuadro se presentan los 15 elementos:

CONTENIDO	PROPIEDAD INTELLECTUAL	INSTALACIÓN/ MANIPULACIÓN
Título	Creador	Fecha
Materia	Editor	Tipo
Descripción	Colaborador	Formato
Fuente	Derechos	Identificador
Idioma		
Relación		
Cobertura		

A continuación se detallan los elementos de Dublin Core (ANSI/NISO Standard Z39.85-2007) presentados de la siguiente forma, primero el nombre del elemento en español, entre paréntesis la etiqueta utilizada para la marcación de documentos, y luego una breve explicación del significado. En la explicación que se proporciona, el nombre de los elementos aparece en español para su mejor comprensión, pero en el marcado de los documentos las etiquetas se utilizan siempre en inglés.

Elementos relacionados con el **contenido**:

Contenido	
Título (<i>Title</i>)	Nombre dado al recurso.
Materia (<i>Subject</i>)	Palabras clave, frases o códigos de clasificación que describan los temas; se sugiere la utilización de vocabularios controlados y sistemas de clasificación formales.
Descripción (<i>Description</i>)	Resumen del contenido del recurso (<i>abstract</i>), tabla de contenido, representación gráfica del contenido, reseña de texto libre.
Fuente (<i>Source</i>)	Referencia a otro recurso del cual se deriva o proviene, en su totalidad o en parte.
Idioma (<i>Language</i>)	Idioma en el que está expresado el contenido. Se recomienda usar la normativa RFC 3066, la cual en conjunto con ISO 639, define códigos de dos y tres letras para los idiomas. Ejemplos “en ” o “eng ” para el inglés y “en-GB ” para inglés del Reino Unido.
Relación (<i>Relation</i>)	Referencia a un recurso relacionado.
Cobertura (<i>Coverage</i>)	Cobertura del contenido intelectual del recurso, incluye ubicación espacial (lugar o coordenadas geográficas), período de tiempo (fecha o rango de fechas), o jurisdicción (para entidades administrativas). Se recomienda el uso de vocabulario controlado, por ejemplo Thesaurus of Geographic Names [TGN].

Elementos relacionados con la **propiedad intelectual**:

Propiedad intelectual	
Creador (<i>Creator</i>)	Entidad primaria responsable de la creación del contenido del recurso, puede ser una persona, una organización o un servicio. Se registra el nombre del creador.
Editor (<i>Publisher</i>)	Nombre del responsable de que el recurso esté

	disponible para su consulta, puede ser una persona, una organización o un servicio.
Colaborador (<i>Contributor</i>)	Nombre del responsable que ha hecho contribuciones al contenido del recurso, puede ser una persona, una organización o un servicio.
Derechos (<i>Rights</i>)	Información sobre derechos de uso y propiedad intelectual, referidos al servicio que provee la información. Si no está detallado en el recurso, no deben asentarse suposiciones.

Elementos relacionados con el **uso y manipulación**:

Uso y manipulación	
Fecha (<i>Date</i>)	Fecha de un evento en el ciclo de vida del recurso; asociada con la creación, modificación o disponibilidad del recurso. Se recomienda codificar la fecha según la Norma ISO 8601 [W3CDTF], en el formato <i>aaaa-mm-dd</i> .
Tipo (<i>Type</i>)	La naturaleza o género del recurso en cuanto a su contenido, incluye términos que describen categorías, funciones, géneros, en relación con el contenido. Se recomienda tomar la tipología de acuerdo a un vocabulario controlado, por ejemplo el DCMI Type Vocabulary [DCT]. Para describir manifestaciones físicas o digitales del recurso, se recomienda usar el elemento “Formato”.
Formato (<i>Format</i>)	Manifestación física o digital del recurso. Incluye tipo de medio y tamaño o duración del recurso. Se utiliza para identificar el software, hardware u otros requerimientos necesarios para la visualización y manipulación del

	recurso. Se recomienda utilizar vocabulario controlado, por ejemplo la lista Internet Media Types [MIME].
Identificador (<i>Identifier</i>)	Referencia unívoca del recurso en un contexto dado. Se recomienda utilizar un identificador formal, por ejemplo Uniform Resource Identifier (URI), Uniform Resource Locator (URL), Digital Object Identifier (DOI), e International Standard Book Number (ISBN).

Para el caso de Dublin Core Calificado se agregan los siguientes elementos:

Audiencia (<i>Audience</i>)	A quiénes está dirigido el recurso. La audiencia puede ser determinada por el creador o editor del recurso o por terceros.
Procedencia (<i>Provenance</i>)	Datos de cualquier cambio en la propiedad y custodia del recurso desde su creación, que sea significativo en cuanto a autenticidad, integridad o interpretación.
Administrador de los derechos (<i>Rights Holder</i>)	Organización o persona que poseen o administran los derechos del recurso.

Calificadores (*qualifiers*)

Hay dos grandes categorías de calificadores (Using Dublin Core. Dublin Core Qualifiers, 2005):

1. **Refinamiento de elementos básicos:** hacen más específico el significado de un elemento.
2. **Esquemas de codificación:** identifican los códigos normalizados que ayudan a la forma de expresar el contenido. Incluyen vocabularios controlados, reglas de catalogación, normas de codificación, esquemas de clasificación. Por ejemplo: codificar la fecha del siguiente modo "2002.12.01".

En el siguiente cuadro se presenta un resumen de los calificadores de acuerdo al documento citado:

Elemento	Calificadores/Refinamiento del elemento	Esquemas de codificación del elemento
Título (Title)	Alternativo (<i>Alternative</i>)	-
Creador (creator)	-	-
Materia (Subject)	-	LCSH MeSH CDD LCC CDU
Descripción (Description)	Tabla de Contenidos (<i>Table Of Contents</i>) Resumen (<i>Abstract</i>)	-
Editor (Publisher)	-	-
Colaborador (Contributor)	-	-
Fecha (Date)	Creado (<i>Created</i>) Válido (<i>Valid</i>) Disponible (<i>Available</i>) Aparición/Publicación (<i>Issued</i>) Modificado (<i>Modified</i>) Fecha de aceptación (<i>Date Accepted</i>) Fecha de Copyright (<i>Date Copyrighted</i>) Fecha de envío (<i>Date Submitted</i>)	DCMI Period W3C-DTF
Tipo (Type)	-	DCMI Type Vocabulary
Formato (Format)	-	IMT
	Extensión (<i>Extent</i>)	-
	Medio (<i>Médium</i>)	-
Identificador (Identifier)	-	URI
	Cita bibliográfica (<i>Bibliographic Citation</i>)	-

Fuente (Source)	-	URI
Idioma (Language)	-	ISO 639-2 RFC 3066
Relación (Relation)	Es versión de (<i>Is Version Of</i>) Tiene versión (<i>Has Version</i>) Es remplazado por (<i>Is Replaced By</i>) Remplaza a (<i>Replaces</i>) Es requerido por (<i>Is Required By</i>) Requiere (<i>Requires</i>) Es parte de (<i>Is Part Of</i>) Tiene parte (<i>Has Part</i>) Es referenciado por (<i>Is Referenced By</i>) Referencia a (<i>References</i>) Es formato de (<i>Is Format Of</i>) Tiene formato (<i>Has Format</i>) Conforme a (<i>Conforms To</i>)	URI
Cobertura (Coverage)	Espacial (<i>Spatial</i>)	DCMI Point ISO 3166 DCMI Box TGN
	Temporal (<i>Temporal</i>)	DCMI Period W3C-DTF
Derechos (Rights)	Derechos de acceso (<i>Access Rights</i>)	-
	Licencia (<i>License</i>)	URI
Audiencia (Audience)	Mediador (<i>Mediator</i>) Nivel educativo (<i>Education Level</i>)	-
Procedencia (Provenance)	-	-
Administrador de derechos (Rights Holder)	-	-

A continuación se detallan los calificadores sugeridos para cada elemento, presentados de la siguiente forma: bajo cada elemento el nombre de los calificadores, entre paréntesis la etiqueta utilizada y debajo una nota con explicación del significado. Para los casos en que se sugiere la aplicación de una normativa para el registro del contenido, se enumeran debajo del calificador correspondiente.

Elemento	Calificadores/ Refinamiento del elemento	Significado	Esquemas de codificación del elemento
Título	Alternativo (<i>Alternative</i>)	Cualquier forma	

		alternativa del título utilizada como sustituto del título formal de un recurso. Puede incluir abreviaturas y traducciones.	
Materias			LCSH Encabezamientos de Materias de la <i>Library of Congress</i> (EEUU). MeSH Encabezamientos de materias de la <i>National Library of Medicine</i> (EEUU). DDC Clasificación Decimal de Dewey . LCC Clasificación de la <i>Library of Congress</i> (EEUU). UDC Clasificación Decimal Universal.
Descripción	Tabla de contenidos <i>(table of contents)</i> Abstract o resumen <i>(abstract)</i>		
Fecha	Creado <i>(created)</i>	Fecha de creación del recurso.	DCMI Period Especificación de los límites de un intervalo de tiempo establecidos por DCMI. W3C-DTF Reglas de codificación para fechas y tiempo del W3C, basado en la norma ISO 8601.
	Válido <i>(valid)</i>	Fecha (o rango de fechas) de la validez de un recurso.	
	Disponible <i>(available)</i>	Fecha o rango de fechas en las que el recurso estuvo o estará disponible.	
	Aparición o publicación <i>(issued)</i>	Fecha de aparición formal (publicación) del recurso.	
	Modificado <i>(modified)</i>	Fecha en que un recurso cambia o es actualizado.	
	Fecha de aceptación <i>(Date Accepted)</i>	Fecha en que fue aceptado el recurso (Por ejemplo: una tesis, un artículo de revista).	
	Fecha de Copyright <i>(Date Copyrighted)</i>	Fecha de la declaración de copyright.	
	Fecha de envío <i>(Date Submitted)</i>	Fecha en que fue enviado o presentado para consideración un	

		recurso (Por ejemplo: tesis, artículo de revista).	
Tipo de recurso			DCMI Type Vocabulary Listado de tipos de recursos utilizado para categorizar la naturaleza o género del contenido de un recurso establecido por DCMI.
Formato	Extensión (<i>extent</i>)	El tamaño o duración de un recurso.	
	Medio (<i>medium</i>)	Soporte físico o material del recurso.	IMT (<i>Internet Media Type</i>) Códigos para diferentes tipos de medios definidos por IANA (<i>Internet Assigned Numbers Authority</i>).
Identificador del recurso	Cita bibliográfica (<i>Bibliographic Citation</i>)	Referencia bibliográfica del recurso.	URI Identificador uniforme del recurso (Uniform Resource Identifier).
Idioma			ISO 639-2 Códigos ISO de idiomas. RFC 4646 Etiquetas para la identificación de idioma, RFC es el nombre que da la IETF (Internet Engineering Task Force) a sus especificaciones.
Relación	Es versión de: (<i>is version of</i>)	Es una versión, edición o adaptación del recurso referenciado aquí. El cambio de versión implica modificaciones significativas del contenido, más allá que un mero cambio de formato.	URI Identificador uniforme del recurso (Uniform Resource Identifier).
	Tiene versión: (<i>has version</i>)	Tiene otra versión, edición o adaptación,	

		que se referencia aquí.	
	Es reemplazado por: (<i>is replaced by</i>)	Es suplantado, desplazado o actualizado por el recurso referenciado aquí.	
	Reemplaza a: (<i>replaces</i>)	Suplanta, desplaza o actualiza al recurso que se referencia aquí.	
	Es requerido por: (<i>is required by</i>)	Es requerido, ya sea física o lógicamente, por el recurso referenciado aquí.	
	Requiere: (<i>requires</i>)	Requiere del recurso referenciado aquí para funcionar, para ser distribuido o para completar su contenido.	
	Es parte de: (<i>is part of</i>)	Es una parte física o lógica del recurso referenciado aquí.	
	Tiene parte: (<i>has part</i>)	Incluye el recurso referenciado aquí.	
	Es referenciado por: (<i>is referenced by</i>)	Es citado, referenciado o apuntado por otro que se nombra aquí.	
	Referencia a: (<i>references</i>)	Cita o apunta a otro nombrado aquí.	
	Es formato de: (<i>is format of</i>)	Tiene el mismo contenido intelectual que el citado aquí, sólo que es presentado en otro formato.	
	Tiene formato: (<i>has format</i>)	Es anterior al recurso referenciado aquí, este último constituye otro formato con el mismo contenido intelectual.	
	Conforme a: (<i>conforms to</i>)	Hace referencia a alguna normativa a la que se ajusta el recurso.	
Cobertura	Espacial (<i>spatial</i>)	Cobertura geográfica del contenido intelectual del recurso.	Punto DCMI (<i>point DCMI</i>) Código que identifica un punto en el espacio utilizando coordenadas geográficas definido por DCMI. ISO 3166 Código ISO para la representación de los nombres de países.

			DCMI Box Código que identifica una región del espacio utilizando límites geográficos. TGN Tesauro Getty de Nombres Geográficos.
	Temporal (<i>temporal</i>)	Características temporales del contenido intelectual del recurso.	Período DCMI Código para especificar los límites de un intervalo de tiempo. W3C-DTF Código W3C para fechas y tiempo, basado en la norma ISO 8601.

CONTENIDO: No está regido por reglas de contenido, sino que se sugiere la normativa a aplicar en algunos elementos, como se vio en el apartado anterior.

NIVEL DE CATALOGACIÓN PERMITIDA: sencilla, concisa. Se lo ubica en un Nivel 2 (Dempsey & Heery, 1998).

ACTUALIZACIÓN: la última actualización es la Versión 1.1. En enero de 2008, se definieron dominios formales y rangos en las definiciones de las propiedades, consignados en el documento “DCMI Metadata Terms” que es más abarcativo que el “Dublin Core Metadata Element Set, Version 1.1”.

CONTRIBUCIÓN A LA NORMALIZACIÓN INTERNACIONAL: es norma de derecho desde octubre de 2001 (ANSI/NISO Standard Z39.85-2001). Las versiones actualizadas son:

- ISO Standard 15836-2003 (Febrero 2003): <http://www.niso.org/international/SC4/n515.pdf>
- ANSI/NISO Standard Z39.85-2007 (Mayo 2007): The Dublin Core Metadata Element Set <http://www.niso.org/standards/resources/Z39-85-2007.pdf>
- IETF RFC 5013 of August 2007 [[RFC5013](#)]

SINTAXIS: posee independencia sintáctica. DC se puede codificar en distintas sintaxis tales como HTML, XML o RDF.

Ejemplo de un registro DC codificado en XML tomado de:

<http://www.loc.gov/standards/marcxml/Sandburg/sandburgdc.xml>

```
<?xml version="1.0" ?>
<dc xmlns="http://purl.org/dc/elements/1.1/">
  <title>Arithmetic </title>
  <creator>Sandburg, Carl, 1878-1967.</creator>
  <creator>Rand, Ted, ill.</creator>
  <type />
  <publisher>San Diego :Harcourt Brace Jovanovich,</publisher>
  <date>c1993.</date>
  <language>eng</language>
  <description>A poem about numbers and their characteristics. Features
  anamorphic, or distorted, drawings which can be restored to normal by viewing
  from a particular angle or by viewing the image's reflection in the provided
  Mylar cone.</description>
  <description>One Mylar sheet included in pocket.</description>
  <subject>Arithmetic</subject>
  <subject>Children's poetry, American.</subject>
  <subject>Arithmetic</subject>
  <subject>American poetry.</subject>
  <subject>Visual perception.</subject>
</dc>
```

Ejemplo del mismo registro DC ahora codificado en HTML:

```
<HTML>
<HEAD>
  <TITLE> Ejemplos</TITLE>
  <META NAME="DC.Title" CONTENT=" Arithmetic /">
  <META NAME="DC.Creator" CONTENT="Sandburg, Carl, 1878-1967.">
  <META NAME="DC.Creator" CONTENT=" Rand, Ted, ill.">
  <META NAME="DC.Publisher" CONTENT="San Diego :Harcourt Brace
  Jovanovich, ">
  <META NAME="DC.Date" CONTENT=" c1993.">
  <META NAME="DC.Type" CONTENT="Texto">
  <META NAME="DC.Format" CONTENT="text/html">
  <META NAME="DC.Language" CONTENT=" eng">
  <META NAME="DC.Description" CONTENT="A poem about numbers and
  their characteristics. Features anamorphic, or distorted, drawings which
  can be restored to normal by viewing from a particular angle or by viewing
  the image's reflection in the provided Mylar cone. >
  <META NAME="DC.Subject" CONTENT="Arithmetic">
  <META NAME="DC.Subject" CONTENT="Children's poetry, American.">
  <META NAME="DC.Subject" CONTENT=" American poetry.">
  <META NAME="DC.Subject" CONTENT=" Visual perception.">
</HEAD>
<BODY>

</BODY>
</HTML>
```

Y por último, el mismo ejemplo de un registro DC codificado en RDF:

```

<RDF xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
<Description about="Aritmetic">
  <dc:title>Arithmetic /</dc:title>
  <dc:creator>Sandburg, Carl, 1878-1967.</dc:creator>
  <dc:creator>Rand, Ted, ill.</dc:creator>
  <dc:publisher>San Diego :Harcourt Brace Jovanovich,</dc:publisher>
  <dc:date>c1993.</dc:date>
  <dc:language>eng</dc:language>
  <dc:description>A poem about numbers and their characteristics. Features
    anamorphic, or distorted, drawings which can be restored to normal by
    viewing from a particular angle or by viewing the image's reflection in the
    provided Mylar cone.</dc:description>
  <dc: description>One Mylar sheet included in pocket.</dc:description>
  <dc:subject>Arithmetic</dc:subject>
  <dc:subject>Children's poetry, American.</dc:subject>
  <dc:subject>Arithmetic</dc:subject>
  <dc:subject>American poetry.</dc:subject>
  <dc:subject>Visual perception.</dc:subject>
</Description>
</RDF>

```

RESULTADOS OBTENIDOS HASTA EL PRESENTE: son muchos los proyectos en marcha basados en DC en diversas áreas del conocimiento y se han desarrollado numerosas aplicaciones informáticas para la asignación automática de metadatos en DC. Como ejemplo se puede mencionar el proyecto **CORC** (*Cooperative Online Resource Catalog*) de OCLC. Se inició como proyecto en 1998 y pasó a ser un servicio OCLC en 1999. CORC impulsó la creación de una base de datos de recursos Web alimentada en forma cooperativa, con registros tanto en formato MARC como en Dublin Core, que luego son almacenados en XML y enviados o exportados al usuario final, tanto en formato MARC como en DC. Desde 2002 forma parte del servicio integrado de catalogación OCLC Connexion.

MODS (<i>Metadata Object Description Schema</i>)

AÑO DE ORIGEN: 2002

ORGANISMO CREADOR: *Library of Congress*

MANTENIMIENTO: *Network Development and MARC Standards Office, Library of Congress.*

DOCUMENTACIÓN EXISTENTE: sitio Web oficial con manuales en línea y lista de distribución.

OBJETIVOS/MISIÓN: es un subconjunto de MARC21 bibliográfico, creado para permitir la extracción de datos de manera selectiva de los registros MARC21 existentes y crear registros originales. MODS puede ser utilizado con los siguientes propósitos:

- Como formato SRU (Search/Retrieval via URL), protocolo para búsquedas en Internet basado en XML.
- Como una extensión de METS (*Metadata Encoding and Transmission Standard*).
- Para representar metadatos para cosecha automática de datos.
- Para realizar descripciones originales de recursos en sintaxis XML.
- Para representar un registro MARC simplificado en XML.
- Para crear metadatos embebidos en XML en un recurso electrónico.

DISCIPLINA: utilizado para distintos propósitos pero fundamentalmente para aplicaciones bibliotecarias.

CORRIENTE TEÓRICA A LA QUE RESPONDE: proviene de la comunidad bibliotecaria.

SIMPLICIDAD/COMPLEJIDAD: se puede decir que MODS posee una complejidad intermedia. Es más rico y complejo que DC. Es más simple y más orientado al usuario que MARC.

INTEROPERABILIDAD: puede ser exportado a distintos formatos bibliográficos y de metadatos al igual que MARC. Mapeos disponibles: MARCXML a MODS; MODS a MARCXML; Dublin Core a MODS; MODS a Dublin Core y entre las distintas versiones de MODS.

EXTENSIBILIDAD: permite extender la descripción conectando con descripciones alojadas fuera del registro, por ejemplo un resumen (*abstract*) disponible en otra dirección Web.

FLEXIBILIDAD: la flexibilidad está dada al ser opcionales todos sus elementos.

ESTRUCTURA: incluye un subconjunto de campos de MARC21 bibliográfico. La mayoría de los elementos de MODS tiene su equivalente en MARC21 aunque también tiene incorporado otros específicos para la descripción de recursos digitales. Consta de 20 elementos principales,

subelementos y atributos que pueden ser usados en uno o más elementos. Utiliza etiquetas basadas en palabras y no numéricas.

Características de los elementos:

- a) Opción: todos los elementos y atributos son opcionales.
- a)b) Repetición: los elementos son repetibles y los atributos no repetibles.
- a)c) Subcampos: contiene subelementos y atributos.
- a)d) Control: en ciertos elementos el atributo “*authority*” permite registrar con qué herramienta se realiza el control del valor.

Se listan, a continuación, los elementos principales presentados de la siguiente manera: el nombre del elemento entre paréntesis la etiqueta y debajo una breve explicación de su significado. Luego, se listan los subelementos y atributos con sus correspondientes etiquetas entre paréntesis y una breve nota aclaratoria en los casos en que se considera necesaria.

Elementos de MODS

1. Información de título (<i>titleInfo</i>)	
Una palabra, frase, carácter o grupo de caracteres que normalmente aparecen en el recurso y le dan nombre a éste o a la obra contenida en él.	
Subelementos:	Título (<i>title</i>) Subtítulo (<i>subtitle</i>) Número de parte (<i>partNumber</i>) Nombre de parte (<i>partName</i>) Texto a ignorar para clasificación (<i>nonSort</i>)
Atributos:	Tipo (<i>type</i>): este atributo se aplica cuando es necesario para identificar el tipo de título que se registra: <div style="text-align: center;"> abreviado (<i>abbreviated</i>) traducido (<i>translated</i>) alternativo (<i>alternative</i>) uniforme (<i>uniform</i>) </div> Autoridad (<i>authority</i>): se refiere al nombre de la lista autorizada de donde se toma un valor controlado que se registra. Etiqueta de visualización (<i>displayLabel</i>): cuando se necesita desplegar un texto adicional asociado con el título. Link externo (<i>xlink</i>): se utiliza para referenciar un link externo al

	<p>recurso.</p> <p>ID – (ID) -: se utiliza para establecer un link interno.</p> <p>Idioma (<i>lang</i>): se utiliza para especificar el idioma utilizado en el elemento, aplicando los códigos ISO 639-2/b.</p> <p>Idioma en XML (<i>xml:lang</i>): se utiliza para especificar el idioma utilizado en el elemento, de acuerdo al estándar XML aplicando la especificación RFC 3066.</p> <p>Alfabeto (<i>script</i>): Especifica el alfabeto utilizado en los elementos, utilizando los códigos de la norma ISO 15924 (<i>Code for the representation of names of scripts</i>).</p> <p>Transliteración (<i>transliteration</i>): especifica la técnica de transliteración utilizada en el elemento.</p>
<p>2. Nombre (<i>name</i>)</p>	
<p>Nombre de una persona, organización o evento (conferencia, reunión) asociada con el recurso.</p>	
<p>Subelementos:</p>	<p>Parte de nombre (<i>namePart</i>) incluye cada parte del nombre que desea describirse por separado.</p> <p>Atributos</p> <p>Tipo:</p> <ul style="list-style-type: none"> Fecha (<i>Date</i>) Familia (<i>Family</i>) Nombre de pila (<i>Given</i>) Términos de tratamiento (<i>termsOfAddress</i>): utilizado para anotar títulos, números asociados con un nombre, por ejemplo, Jr., II, etc. <p>Nombre no formateado (<i>displayForm</i>): es utilizado para indicar la forma no estructurada del nombre, tal como aparece en el recurso.</p> <p>Filiación (<i>affiliation</i>): contiene el nombre de una organización, institución etc., con la cual la persona o entidad registrada en <name> estuvo asociada al tiempo de la creación del recurso. Puede contener también otros elementos que son parte de la filiación, tal como dirección de mail, dirección postal, cargo, etc.</p> <p>Rol (<i>role</i>): término que designa la relación de la entidad registrada en <name> con el recurso.</p>

	<p>Subelemento:</p> <p>Término de rol (<i>roleTerm</i>): contiene la forma textual o codificada que designa el rol.</p> <p>Atributos:</p> <p>Tipo (<i>type</i>):</p> <p> Texto (<i>text</i>): rol expresado en palabras.</p> <p> Código (<i>code</i>): rol expresado a través de un código.</p> <p>Autoridad (<i>authority</i>)</p> <p>Descripción (<i>description</i>): puede ser usado para agregar una descripción textual para el nombre cuando sea necesario, por ejemplo: para distinguirlo de otros nombres.</p>
Atributos:	<p>ID (<i>ID</i>)</p> <p>Link externo (<i>xlink</i>)</p> <p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p> <p>Tipo (<i>type</i>):</p> <p> Personal (<i>personal</i>)</p> <p> Corporativo (<i>corporate</i>)</p> <p> Conferencia (<i>conference</i>)</p> <p>Autoridad (<i>authority</i>)g</p>
3. Tipo de recurso (<i>typeOfResource</i>)	
<p>Término que especifica el tipo general de contenido y características de un recurso.</p> <p>Valores utilizados:</p> <p> texto (<i>text</i>)</p> <p> material cartográfico (<i>cartographic</i>)</p> <p> música impresa (<i>notated music</i>)</p> <p> grabación sonora musical (<i>sound recording-musical</i>)</p> <p> grabación sonora no musical (<i>sound recording-nonmusical</i>)</p> <p> grabación sonora (<i>sound recording</i>)</p> <p> imagen fija (<i>still image</i>)</p> <p> imagen en movimiento (<i>moving image</i>)</p>	

<p>objeto tridimensional (<i>three dimensional object</i>)</p> <p>software, multimedia (<i>software, multimedia</i>)</p> <p>material mixto (<i>mixed material</i>)</p>	
Subelementos:	No posee.
Atributos:	<p>Colección (<i>collection=yes</i>): se utiliza cuando el recurso es una colección de ítems agrupados no publicados en conjunto originalmente.</p> <p>Manuscrito (<i>manuscript=yes</i>): para ítems escritos a mano o tipeados.</p> <p>Los atributos “colección” y “manuscrito” se utilizan asociados a un valor en particular utilizado en “Tipo de recurso”.</p>
4. Género (<i>Genre</i>)	
<p>Término que designa una categoría que caracteriza un estilo, forma o contenido particular, por ejemplo: artístico, musical, literario. Más específico que “Tipo de recurso”.</p>	
Subelementos:	No posee.
Atributos:	<p>Tipo (<i>type</i>): se utiliza si se desea distinguir entre diferentes aspectos del género, como clase, tipo de trabajo o estilo.</p> <p>Autoridad (<i>authority</i>)</p> <p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p>
5. Información de origen (<i>OriginInfo</i>)	
<p>Información sobre el origen del recurso, incluye lugar de origen o publicación, editor original y fechas asociadas.</p>	
Subelementos:	<p>Lugar (<i>place</i>): nombre de un lugar asociado con la publicación, tirada, fabricación, producción u origen del recurso.</p> <p>Término de lugar (<i>placeTerm</i>): utilizado para consignar el lugar en forma codificada o normalizada.</p> <p style="text-align: center;">Atributos:</p>

	<p>Tipo (<i>type</i>)</p> <p>Código (<i>Code</i>)</p> <p>Texto (<i>Text</i>)</p> <p>Autoridad (<i>authority</i>)</p> <p>Código países MARC (<i>Marccountry</i>)</p> <p><i>iso3166</i></p> <p>Editor (<i>publisher</i>): el nombre de la entidad que publica, imprime, distribuye, emite o produce el recurso.</p> <p>Fecha de emisión (<i>dateIssued</i>): fecha en que el recurso fue publicado o emitido.</p> <p>Atributos:</p> <p>Codificación (<i>encoding</i>)</p> <p><i>w3cdf:</i> fecha codificada según patron AAAA-MM-DD.</p> <p><i>iso8601:</i> AAAAMMDD (se pueden agregar horas, minutos y segundos).</p> <p><i>marc:</i> fecha codificada según campo 008/07-14.</p> <p>Punto (<i>point</i>):</p> <p>Inicio(<i>Start</i>): indica que es la primera fecha de un rango.</p> <p>Fin (<i>end</i>): indica que es la última fecha de un rango.</p> <p>Fecha clave (<i>keyDate</i>)</p> <p><i>Yes:</i> para distinguir una fecha entre otras.</p> <p>Calificador (<i>qualifier</i>)</p> <p>Aproximada (<i>aproximate</i>): fecha aproximada.</p> <p>Inferida (<i>inferred</i>): no tomada del recurso en sí.</p> <p>Cuestionable (<i>questionable</i>): fecha probable.</p>
--	---

Fecha de creación (*dateCreated*)

Atributos:

Codificación (*encoding*)

Punto (*point*)

Fecha clave (*keyDate*)

Calificador (*qualifier*)

Fecha de captura (*dateCaptured*)

Atributos:

Codificación (*encoding*)

Punto (*point*)

Fecha clave (*keyDate*)

Calificador (*qualifier*)

Fecha válida (*dateValid*)

Atributos:

Codificación (*encoding*)

Punto (*point*)

Fecha clave (*keyDate*)

Calificador (*qualifier*)

Fecha de modificación (*dateModified*)

Atributos:

Codificación (*encoding*)

Punto (*point*)

Fecha clave (*keyDate*)

Calificador (*qualifier*)

Fecha de copyright (*copyrightDate*)

Atributos:

Codificación (*encoding*)

Punto (*point*)

Fecha clave (*keyDate*)

Calificador (*qualifier*)

Otra fecha (*dateOther*)

	<p>Atributos:</p> <p>Codificación (<i>encoding</i>)</p> <p>Punto (<i>point</i>)</p> <p>Fecha clave (<i>keyDate</i>)</p> <p>Calificador (<i>qualifier</i>)</p> <p>Tipo (<i>type</i>)</p> <p>Edición (<i>edition</i>): edición o versión a la que pertenece el recurso.</p> <p>Emisión (<i>issuance</i>): indica cómo es emitido en recurso.</p> <p>Continuo (<i>continuing</i>)</p> <p>Monográfico (<i>monographic</i>)</p> <p>Frecuencia (<i>frequency</i>)</p> <p>Atributo:</p> <p>Autoridad (<i>authority</i>)</p>
Atributos:	<p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p>
6. Idioma (<i>language</i>)	
Subelementos:	<p>Término de idioma (<i>languageTerm</i>)</p> <p>Atributos:</p> <p>Tipo (<i>type</i>)</p> <p><i>text</i> : para expresar el idioma con palabras.</p> <p><i>Code</i>: para expresar el idioma codificado.</p> <p>Autoridad (<i>authority</i>)</p> <p><i>iso639-2b</i> – código de idiomas ISO 639-2.</p> <p><i>rfc3066</i> – código de idiomas RFC3066.</p>
Atributos:	Parte del objeto (<i>objectPart</i>): indica a qué parte del objeto aplica el

	idioma consignado.
7. Descripción física (<i>physicalDescription</i>)	
Subelementos:	<p>Forma (<i>form</i>): designación de una presentación física en particular del recurso.</p> <p>Atributos:</p> <p>Autoridad (<i>authority</i>)</p> <p>Tipo (<i>type</i>) (Ejemplos: material, técnica)</p> <p>Cualidades de reformato (<i>reformattingQuality</i>): un término que indica una evaluación general de las cualidades físicas de un recurso electrónico en relación con el uso que se pretenda dar.</p> <p>Acceso (<i>access</i>): el recurso provee acceso al ítem original.</p> <p>Preservación (<i>preservation</i>): es un recurso creado vía reformato para preservación del original.</p> <p>Respaldo (<i>replacement</i>): el recurso es de calidad suficiente para servir como respaldo si el original se pierde, se daña o se destruye.</p> <p>Tipo de medio (<i>internetMediaType</i>): identificación del tipo de formato electrónico.</p> <p>Extensión (<i>extent</i>)</p> <p>Origen digital (<i>digitalOrigin</i>):</p> <p>Nacido digital (<i>born digital</i>).</p> <p>Digitalizado (<i>reformatted digital</i>): digitalizado a partir de un original no digital.</p> <p>Microfilm digitalizado (<i>digitized microfilm</i>): digitalizado a partir de una microforma.</p> <p>Digitalizado de otra forma no digital (<i>digitized other analog</i>): ej. de una fotocopia.</p> <p>Nota(<i>note</i>): información adicional en forma de texto sobre descripción física.</p> <p>Atributos:</p>

	<p>Link externo (<i>xlink</i>)</p> <p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p> <p>Etiqueta de visualización (<i>displayLabel</i>)</p> <p>Tipo (<i>type</i>):</p>
Atributos:	<p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p>
8. Resumen (<i>abstract</i>)	
Subelementos:	No posee.
Atributos:	<p>Link externo (<i>xlink</i>)</p> <p>Idioma (<i>lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p> <p>Etiqueta de visualización (<i>displayLabel</i>)</p> <p>Tipo (<i>type</i>)</p>
9. Tabla de contenidos (<i>tableOfContents</i>)	
Subelementos:	No posee
Atributos:	<p>Link externo (<i>xlink</i>)</p> <p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p> <p>Etiqueta de visualización (<i>displayLabel</i>)</p> <p>Tipo (<i>type</i>):</p> <p>Ej. Contenido incompleto, contenido parcial.</p>
10. Audiencia (<i>targetAudience</i>)	
Subelementos:	No posee.
Atributos:	<p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p>

	<p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p> <p>Autoridad (<i>authority</i>): se sugiere Lista MARC21 de audiencia.</p>
11. Nota (<i>note</i>)	
Subelementos:	No posee.
Atributos:	<p>ID (<i>ID</i>)</p> <p>Link externo (<i>xlink</i>)</p> <p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p> <p>Etiqueta de visualización (<i>displayLabel</i>)</p> <p>Tipo (<i>type</i>): la mayoría de las notas implementadas coinciden con las de MARC21</p>
12. Materia (<i>subject</i>)	
Subelementos:	<p>Tópico (<i>topic</i>)</p> <p>Geográfico(<i>geographic</i>)</p> <p>Temporal (<i>temporal</i>)</p> <p style="padding-left: 40px;">Atributos:</p> <p style="padding-left: 80px;">Codificación (<i>encoding</i>)</p> <p style="padding-left: 80px;">Punto (<i>point</i>)</p> <p style="padding-left: 80px;">Fecha clave (<i>keyDate</i>)</p> <p style="padding-left: 80px;">Calificador (<i>qualifier</i>)</p> <p>Título (<i>titleInfo</i>)</p> <p>Nombre (<i>name</i>)</p> <p>Código geográfico (<i>geographicCode</i>)</p> <p style="padding-left: 40px;">Atributo:</p> <p style="padding-left: 80px;">Autoridad (<i>authority</i>)</p> <p style="padding-left: 120px;">Código países MARC (<i>Marccountry</i>)</p> <p style="padding-left: 120px;"><i>iso3166</i></p> <p>Género (<i>genre</i>)</p> <p>Jerarquía geográfica (<i>hierarchicalGeographic</i>)</p> <p style="padding-left: 40px;">Continente (<i>continent</i>)</p> <p style="padding-left: 40px;">País (<i>country</i>)</p>

	<p>Provincia (<i>province</i>)</p> <p>Región (<i>region</i>)</p> <p>Estado (<i>state</i>)</p> <p>Territorio (<i>territory</i>)</p> <p>Condado, distrito (<i>county</i>)</p> <p>Ciudad (<i>city</i>)</p> <p>Isla (<i>island</i>)</p> <p>Área (<i>area</i>)</p> <p>Área extraterreste (<i>extraterrestrialArea</i>)</p> <p>Sección de ciudad (<i>citySection</i>)</p> <p>Datos cartográficos (cartographics)</p> <p>Escala (<i>scale</i>)</p> <p>Proyección (<i>projection</i>)</p> <p>Coordenadas (<i>coordinates</i>)</p> <p>Profesión u ocupación (occupation)</p> <p>Si el contenido del recurso está relacionado con alguna profesión u ocupación, no se utiliza para la ocupación de los creadores del recurso.</p>
Atributos:	<p>ID (<i>ID</i>)</p> <p>Link externo (<i>xlink</i>)</p> <p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p> <p>Autoridad (<i>authority</i>): de acuerdo a códigos de encabezamientos de materias mantenido por MARC.</p>
13. Clasificación (classification)	
Subelementos:	No posee.
Atributos:	<p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p> <p>Autoridad (<i>authority</i>): de acuerdo a códigos de esquemas de</p>

	<p>clasificación mantenido por MARC.</p> <p>Edición (<i>edition</i>): edición del esquema de clasificación utilizado.</p> <p>Etiqueta de visualización (<i>displayLabel</i>)</p>
14. Item relacionado (<i>relatedItem</i>)	
Información que identifica otro recurso relacionado con el que se está describiendo.	
Subelementos:	Cualquier elemento de MODS puede ser tomado como subelemento.
Atributos:	<p>ID (<i>ID</i>)</p> <p>Link externo (<i>xlink</i>)</p> <p>Etiqueta de visualización (<i>displayLabel</i>)</p> <p>Tipo (<i>type</i>): describe la relación del recurso relacionado con el que se está describiendo.</p> <p>Valores:</p> <ul style="list-style-type: none"> Precedente (<i>preceding</i>) Continuación (<i>succeeding</i>) Original (<i>original</i>) Anfitrión (<i>host</i>) Componente (<i>constituent</i>) Serie (<i>series</i>) Otra version (<i>otherVersion</i>) Otro formato (<i>otherFormat</i>) Es referenciado por (<i>isReferencedBy</i>)
15. Identificador (<i>identifier</i>)	
Subelementos:	No posee.
Atributos:	<p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p> <p>Tipo (<i>type</i>)</p> <p>Valores sugeridos:</p> <ul style="list-style-type: none"> doi (Digital Objects Identifier) hdl (Handle) isbn (International Standard Book Number) ismn (International Standard Music Number)

	<p>isrc (International Standard Recording Code)</p> <p>issn (International Standard Serials Number)</p> <p>issue number</p> <p>istc (International Standard Text Code)</p> <p>lccn (Library of Congress Control Number)</p> <p>local</p> <p>matrix number</p> <p>music plate</p> <p>music publisher</p> <p>sici (Serial Item and Contribution Identifier)</p> <p>stock number</p> <p>upc (Universal Product Code)</p> <p>uri (Uniform Resource Identifier)</p> <p>videorecording identifier</p> <p>Invalid (yes): en el caso de un identificador inválido o cancelado.</p>
<p>16. Ubicación (location)</p>	
<p>Se utiliza para nombrar la institución que lo mantiene o aloja, puede utilizarse texto o códigos.</p>	
<p>Subelementos:</p>	<p>Ubicación física (<i>physicalLocation</i>)</p> <p>Atributos:</p> <p>Link externo (<i>xlink</i>)</p> <p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p> <p>Etiqueta de visualización (<i>displayLabel</i>)</p> <p>Autoridad (<i>authority</i>): si se utilizan códigos, pueden utilizarse los listados de organizaciones mantenidos por MARC.</p> <p>Tipo (<i>type</i>): utilizado para distinguir entre diferentes tipos de ubicaciones, por ejemplo: actual, anterior, de</p>

	<p>creación.</p> <p>Ubicación de estante (<i>shelfLocator</i>)</p> <p>URL (<i>url</i>)</p> <p>Atributos:</p> <p>Última fecha de acceso (<i>dateLastAccessed</i>)</p> <p>Etiqueta de visualización (<i>displayLabel</i>)</p> <p>Nota (<i>note</i>)</p> <p>Acceso (<i>access</i>)</p> <p>Valores:</p> <p>Vista preliminar (<i>preview</i>)</p> <p>Recurso en sí (<i>raw object</i>): cuando el objeto completo está en un único archivo</p> <p>Recurso en contexto (<i>object in context</i>)</p> <p>Uso (<i>usage</i>)</p> <p>Mantenimiento simple (<i>holdingSimple</i>) (no repetible)</p> <p>Información sobre copia (<i>copyInformation</i>)</p> <p>Forma (<i>form</i>) (no repetible)</p> <p>Atributo:</p> <p>Autoridad (<i>authority</i>)</p> <p>Sububicación (<i>sublocation</i>)</p> <p>Ubicación de estante (<i>shelfLocator</i>)</p> <p>Ubicación electrónica (<i>electronicLocator</i>)</p> <p>Nota (<i>note</i>)</p> <p>Atributos:</p> <p>Etiqueta de visualización (<i>displayLabel</i>)</p> <p>Tipo (<i>type</i>)</p> <p>Enumeración y cronología (<i>enumerationAndChronology</i>)</p> <p>Atributos:</p> <p>Tipo de unidad (<i>unitType</i>)</p>
--	---

	<p style="text-align: right;"><i>Nota: 1=basic bibliographic unit; 2=supplement; 3=index</i></p> <p>Mantenimiento externo (holdingExternal) (no repetible)</p>
17. Condiciones de acceso (accessCondition)	
Subelementos:	No posee.
Atributos:	<p>Link externo (xlink)</p> <p>Idioma (lang)</p> <p>Idioma en XML (xml:lang)</p> <p>Alfabeto (script)</p> <p>Transliteración (transliteration)</p> <p>Etiqueta de visualización (displayLabel)</p> <p>Tipo (type):</p> <p style="padding-left: 40px;">Valores sugeridos: restricción de acceso, uso y reproducción.</p>
18. Parte (part)	
Designación de partes físicas de un recurso en forma detallada.	
Subelementos:	<p>Detalle (detail)</p> <p style="padding-left: 40px;">Número (<i>number</i>)</p> <p style="padding-left: 40px;">Encabezamiento de la parte (<i>caption</i>)</p> <p style="padding-left: 40px;">Título (<i>title</i>)</p> <p style="padding-left: 40px;">Atributos:</p> <p style="padding-left: 80px;">Tipo (<i>type</i>)</p> <p style="padding-left: 80px;">Valores sugeridos: parte, volumen, número, capítulo, sección, párrafo, pista.</p> <p style="padding-left: 40px;">Nivel de la numeración (<i>level</i>)</p> <p>Extensión (extent)</p> <p style="padding-left: 40px;">Comienzo (<i>start</i>)</p> <p style="padding-left: 40px;">Fin (<i>end</i>)</p> <p style="padding-left: 40px;">Total (<i>total</i>)</p> <p style="padding-left: 40px;">Abarque o lista (<i>list</i>)</p> <p style="padding-left: 40px;">Atributo: Unidad (<i>unit</i>) Por ejemplo: Páginas, minutos</p>

	<p>Fecha (<i>date</i>)</p> <p>Atributos:</p> <p>Codificación (<i>encoding</i>)</p> <p>Punto (<i>point</i>)</p> <p>Calificador (<i>qualifier</i>)</p> <p>Texto (<i>text</i>)</p> <p>Atributos:</p> <p>Link externo (<i>xlink</i>)</p> <p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p> <p>Transliteración (<i>transliteration</i>)</p> <p>Etiqueta de visualización (<i>displayLabel</i>)</p> <p>Tipo (<i>type</i>)</p>
Atributos:	<p>ID (<i>ID</i>)</p> <p>Tipo (<i>type</i>)</p> <p>Valores sugeridos: volumen, número, capítulo, párrafo, pista.</p> <p>Orden (<i>order</i>)</p>
19. Extensión (<i>extension</i>)	
Se utiliza para agregar otra información no cubierta por los elementos de MODS en texto libre.	
Subelementos:	No posee.
Atributos:	No posee.
20. Información del registro (<i>recordInfo</i>)	
Contiene información sobre el registro de metadatos.	
Subelementos:	<p>Creador del contenido registro (<i>recordContentSource</i>)</p> <p>Atributos:</p> <p>Autoridad (<i>authority</i>): código MARC de la organización.</p> <p>Idioma (<i>lang</i>)</p> <p>Idioma en XML (<i>xml:lang</i>)</p> <p>Alfabeto (<i>script</i>)</p>

	<p style="text-align: center;">Transliteración (<i>transliteration</i>)</p> <p>Fecha de creación del registro (<i>recordCreationDate</i>)</p> <p style="padding-left: 40px;"><i>Atributos:</i></p> <p style="padding-left: 80px;">Codificación (<i>encoding</i>)</p> <p style="padding-left: 80px;">Punto (<i>point</i>)</p> <p style="padding-left: 80px;">Fecha clave (<i>keyDate</i>)</p> <p style="padding-left: 80px;">Calificador (<i>qualifier</i>)</p> <p>Fecha de modificación del registro (<i>recordChangeDate</i>)</p> <p style="padding-left: 40px;"><i>Atributos:</i></p> <p style="padding-left: 80px;">Codificación (<i>encoding</i>)</p> <p style="padding-left: 80px;">Punto (<i>point</i>)</p> <p style="padding-left: 80px;">Fecha clave (<i>keyDate</i>)</p> <p style="padding-left: 80px;">Calificador (<i>qualifier</i>)</p> <p>Identificador del registro (<i>recordIdentifier</i>)</p> <p style="padding-left: 40px;"><i>Atributo:</i></p> <p style="padding-left: 80px;">Fuente (<i>source</i>): organización que le asigna el identificador.</p> <p>Origen del registro (<i>recordOrigin</i>)</p> <p>Idioma de la catalogación (<i>languageOfCataloging</i>)</p> <p style="padding-left: 40px;">Término de idioma (<i>languageTerm</i>)</p> <p style="padding-left: 80px;"><i>Atributos:</i></p> <p style="padding-left: 120px;">Tipo (<i>type</i>)</p> <p style="padding-left: 160px;"><i>text</i>: para expresar el idioma con palabras.</p> <p style="padding-left: 160px;"><i>code</i>: para expresar el idioma codificado.</p> <p style="padding-left: 80px;">Autoridad (<i>authority</i>)</p> <p style="padding-left: 120px;"><i>iso639-2b</i> – código de idiomas ISO 639-2.</p> <p style="padding-left: 120px;"><i>rfc3066</i> – código de idiomas RFC3066.</p>
--	---

	Esquema de catalogación (<i>descriptionStandard</i>)
Atributos:	Idioma (<i>lang</i>) Idioma en XML (<i>xml:lang</i>) Alfabeto (<i>script</i>) Transliteración (<i>transliteration</i>)

Elementos Raíz (*Root elements*)

1. Registro MODS (<i>mods</i>)	
Para un registro MODS simple	
Subelementos:	Todos los elementos de MODS
Atributos:	ID (<i>ID</i>) Versión de MODS (<i>version</i>)
2. Colección de registros MODS (<i>modsCollection</i>)	
Para un conjunto de registros MODS	
Subelementos:	MODS (<i>mods</i>)
Atributos:	No posee.

CONTENIDO: es un estándar de estructura, el contenido de la mayoría de los elementos se anota de acuerdo a normativas externas (p. ej.: AACR2) de manera análoga a MARC21.

NIVEL DE CATALOGACIÓN PERMITIDA: más sencillo que MARC21 y más orientado al usuario que MARC XML. Permite la conversión de los campos centrales de un registro MARC21, aunque pueden perderse algunos datos específicos. En el caso de una descripción original, permite crear un registro simple, en algunos casos usando mayor cantidad de etiquetas generales que las disponibles en un registro MARC. Atiende a la granularidad, se puede describir el todo o una parte de un recurso.

ACTUALIZACIÓN: la última versión es MODS 3.3 (2008). Durante el 2009 se han aprobado las modificaciones para incluir en la versión MODS 3.4

CONTRIBUCIÓN A LA NORMALIZACIÓN INTERNACIONAL: en 2004 ha sido registrado en NISO. Un aspecto importante a la hora de convertirse en estándar es la total compatibilidad con MARC21.

SINTAXIS: se codifica en XML.

Ejemplo de un registro MODS codificado en XML tomado de:

<http://www.loc.gov/standards/marcxml/Sandburg/sandburgmods.xml>

```

<?xml version="1.0" ?>
= <mods xmlns:xlink="http://www.w3.org/TR/xlink"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="http://www.loc.gov/mods/" xsi:schemaLocation="http://www.loc.gov/mods/
  http://www.loc.gov/standards/mods/mods.xsd">
= <titleInfo>
  <title>Arithmetic /</title>
  </titleInfo>
= <name type="personal">
  <namePart>Sandburg, Carl</namePart>
  <namePart type="date">1878-1967</namePart>
= <role>
  <text>creator</text>
  </role>
  </name>
= <name type="personal">
  <namePart>Rand, Ted</namePart>
= <role>
  <text>ill.</text>
  </role>
  </name>
  <typeOfResource>text</typeOfResource>
= <originInfo>
= <place>
  <code authority="marc">cau</code>
  <text>San Diego</text>
  </place>
  <publisher>Harcourt Brace Jovanovich</publisher>
  <dateIssued>c1993</dateIssued>
  <dateIssued encoding="marc">1993</dateIssued>
  <edition>1st ed.</edition>
  <issuance>monographic</issuance>
  </originInfo>
  <language authority="iso639-2b">eng</language>
= <physicalDescription>
  <form authority="marcform">print</form>
  <extent>1 v. (unpaged) : ill. (some col.) ; 26 cm.</extent>
  </physicalDescription>
  <abstract>A poem about numbers and their characteristics. Features anamorphic, or
    distorted, drawings which can be restored to normal by viewing from a particular
    angle or by viewing the image's reflection in the provided Mylar cone.</abstract>
  <targetAudience>juvenile</targetAudience>

```

```

<note type="statement of responsibility">Carl Sandburg ; illustrated as an anamorphic
  adventure by Ted Rand.</note>
<note>One Mylar sheet included in pocket.</note>
= <subject authority="lcs" >
  <topic>Arithmetic</topic>
  <topic>Juvenile poetry</topic>
  </subject>
= <subject authority="lcs" >
  <topic>Children's poetry, American</topic>
  </subject>
= <subject authority="lcs" >
  <topic>Arithmetic</topic>
  <topic>Poetry</topic>
  </subject>
= <subject authority="lcs" >
  <topic>American poetry</topic>
  </subject>
= <subject authority="lcs" >
  <topic>Visual perception</topic>
  </subject>
<classification authority="lcc">PS3537.A618 A88 1993</classification>
<classification edition="20" authority="ddc">811/.52</classification>
<identifier type="isbn">0152038655 :</identifier>
<identifier type="lcn">92005291</identifier>
= <recordInfo>
  <recordContentSource>DLC</recordContentSource>
  <recordCreationDate encoding="marc">920219</recordCreationDate>
  <recordChangeDate encoding="iso8601">19930521155141.9</recordChangeDate>
  <recordIdentifier source="DLC">92005291</recordIdentifier>
  </recordInfo>
</mods>

```

RESULTADOS OBTENIDOS HASTA EL PRESENTE: en la Web de la *Library of Congress*, en diciembre de 2009, había 30 proyectos listados de implementaciones de este esquema. Entre las instituciones figuran la *LC*, la *National Library of Australia*, la *Digital Library Federation*, varias universidades de Estados Unidos y la *Oxford University*.

TEI (*Text Encoding Initiative*)

AÑO DE ORIGEN: el proyecto se inicia en 1987; en 1994 se publican las Recomendaciones TEI (*TEI Guidelines*) y en diciembre del 2000 se crea un consorcio para su mantenimiento (*TEI-C*).

ORGANISMO CREADOR: el proyecto TEI surge de una investigación cooperativa entre tres asociaciones: *Association of Computers in the Humanities* (ACH), *Association for*

Computational Linguistics (ACL), y *Association of Literary and Linguistic Computing (ALLC)* y con fondos de: *U.S. National Endowment for the Humanities (NEH)*, *European Community*, *Mellon Foundation* y *Social Science and Humanities Research Council of Canada*.

Las recomendaciones TEI fueron publicadas en 1994 como resultado de un proyecto de investigación que comenzó en el año 1987. TEI establece que cualquier texto codificado según TEI debe tener, antes que nada, una cabecera que contenga la descripción de dicho texto. Las especificaciones de la cabecera fueron formuladas por el Comité de Documentación del proyecto, el cual estaba integrado por bibliotecarios y archivistas de Europa y Norteamérica, y están inspiradas en la catalogación tradicional. Esta cabecera incluye los metadatos.

MANTENIMIENTO: en el año 2000 se establece una nueva corporación sin fines de lucro denominada *TEI Consortium* para el desarrollo y mantenimiento del esquema, cuya oficina ejecutiva se ubica en Charlottesville, Virginia, con sedes en Brown University, Oxford University y University of Virginia.

DOCUMENTACIÓN EXISTENTE: manuales y tutoriales en línea en el sitio Web oficial, software, varios grupos de discusión, presentaciones para uso educativo y un wiki.

OBJETIVOS/MISIÓN: TEI ha sido desarrollado como un estándar internacional e interdisciplinario para bibliotecas, museos, editoriales corporativas y editores individuales para representar una gran variedad de textos literarios y lingüísticos para la investigación, preservación y búsqueda en línea. Se trata de representar los materiales textuales en formato electrónico de tal forma que permita, a los investigadores de cualquier disciplina, intercambiar y reutilizar los recursos independientemente del hardware y software que utilicen.

DISCIPLINA/ÁMBITO DE APLICACIÓN: se aplica especialmente en el campo de la lengua, la literatura, las humanidades y las ciencias sociales. Se utiliza para describir textos literarios y lingüísticos en bibliotecas, museos y editoriales. Sin embargo, puede aplicarse para la estructuración de información electrónica de todo tipo, incluso imágenes.

CORRIENTE TEÓRICA A LA QUE RESPONDEN: proviene de la comunidad de estudio de la lingüística.

SIMPLICIDAD/COMPLEJIDAD: la complejidad está asociada a la cantidad de información que se quiera incluir y el nivel de normalización que se pretenda. Cada proyecto que utilice TEI debe tomar sus propias decisiones respecto a qué partes del esquema utilizar y el nivel de

complejidad o sofisticación a emplear. Requiere del conocimiento del modelo y de las reglas bibliotecarias.

INTEROPERABILIDAD: ha sido concebido para intercambio, con independencia del software y hardware.

EXTENSIBILIDAD: permite un marcado semántico de textos muy detallado.

FLEXIBILIDAD: en los instructivos se sugiere que el usuario adecue TEI a sus necesidades, para obtener la máxima prestación.

ESTRUCTURA: se comenzó en 1994 con la definición de 600 elementos en SGML, luego surgieron las versiones TEI P3 (1998) -más operativa y modular-, TEI P4 (2001) en XML y TEI Lite, orientada a bibliotecas digitales; la versión actual es TEI P5 versión 1.5.0 (2009). Utiliza etiquetas basadas en palabras y no numéricas. Uno de los componentes más significativos de TEI es la cabecera TEI (*TEI Header*), que permite una descripción bibliográfica detallada para cada texto codificado TEI como parte de su estructuración original.

Características generales de los elementos:

- a) Opción: existen elementos obligatorios, opcionales y recomendados.
- a)b) Repetición: posee elementos repetibles y no repetibles.
- a)c) Subcampos: no se utilizan subcampos, sí atributos dentro de los campos.
- a)d) Control: se sugiere establecer control para ciertos elementos (p. ej.: Series).

Los elementos de TEI pueden albergar diferentes tipos de contenidos:

- **Texto libre (*free prose*):** los diferentes párrafos se marcan con el elemento *<p>*.
- **Elementos de conjunto (*grouping elements*):** elementos cuyos nombres llevan el sufijo *Stmt*, usualmente agrupan una serie de elementos especializados. En el caso de los elementos bibliográficos corresponden a las áreas ISBD.
- **Declaraciones (*declarations*):** elementos con el sufijo *Decl*, albergan datos específicos sobre la codificación del texto electrónico, usualmente registrados en forma codificada.
- **Descripciones (*descriptions*):** elementos con el sufijo *Desc* contienen una descripción organizada bajo encabezamientos específicos de acuerdo a los subelementos sugeridos.

A continuación se listan los principales elementos de TEI Header de la versión TEI P5 (2009).

Los elementos de la cabecera TEI están estructurados en 4 secciones:

- **descripción de archivo (*fileDesc*)**, que contiene la descripción bibliográfica completa del texto similar a AACR2R e ISBD;
- **descripción de codificación (*encodingDesc*)**, con elementos sobre características específicas del archivo electrónico, relaciona el texto electrónico con la fuente o fuentes de las cuales deriva;
- **descripción de perfil (*profileDesc*)**, con información similar a los puntos de acceso secundarios de AACR2R, datos de lengua, clasificación, materias y,
- **descripción de revisión (*revisionDesc*)**, que contiene un registro de cada cambio que se le ha hecho al texto y la fecha en que ha sido cambiado.

TEI posee un sistema de clases que permite adaptar el esquema a las necesidades del usuario. Por ejemplo: en la cabecera, para refinar la descripción se pueden aplicar elementos agrupados en las siguientes clases:

- **model.applicationLike**: agrupa elementos usados para registrar un tipo específico de información acerca de un documento en la cabecera. Por ejemplo: model.listLike (tipo lista); model.biblLike (bibliográfica).
- **model.catDescPart**: agrupa elementos sobre las categorías de descripción de la cabecera TEI.
- **model.editorialDeclPart**: agrupa elementos que pueden usarse dentro del elemento *<editorialDecl>*.
- **model.encodingPart**: agrupa elementos que pueden ser utilizados dentro del elemento *<encodingDesc>*.
- **model.profileDescPart**: agrupa elementos que pueden usarse en el elemento *<profileDesc>*.
- **model.headerPart**: agrupa elementos de alto nivel que pueden aparecer más de una vez en la cabecera TEI.
- **model.sourceDescPart**: agrupa elementos que pueden ser utilizados en el elemento *<sourceDesc>*.
- **model.textDescPart**: agrupa elementos que sirven para categorizar un texto, por ejemplo, en términos de sus parámetros situacionales.

El elemento superior es la etiqueta de cabecera TEI *<teiHeader>*

El elemento **descripción de archivo** *<fileDesc>* es obligatorio. Contiene a su vez los siguientes elementos:

- **Título** (*titleStmt*) (title statement)

Título (*title*): contiene el título completo de una obra de cualquier tipo.

Autor (*author*): contiene el nombre del (los) autor(es), personales o corporativos.

Patrocinante (*sponsor*)

Financiador (*funder*) (funding body)

Principal (*principal*) (principal researcher): nombre del investigador principal responsable de la creación del texto electrónico.

Mención de responsabilidad (*respStmt*) (statement of responsibility): se registra aquí una mención de responsabilidad cuando otros elementos no son suficientes o no se aplican.

Responsabilidad (*resp*) (responsibility): frase que describe la naturaleza de la responsabilidad intelectual de una persona.

Nombre (*name*) (name, proper noun): contiene un nombre propio.

- **Edición** (*editionStmt*) (edition statement)

Edición (*edition*)

Mención de responsabilidad (*respStmt*)

Nombre (*name*)

Responsabilidad (*resp*)

- **Extensión** (*extent*)

- **Publicación** (*publicationStmt*) (publication statement)

Editorial (*publisher*)

Distribuidor (*distributor*)

Autoridad de publicación (*authority*) (release authority)

Uno de estos tres elementos mencionados estará presente en la descripción y contendrá alguno o todos de los siguientes elementos:

Lugar de publicación (*pubPlace*)

Dirección postal (*address*)

Número de identificación (*idno*) (identifying number) se registra el número estándar o no estándar utilizado por el editor para identificar el ítem.

Tipo (*type*): caracteriza el número, por ej. ISBN.

Disponibilidad (*availability*) información sobre condiciones de uso de un texto, por ejemplo: alguna restricción, copyright, etc.

Estado (*status*): información sobre la disponibilidad actual del texto.

Fecha (*date*)

- **Serie** (*seriesStmt*) (series statement)

Título (*title*)

Número de identificación (*idno*) (identifying number)

Mención de responsabilidad (*respStmt*)

Responsabilidad (*resp*)

Nombre (*name*)

- **Notas** (*notesStmt*) (notes statement)

Nota (*note*): se utiliza para cada nota individual

- **Fuente** (*sourceDesc*) (**source description**): es un elemento obligatorio que detalla la fuente o fuentes de donde deriva el texto electrónico. Puede ser un texto impreso o manuscrito, otro archivo electrónico, una grabación de audio o video, o una combinación de los nombrados. Puede no tener fuente, es decir, ser un documento

original creado electrónicamente. Puede contener una declaración simple o utilizar clases y elementos específicos provistos por TEI para una información detallada.

El elemento **descripción de codificación** *<encodingDesc>* especifica los métodos y principios editoriales que se utilizaron en la transcripción del texto. El uso es altamente recomendado.

Se puede incluir información en texto libre con la codificación *<p>*, o contener elementos específicos tomados de la clase *model.encodingPart*. Por defecto dicha clase especifica los siguientes elementos:

- **Descripción de proyecto** (*projectDesc*): describe en detalle el propósito para el cual fue codificado el archivo electrónico y otra información relevante concerniente al proceso de realización.
- **Declaración de muestreo** (*samplingDecl*): describe la metodología utilizada para la selección de los textos a incluir en el recurso.
- **Declaración de prácticas editoriales** (*editorialDecl*): detalla los principios y prácticas involucrados durante la edición de un texto electrónico. Puede ser una aclaración en texto libre o ampliarse según los elementos de la clase *model.editorialDeclPart*.
- **Declaración de etiquetado** (*tagsDecl*): provee información detallada acerca del etiquetado aplicado a un documento.
- **Declaración del sistema de referencias** (*refsDecl*): especifica como funciona el sistema de referencias construido en la codificación del texto.
- **Declaración de clasificación** (*classDecl*): contiene elementos que definen los códigos de clasificación utilizados en cualquier parte del texto.
- **Información de aplicación** (*appInfo*): registra información sobre la aplicación que se utilizó en la edición del archivo TEI.

La descripción de perfil (*profileDesc*) provee información que caracteriza varios aspectos descriptivos no bibliográficos de un texto. Es de uso opcional e incluye tres elementos opcionales:

- **Creación** (*creation*): contiene información acerca del origen de un texto.

- **Uso de idioma (*langUsage*):** describe los idiomas, subidiomas, registros y dialectos representados dentro del texto.
- **Clasificación de texto (*textClass*):** agrupa información que describe la naturaleza o los tópicos de un texto, en términos de un esquema de clasificación estándar, tesauro, etc.

Dentro de (*textClass*) se proveen los siguientes elementos:

Palabras clave (*keywords*)

Esquema (*scheme*), identifica el vocabulario controlado utilizado.

Código de clasificación (*classCode*)

Esquema (*scheme*), identifica el código de clasificación utilizado.

Referencia de categoría (*catRef*): especifica una o más categorías definidas dentro de una taxonomía o tipología de texto.

Otros tres elementos que pueden aparecer dentro de (*profileDesc*) son:

- **Descripción de texto (*textDesc*):** provee una descripción del texto en términos de sus parámetros situacionales.
- **Descripción de participación (*particDesc*):** se registran locutores, voces u otros participantes en una interacción lingüística.
- **Descripción de contexto (*settingDesc*):** describe el marco dentro del cual una interacción lingüística tiene lugar.

El elemento **descripción de revisión (*revisionDesc*)** registra los cambios que va teniendo en texto. Es de uso opcional pero altamente recomendable. Los cambios pueden registrarse como una secuencia de elementos (*change*) que contengan una breve descripción. Los atributos de fecha *<date>* y quién *<who>* pueden utilizarse para identificar cuándo se realizó el cambio y el agente responsable de la modificación.

Los elementos básicos son:

descripción de revisión (*revisionDesc*) reúne el historial de cambios

cambio (*change*): registra un cambio particular o corrección realizada a una versión del texto electrónico.

Como puede verse TEI es un esquema muy completo y permite una gran extensibilidad.

CONTENIDO: es un estándar de estructura que en ciertos elementos prescribe el contenido, está implícito el uso de AACR2 o ISBD. Permite además información en texto libre.

NIVEL DE CATALOGACIÓN PERMITIDA: formato rico, información descriptiva altamente estructurada, se ubica en un Nivel 3 (Dempsey & Heery, 1998). Permite realizar adaptaciones, según las necesidades, a través de su estructuración en clases, que, a su vez, agrupan elementos que permiten incluir más detalle en otros elementos.

ACTUALIZACIÓN: la versión actual de esquema es TEI P5 de noviembre de 2009.

CONTRIBUCIÓN A LA NORMALIZACIÓN INTERNACIONAL: se puede considerar que actualmente es un estándar de hecho.

TEI hace referencia a los siguientes estándares:

- ISO 646: Information technology -- ISO 7-bit coded character set for information interchange
- ISO 10646: Information technology -- Universal Multiple-Octet Coded Character Set (UCS)
- Unicode

SINTAXIS: se codifica en SGML y XML.

Ejemplo de un registro TEI en sintaxis XML:

```
<tei header>
<filedesc>
  <titlestmt>
    <title>Arithmetic /</title>
    <author>Sandburg, Carl, 1878-1967.</ author>
    <respStmt>
      <resp> illustrated as an anamorphic adventure by </resp>
      <name>Ted Rand</name>
    </respStmt>
  </titlestmt>
  <publicationstmt>
    <publisher>Harcourt Brace Jovanovich,</publisher>
    <pubPlace> San Diego</pubPlace>
    <idno type='ISBN'>0152038655</idno>
    <date>c1993</date>
  </publicationstmt>
  <sourcedesc>
    <p> </p>
  </sourcedesc>
</filedesc>
<profiledesc>
```

```
<langUsage>
<language id="en">english</language>
</langUsage>
<textClass>
<keywords scheme="lsh">
<list>
<item>Arithmetic</item>
<item>Children's poetry, American.</item>
<item>Arithmetic</item>
<item>American poetry.</item>
<item>Visual perception.</item >
</list>
</keywords>
</textClass>
</profiledesc>

<encodingdesc>
</encodingdesc>

<revisiondesc>
</revisiondesc>

</tei header>
```

RESULTADOS OBTENIDOS HASTA EL PRESENTE: a diciembre de 2009 existen 125 proyectos registrados en la página Web de TEI.

LENGUAJES DE MARCADO

Los esquemas de metadatos analizados en este trabajo se codifican con los denominados lenguajes de marcas, como SGML, HTML, XML o RDF. En este apartado se describen las características principales de éstos últimos.

SGML

El SGML, *Standard Generalized Markup Language*, está definido por la norma ISO 8879 aprobada en 1986 y constituye un modelo abierto para el diseño de documentos electrónicos. Es un metalenguaje o marco normativo que permite definir diferentes lenguajes de marcación de textos compatibles entre sí. Aplica el concepto de documento distribuido, es decir, la posibilidad de generar un documento distribuido en distintos archivos que se pueden almacenar en diferentes servidores. Se basa en un marcado genérico del documento para identificar sus partes constitutivas de manera jerárquica. La importancia del marcado textual reside en que asegura que la indización no se separe del texto del documento, lo que asegura la persistencia en el tiempo y en la arquitectura distribuida de la Web. SGML ha sido la base de modelos de metadatos como TEI y desarrollos posteriores como HTML y XML. Como es muy complejo y difícil de manejar, aplicar y visualizar en la Web tiene una escasa aplicación.

HTML

El HTML, Lenguaje de Marcación de Hipertexto (*Hiper Text Mark-up Language*), es un subconjunto simplificado o aplicación del SGML. Provee una forma simple de estructurar documentos de hipertexto y de realizar referencias desde un documento a otro. En esencia, los documentos HTML son documentos de texto plano (ASCII). Sirve para marcado estructural (elementos del contenido) y para presentación

La estructura básica de un archivo HTML es la siguiente:

```
<HTML>
```

```
  <HEAD>
```

```
    <TITLE>
```

```
  </TITLE>
```

```
  </HEAD>
```

```
  <BODY>
```

```
  </BODY>
```

```
</HTML>
```

Dentro de esta estructura, el documento se divide en una *cabecera* y un *cuerpo*. La cabecera, etiqueta <HEAD>, contiene información sobre la página. Los datos aquí incluidos no se ven con la visualización del documento a través del navegador. Dentro de la cabecera se encuentra la etiqueta <TITLE> que brinda información acerca del nombre del documento. Este nombre aparece en la barra de estado del navegador. El cuerpo del documento, etiqueta <BODY>, contiene toda la información que el usuario visualiza a través del navegador.

Dentro de la etiqueta <HEAD>, el HTML permite incorporar metadatos mediante el uso de la etiqueta META. Cada etiqueta META tiene mínimamente dos atributos:

- NAME: indica el tipo de propiedad.
- CONTENT: indica el valor de la propiedad.

Ejemplo:

```
<META NAME="Author" CONTENT="Umberto Eco">
```

También puede incluir otros atributos tales como:

- LANG: indica la lengua.
- SCHEME: indica el esquema utilizado.

Ejemplos:

```
<META NAME="Author" LANG="es" CONTENT="Umberto Eco">
```

```
<META NAME="identifier" SCHEME="ISSN" CONTENT="1056-3456">
```

Como señala Méndez Rodríguez (2002), es un lenguaje diseñado especialmente para soportar una presentación consistente, y se ha convertido predominantemente en un lenguaje de formato más que en un código semántico, por lo que actualmente proporciona muy poca ayuda para la recuperación de información. Tiene un conjunto de etiquetas prefijado y finito, no se pueden ampliar o crear nuevas etiquetas más significativas, sólo en Consorcio Web es posible hacerlo. La última versión es HTML 4.01 de 1999.

La complejidad de SGML y las debilidades de HTML en materia de recuperación de la información han propiciado el surgimiento de XML. El sucesor es XHTML, *Extensible HyperText Markup Language*, que sucede y extiende HTML reformulado en XML. Actualmente el W3C está trabajando en XHTML2.

XML

El XML, *eXtensible Mark-up Language*, es una aplicación del SGML desarrollada por el *World Wide Web Consortium* (W3C) como un formato universal para documentos estructurados y datos en la Web. La versión 1.0 de XML surge en 1998.

Mientras que HTML es una sintaxis no extensible, con etiquetas fijas; XML, al ser extensible, permite definir localmente las etiquetas que se quieren crear y permite construir productos basados en el intercambio estructurado de datos entre aplicaciones y admite en sí mismo la creación de metadatos. Una ventaja del XML es su gran funcionalidad en la búsqueda. Es un estándar de intercambio de documentos en la Web, que mejora la recuperación ya que permite el marcado según el contenido semántico, a través de elementos y atributos.

Como señalan De la Rosa y Senso (1999), XML simplifica SGML, eliminando parte de sus operaciones sintácticas pero dotando a la información de una estructura, XML permite la creación de etiquetas propias dentro de cada documento, lo que añade precisión y detalle, y es válido para la representación digital de documentos de cualquier tipo y con cualquier extensión. Todos los documentos XML deben comenzar con una etiqueta `<?XML....?>` que indica que el tipo de documento es XML. Luego pueden tener varios atributos, como por ejemplo la versión.

Ejemplo de un archivo XML (De la Rosa y Senso, 1999, p. 493.):

```
<?XML version="1.0"? >
<DOCUMENTO>
  <libro fechalibro="1975">

    <!__ La primera edición en 1970__>
    <título>Organización de las bases de datos</título>
    <autor><apellido>Martin</apellido></autor>
    <autor><nombre>James</nombre></autor>
    <traductor><apellido>Marco</apellido></traductor>
    <traductor><nombre>Alfredo di</nombre></traductor>

  </libro>
</DOCUMENTO>
```

Las perspectivas del XML para estructurar, intercambiar y recuperar información electrónica son muy promisorias, en tanto se generalice su aplicación en la Web.

RDF

RDF, *Resource Description Framework*, es una iniciativa del W3C que surge como primer borrador en el año 1997. En 1999 se publica como recomendación del W3C "*RDF Model and Syntax Specification*". Las últimas actualizaciones se realizaron en 2004. RDF es la base para el actual proyecto de Web Semántica de W3C.

Según la recomendación del W3 Consortium (1999), RDF, o Infraestructura para la Descripción de Recursos, es una base para procesar metadatos que proporciona interoperabilidad entre aplicaciones que intercambian información legible por máquina en la Web. RDF puede utilizarse en distintas áreas de aplicación, por ejemplo: en *recuperación de recursos*, para proporcionar mejores prestaciones a los motores de búsqueda; en *catalogación*, para describir el contenido y las relaciones de contenido disponibles en un sitio Web, una página Web, o una biblioteca digital particular; por los *agentes de software inteligentes*, para facilitar el intercambio y para compartir conocimiento; en la *calificación de contenido*; en la descripción de *colecciones de páginas* que representan un "documento" lógico individual; para describir los *derechos de propiedad intelectual* de las páginas Web; y para expresar las *preferencias de privacidad* de un usuario, así como las *políticas de privacidad* de un sitio Web. RDF, junto con las *firmas digitales*, será la clave para construir el "Web de confianza" para el comercio electrónico, la cooperación y otras aplicaciones.

RDF y XML son complementarios: RDF es un modelo de metadatos y sólo dirige por referencia muchos de los aspectos de codificación que requiere el almacenamiento y la transferencia de archivos (tales como internacionalización, conjuntos de caracteres, etc.). Para estos aspectos, RDF cuenta con el soporte de XML. Es importante también entender que esta sintaxis XML es sólo una sintaxis posible para RDF y que pueden surgir formas alternativas para representar el mismo modelo de datos RDF.

El fundamento o base de RDF es un modelo para representar propiedades designadas y valores de propiedades. El modelo de datos básico consiste en tres tipos de objetos:

Recursos: todas las cosas descritas por expresiones RDF se denominan *recursos*. Un recurso puede ser una página Web completa, una parte de una página Web; por ejemplo: un elemento HTML o XML específico dentro del documento fuente. Un recurso puede ser también una colección completa de páginas, por ejemplo: un sitio Web completo.

Propiedades: una *propiedad* es un aspecto específico, característica, atributo o relación utilizado para describir un recurso. Cada propiedad tiene un significado específico, define sus valores permitidos, los tipos de recursos que puede describir y sus relaciones con otras propiedades.

Sentencias: declaraciones o enunciados, un recurso específico junto con una propiedad denominada, más el valor de dicha propiedad para ese recurso es una *sentencia RDF* (RDF statement). Estas tres partes individuales de una sentencia se denominan, respectivamente, sujeto, predicado y objeto. El objeto de una sentencia (es decir, el valor de la propiedad) puede ser otro recurso o puede ser un literal.

RDF permite intercambiar metadatos de recursos web entre distintas aplicaciones. Es decir que permite la codificación e intercambio de metadatos generados, bajo cualquier formato de descripción de metadatos (DC, PICS, IAFA, etc.). El modelo es independiente de la sintaxis. A continuación se muestra un ejemplo simple de RDF tomado de la Web del W3C (<http://www.w3.org/Metadata/Activity>)

```
<RDF xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
<Description about="http://www.w3.org/Press/99Folio.pdf">
<dc:title>The W3C Folio 1999</dc:title>
<dc:creator>W3C Communications Team</dc:creator>
<dc:date>1999-03-10</dc:date>
<dc:subject>Web development, World Wide Web
Consortium, Interoperability of the Web</dc:subject>
</Description>
</RDF>
```

COMPARACIÓN DE LOS ESQUEMAS DE METADATOS

A continuación se presenta el resultado de la comparación de los metadatos analizados, según los criterios propuestos.

CRITERIO	MARC21	DC	TEI	MODS
Año de creación	1960 1992, campo 856 2002, MARCXML	1995	1987 1994, TEI Guidelines 2000, TEI Consortium	2002
Organismo creador	LC	OCLC NCSA (National Center for Supercomputing Applications)	Association of Computer in the Humanities, Association for Computational Linguistics y Association of Literary and Linguistic Computing	LC
Mantenimiento	LC	DC Metadata Initiative	TEI Consortium	LC
Documentación	Manuales en línea, varios idiomas, lista de discusión.	Manuales en línea.	Manuales en línea, tutoriales, software, grupos de discusión, wiki	Manuales en línea y lista de discusión.
Misión/ propósito	Inicialmente para documentos impresos, fue incorporando otros formatos.	Ad-hoc para recursos electrónicos, simplicidad, interoperabilidad.	Para textos electrónicos literarios y lingüísticos, intercambio de recursos.	Para recursos electrónicos, subconjunto de MARC, totalmente compatible con

				MARC.
Disciplina de Aplicación (actualmente)	Todas las disciplinas. Bibliotecas.	Todas las disciplinas. Bibliotecas, museos, educación, gobierno, archivos, etc.	Se aplica en el campo de la lengua, la literatura, las ciencias sociales. Bibliotecas, museos, editoriales.	Especialmente en bibliotecas.
Corriente teórica	De la bibliotecología	Interacción bibliotecología e informática	De la lingüística	De la bibliotecología
Actualización	2008	Versión 1.1 2008	TEI P5 2009	Versión 3.3 2008
Normalización	Estándar de hecho. ISO 2709 ANSI Z39.2	ISO 15863-2003 ANSI/NISO Z39.85-2007	Estándar de hecho.	Proyecto registrado en NISO, 2004

En cuanto a su **origen**, MARC, MODS y Dublin Core surgen de instituciones ligadas a la bibliotecología como son LC y OCLC. TEI proviene del ámbito de la lingüística, pero para diseñar la descripción de archivo se interactuó con bibliotecarios, entre otros profesionales. MARC es el formato más antiguo (1960), TEI se inicia en 1987 y en 1994 se publican las primeras recomendaciones, posteriormente aparece Dublin Core (1995) y finalmente MODS nace en 2002.

Se puede afirmar que se ha partido de formatos sumamente complejos como MARC y TEI para luego pasar a un formato más simple como DC, y, finalmente, surge un formato intermedio como lo es MODS. Todos provienen de instituciones altamente reconocidas, está previsto su mantenimiento y desarrollo y cuentan con documentación detallada en línea.

En cuanto al **propósito**, MARC ha sido creado para catalogación de documentos impresos y se fue extendiendo para los nuevos soportes, para el caso de recursos electrónicos se incorporaron nuevos campos, pero siempre respondiendo al mismo modelo. El último avance ha sido la codificación en XML para adaptarlo al ambiente Web.

Dublin Core, TEI y MODS han sido creados *ad hoc* para descripción de recursos electrónicos. Dublin Core y TEI priorizan la facilidad de recuperación a través de motores de búsqueda y el intercambio de información. MODS apunta a una descripción más simple y flexible que MARC21, pero a la vez totalmente compatible con él.

ESTRUCTURA. CONTENIDO. SINTAXIS

El formato **MARC** se estructura en campos y subcampos, algunos obligatorios y otros opcionales. Utiliza etiquetas numéricas, lo que lo independiza del lenguaje. Es un estándar de estructura, donde el contenido está regido por normas externas al formato. Se codifica en HTML y XML.

Dublin Core se estructura en elementos básicos y calificadores, utiliza etiquetas basadas en palabras. Todos los elementos son opcionales y repetibles. Es un estándar de estructura, que no está regido por reglas de contenido, sino que se sugiere la normativa a aplicar. Se codifica en HTML, XML y RDF.

MODS consta de elementos principales, subelementos y atributos que pueden ser usados en uno o más elementos. Utiliza etiquetas basadas en palabras. Todos los elementos y atributos son opcionales. Es un estándar de estructura, el contenido de la mayoría de los elementos se anota de acuerdo a normativas externas de manera análoga a MARC. Se codifica en XML.

TEI se estructura en elementos y atributos, algunos obligatorios, otros opcionales o recomendados. Utiliza etiquetas basada en palabras. Es un estándar de estructura que en ciertos elementos prescribe el contenido. Se codifica en SGML y XML.

Como puede apreciarse, todos son estándares de estructura, utilizan elementos principales que pueden refinarse de acuerdo a la necesidad del usuario y pueden codificarse en XML. TEI en ciertos elementos prescribe el contenido. Dublin Core da un paso más en cuanto a la sintaxis y puede ser incluido en el marco RFD. Excepto MARC, se utilizan etiquetas en palabras en inglés.

NIVEL DE CATALOGACIÓN PERMITIDA

	MARC21	DC	TEI	MODS
Nivel 1				
Nivel 2				
Nivel 3				

En cuanto a esta característica, MARC y TEI, según la clasificación de Dempsey y Heery (1998), se ubican en un nivel 3, son formatos ricos altamente estructurados. Dublin Core se ubica en la banda 2, con una descripción sencilla. MODS estaría moviéndose entre los niveles 2 y 3 dependiendo si se incluyen o no todos los subelementos y atributos.

SIMPLICIDAD. EXTENSIBILIDAD. FLEXIBILIDAD

En cuanto a la **complejidad** de creación de los registros, DC es sin duda el más sencillo, está pensado para que el propio autor pueda asignar los metadatos en el momento de creación del archivo digital. De todas maneras, se complejiza al aumentar el nivel de detalle de la descripción a través del empleo de los calificadores.

MARC y TEI son complejos por el nivel de detalle, la cantidad y la alta estructuración de los elementos, y están asociados a las normativas bibliotecarias para registro del contenido de la descripción. MARC de manera obligatoria y TEI, si bien deja opción, sugieren la aplicación de las AACR2 e ISBD en varios de sus elementos.

MODS estaría en un intermedio, al tener menos elementos, pero se requiere el conocimiento de las normativas bibliotecarias, porque los elementos se equiparan a MARC. La simplicidad relativa, la flexibilidad y la compatibilidad con MARC, dan a MODS muchas posibilidades para las bibliotecas.

Todos son **flexibles** y **extensibles** a la hora de seleccionar los elementos a incluir, el usuario decide el **nivel de detalle de la descripción**, aunque MARC tiene un nivel menor de flexibilidad.

INTEROPERABILIDAD

La interoperabilidad es un criterio con fuerte presencia en todos ellos, actualmente todos se pueden codificar en XML. DC además es compatible con el modelo RDF.

METADATOS Y RECUPERACIÓN DE LA INFORMACIÓN

Como señala Méndez Rodríguez (2002), las herramientas de búsqueda y recuperación de información en Internet se fundamentan en la automatización de las tareas tradicionales de organización del conocimiento; se puede decir que un software clasifica y/o indiza automáticamente los datos de Internet, sistematizando de alguna manera la información digital existente. El tratamiento automático se presenta como la única opción real frente a la inmensa cantidad de datos, sin embargo, estas aplicaciones en general no identifican las características de un documento de la forma detallada y precisa en que lo haría un indizador humano.

Las herramientas generalizadas disponibles son: buscadores o motores de búsqueda, metabuscadores e índices o directorios.

1- Buscadores o motores de búsqueda:

Son sistemas poscoordinados de recuperación de información y se basan en la búsqueda por palabras clave. Un buscador es una base de datos que contiene páginas Web indexadas, cuya función es recuperar y mostrar una lista de resultados que satisfacen una determinada búsqueda. Generalmente proveen un área de entrada (formulario) donde se escribe la expresión de búsqueda. Algunos ejemplos de buscadores son: Altavista, Google, Excite, Hotbot, Lycos, Infoseek.

Existen básicamente dos formas por medio de las cuales los buscadores indexan páginas Web: una de ellas es a través del **registro de los usuarios**, que crean sus páginas y las incorporan a los distintos buscadores por medio de las opciones: **sugerir una página** (*suggest a page*) o **agregar un sitio** (*add a page*). Este es un método simple y generalmente contiene un formulario con campos para el título, el *Uniform Resource Locator* (URL), la persona de contacto, ubicación geográfica, comentario, etc. Luego, los catalogadores del servicio seleccionan las categorías y agregan el sitio a la base de datos. Por último, se le envía un mail a la persona informándola sobre la inclusión o no del sitio. Este método es sencillo y fácil de usar,

pero carece de normalización. La otra forma es a través de **software automático** [denominados **arañas** (*spiders*), **agentes** o **robots**] que capturan páginas Web y crean índices de algunos términos del texto completo o parte del documento y los almacenan en una base de datos. El software realiza el seguimiento automático de los enlaces de cada página para indexar nuevos recursos o para actualizar los existentes. Este método también carece de normalización.

Para la búsqueda, los buscadores utilizan expresiones que son combinaciones de **palabras clave** (*keywords*) que el usuario introduce en el campo de texto. La combinación de estas palabras clave es enviada al motor de búsqueda, que se encarga de rastrear en la base de datos documentos que las contengan.

Existen básicamente tres formas de armar una expresión de búsqueda:

1. **Palabras clave** (*keyword*)
2. **Preguntas en lenguaje natural** (*natural language search queries*)

Se rastrea cada término introducido, a excepción de las palabras vacías (*stop words*) tales como: y (*and*), con (*with*).

Con el lenguaje natural el usuario puede introducir una expresión en forma de pregunta tal como: ¿Dónde puedo encontrar información acerca de? ("*Where can I find information on...?*"), o ¿Qué es...? ("*What is the...?*"), o ¿Dónde hay...? ("*Where is...?*").

Los resultados de la interrogación en lenguaje natural varían de un tópico a otro.

3. **Conceptos** (*concept searching, fuzzy searching, query by theme*)

A veces el usuario no utiliza las palabras clave precisas para encontrar toda la información que necesita, razón por la cual las búsquedas por conceptos permiten ampliar el espectro a través de ciertos tópicos relacionados.

La mayoría de los buscadores permite operaciones booleanas y/o relacionales, truncamiento y adyacencia.

Existen dos formas de realizar búsquedas:

- a través de una expresión de búsqueda construida por el usuario,
- a través de opciones que ofrece el buscador donde se limita la interrogación a una parte específica de la base de datos.

Cada buscador tiene su propio criterio para clasificar los resultados. Dichos criterios se basan en métodos de análisis estadístico y/o ponderación, y presentan los resultados organizados según un ranking de relevancia. A continuación, se citan algunos criterios:

- De acuerdo a la **cantidad de enlaces** que tienen otros sitios Web a la página que contiene el término de búsqueda ingresado por el usuario. Es decir, la cantidad de veces que una página es citada. Este criterio, en cierta forma, excluye aquellos recursos nuevos y de muy buena calidad que en un principio son poco citados por el sólo hecho de ser recientes y aún pocos conocidos. En estos casos predomina la cantidad de enlaces en lugar de la calidad.
- **Frecuencia de aparición de las palabras clave** y frases en el documento recuperado.

- **Ubicación de las palabras clave**, generalmente los documentos que las contienen en el título tienen mayor relevancia que los que las incluyen en el resto del texto o en el URL.

Ventajas:

- Son fáciles de usar.
- Son rápidos.
- Son automáticos.

Desventajas:

- Están preparados para leer texto en ASCII, por lo tanto no pueden detectar el texto comprendido dentro de gráficos. Si éstos no están acompañados por un texto, o no contienen metadatos, resultan inaccesibles.
- Baja eficacia, muchas veces se recupera demasiada información no pertinente. Una palabra clave puede aparecer en un texto que no tiene ninguna relación con la información que interesa.
- En general, no permiten la búsqueda por varios campos específicos, como por ejemplo: autor, título, etc.
- No manejan sinónimos ni homónimos.
- No recuperan información de sitios a los cuales se accede con clave, ni al contenido de bases de datos. A esta parte de la Web comúnmente se la conoce como la “Web invisible”.

2.- Metabuscadores (*metasearch engines*)

En la actualidad, ningún buscador puede indexar toda la World Wide Web, cuando no se encuentra material a través de uno hay que recurrir a otros. Los metabuscadores ofrecen una solución a este problema utilizando una única interface de búsqueda para rastrear en múltiples buscadores. No tienen una base de datos propia sino que buscan simultáneamente, recopilan y organizan los resultados que arrojan otros. Algunos ofrecen un formulario único a través del cual se envía la expresión de búsqueda a cada uno de los buscadores incluidos, mientras que otros simplemente ofrecen un campo de texto para realizar el rastreo en uno de ellos en particular. Ejemplos de metabuscadores son: Dogpile, Ixquick, Metacrawler.

La eficacia en la recuperación de un metabuscadore está supeditada al funcionamiento de cada buscador individual con el que trabaja. En general, tienen las mismas ventajas y desventajas de los buscadores. Algunas ventajas con respecto a estos últimos son que no muestran todos los resultados de cada buscador sino que hace una selección de los más relevantes, suelen ser muy rápidos y simplifican la búsqueda al usuario evitándole repetirla en varios buscadores.

3.- Directorios (*Web directories*)

Son también conocidos como catálogos Web (*Web catalogs*). Se estructuran en forma de listas de categorías en un orden jerárquico, incluyen una limitada cantidad de sitios y se basan en la clasificación humana. Se trata de sistemas precoordinados de recuperación de información. Se basan en criterios semánticos de organización del conocimiento. Algunos ejemplos de directorios son: Google Directory, The Environment Directory, Yahoo! Directory, LookSmart, Open Directory Project.

Ventajas:

Según señala Callery (1997):

- Los resultados de la búsqueda son más relevantes.
- El usuario no necesita conocer todos los sinónimos de un término de búsqueda.
- Ofrecen la posibilidad de descubrir ítems relacionados incluidos en el mismo tópico ya que las categorías vinculadas actúan como referencias de "*véase además*" guiando al usuario.

Desventajas:

- Tienen dos limitaciones:
 - a) cantidad de información, abarcan en forma limitada cierta cantidad de enlaces, el trabajo es realizado por personas y esto no permite asumir el crecimiento incesante de la Web.
 - b) falta de estandarización en cuanto al nombre de las categorías entre los distintos directorios.
- Insumen mayor tiempo que una búsqueda a través de formulario de interrogación.

Las desventajas reseñadas en los sistemas de uso generalizado de recuperación de información han impulsado el desarrollo de herramientas que intentan mejorar los resultados, por ello, algunos de los sistemas se basan en metadatos y/o estructuración de la información.

La mayoría de los motores de búsqueda no registran metadatos, y en el caso de que estén preparados para reconocerlos, la mayor parte de los autores/editores de las páginas Web no utilizan un modelo de metadatos de forma generalizada que permita insertar en la cabecera del documento esta información.

Los motores de búsqueda, si es que soportan metadatos, operan con los metadatos de banda 1 según la clasificación de Demspey y Heery (1998), con lo cual no existe una verdadera estructuración previa de la información, sino que los datos sobre los datos se extraen de manera

automática y sin diferenciación. Algunos buscadores pueden basar su recuperación en metadatos como DC, tal es el caso de Ultraseek y MetaStar.

Por otra parte, ha surgido el concepto de Web semántica. Según lo explican Berners Lee, Hendler y Lassila (2001), la Web semántica brindará una estructura a los contenidos significativos de las páginas Web, creando un ambiente donde el software de rastreo pueda llevar a cabo rápidamente tareas sofisticadas para los usuarios. Para que la Web semántica funcione, las computadoras tienen que tener acceso a colecciones estructuradas de información.

Como señala Méndez Rodríguez (2002), los metadatos serán un método efectivo para la descripción de información en la Web, si además de convertirse en estándares nacionales o internacionales, se aplican consistentemente por los creadores de los recursos y las aplicaciones de búsqueda se basan en ellos.

CONCLUSIONES

- La catalogación es una representación que consiste en la especificación de una serie de características del documento, y es el primer puente que media entre los documentos y el usuario. Estas representaciones documentales han recibido diferentes nombres a lo largo del tiempo: ficha catalográfica, registro bibliográfico, y, en la actualidad, se las denomina con el término metadatos. Acordamos con Gorman (2000) y con Pinto Molina (2002) en que, en teoría, catalogación y metadatos son equivalentes, pero como bien sintetiza Méndez Rodríguez (2002), la descripción de los recursos electrónicos en Internet, realizada manualmente por profesionales, sólo puede abarcar un universo muy pequeño; si se piensa en toda la Web, es ilusorio e impracticable, amén de excesivamente costoso. La estructuración de la información a través de la inclusión de metadatos desde el nacimiento del recurso electrónico es por el momento la alternativa viable para la descripción de contenidos en Internet.
- A nuestro juicio, los diferentes esquemas de metadatos van a coexistir, se aplicarán unos u otros de acuerdo con las necesidades, por lo que el desarrollo deberá ir por el camino de la interoperabilidad, para posibilitar el intercambio entre los diferentes esquemas. En este sentido, *Resource Description Framework* (RDF), que sólo se menciona brevemente en este trabajo, es un marco adecuado para lograr la interoperabilidad de los diversos modelos de metadatos, en el ambiente Web basado en XML.

- En el ámbito de las bibliotecas, MODS se perfila como una opción viable, ya que es compatible con MARC que es ampliamente utilizado.
- Se pasaría del catálogo único y centralizado a uno distribuido, con metadatos asociados a los documentos, registros adicionales creados por bibliotecas y centros de información, que puedan ser recuperados e interpretados por motores de búsqueda.
- Consideramos muy importante la opción de los metadatos embebidos, que nacen con el documento electrónico, como catalogación en fuente o “portada electrónica normalizada” a través de una cabecera. De todos modos, no será tarea fácil conseguir que autores y editores agreguen esta información y lo hagan de manera normalizada, pero un documento en la red sin metadatos, tendrá escasa o nula visibilidad. Tendrá que surgir algún software de asistencia para la asignación de metadatos de manera ágil.
- Como se dijo anteriormente, los metadatos serán un método efectivo para la descripción de información en la Web, si los creadores de los recursos los aplican consistentemente y los motores de búsqueda se basan en ellos.
- Los bibliotecarios tienen un campo de acción destacado en este ámbito de estudio, al tener una larga tradición en la metodología de estructuración de sistemas de información.
- La descripción de recursos electrónicos disponibles en la Web requiere de un trabajo interdisciplinario entre los especialistas autores del contenido de los documentos, en estrecha vinculación con informáticos y bibliotecarios.
- Los proyectos cooperativos jugarán un papel importante en la formalización de estándares de metadatos de reconocimiento internacional, que tienda a la interoperabilidad en el escenario de la sociedad global del conocimiento.
- Resultaría muy interesante para un estudio posterior investigar sobre las iniciativas de Topic Maps, Web Ontology Language OWL, que surgen como complemento de RDF para la construcción de la Web Semántica.

BIBLIOGRAFÍA CONSULTADA¹

- Beall, J.** (1997). Cataloging World Wide Web sites consisting mainly of links. *Journal of Internet Cataloging*, 1 (1), 83-92
- Berners Lee, T., Hendler, J. & Lassila, O.** (2001). *The Semantic Web*. Recuperado el 6 de junio de 2008, de http://www-personal.si.umich.edu/~rfrost/courses/SI110/readings/In_Out_and_Beyond/Semantic_Web.pdf
- Burnett, K., Bor Ng, K. & Park, S.** (1999). A comparison of the two traditions of metadata development. *JASIS*, 50 (13), 1209-1217
- Callery, A. & Tracy-Proulx, D.** (1997). Yahoo! cataloging the web. *Journal of Internet Cataloging.*, 1 (1), 57-64
- Caplan, P.** (1997). To Hel(sinki) and back for the Dublin Core. *Public Access Computer Systems Review*. 8 (4). Recuperado el 10 de octubre de 2009, de <http://info.lib.uh.edu/pr/v8/n4/capl8n4.html>
- Caplan, P.** (1995). You call it corn, we call it syntax-independent metadata for document-like objects. *The public access computer systems review* , 6 (4),19-23
- De la Rosa, A. & Senso, J.** (1999). XML como medio de normalización y desarrollo documental. *Revista Española de Documentación Científica*, 22 (4), 488-504.
- Dempsey, L. & Heery, R.** (1998). Metadata: a current view of practice and issues. *Journal of Documentation*, 54 (2), 145-172
- Dovey, M. J.** (1999). “Stuff” about “Stuff” the different meanings of metadata. *Vine*, 116, 6-13
- Dublin Core Metadata Initiative** (2009). Recuperado el 22 de octubre de 2009 del sitio Web oficial del DCMI, <http://dublincore.org/>
- Ede, S.** (1999). Digital indigestion – is there a cure? EN: *International Cataloguing and Bibliographic Control*. 28(1), 3-7
- Eito Brun, R.** (2008). *Lenguajes de marcas para la gestión de recursos digitales: aproximación técnica, especificaciones y referencia*. Gijón: Trea.
- Gartner, R.** (2003). *MODS: Metadata Object Description Schema*. Recuperado el 20 de septiembre de 2009, de http://www.jisc.ac.uk/uploaded_documents/tsw_03-06.pdf
- Gimeno Montoro, M.J., Barrueco Cruz, J.M. & Garcia Testal, C.** *Catalogación de recursos electrónicos accesibles en Internet: revisión de propuestas para una normativa*. FESABIB 98. VI Jornadas Españolas de Documentación. Recuperado el 3 de noviembre de 2009, de http://www.ciepi.org/fesabid98/Comunicaciones/mj_gimeno.htm

¹ Se utilizó el sistema de citas de la American Psychological Association.

- Gorman, M.** (2000). *¿Metadatos o catalogación? Un cuestionario erróneo*. En: F.F.Martínez Arellano & L. Escalona Ríos (Comps.). *Internet, metadatos y acceso a la información en bibliotecas y redes en la era electrónica* (pp.1-20). México: Universidad Nacional Autónoma.
- Hedstrom, M & Montgomery, S.** (1998). *Digital preservation needs and requirements in RLG members institutions*. Mountain View, California: Research Library Group. Recuperado el 11 de agosto de 2009, de <http://www.oclc.org/research/activities/past/rlg/digpresneeds/digpres.pdf>
- Heery, R.** (1996). Review of Metadata Formats. *Program*, 30(4), 345-373.
- Heery, R.** (1998). What is ... RDF? *Ariadne*, 14. Recuperado el 10 de octubre de 2009, de <http://www.ariadne.ac.uk/issue14/what-is/>
- Hopkins, J.** (2000). *USMARC como estructura de metadatos*. En: F.F. Martínez Arellano & L. Escalona Ríos, L. (Comps.). *Internet, metadatos y acceso a la información en bibliotecas y redes en la era electrónica* (pp.21-35). México: Universidad Nacional Autónoma.
- Hsieh-Yee, I.** (2000). Organizing Internet resources: teaching cataloging standards and beyond. *OCLC Systems & Services*, 16 (3), 130-145
- Iannella, R. & Waugh, A.** (1997). *Metadata. Enabling the Internet*. Recuperado el 23 de agosto de 2009, de <http://ifla.queenslibrary.org/documents/libraries/cataloging/metadata/ianr1.pdf>
- IFLA.** Universal Bibliographic Control and International MARC Programme. (1997) *International standard bibliographic description for electronic resources: ISBD(ER)*; revised from the ISBD(CF): International standard bibliographic description for computer files. München : K.G.Saur.
- Jul, E.** (1997). Cataloguing Internet resources: surveys and prospectus. *Bulletin of the American Society for Information Science*, 24 (1), 6-9
- Lagoze, C.** (1996). The Warwick Framework: a container architecture for diverse sets of metadata. *D-Lib Magazine*, July/August. Recuperado el 17 de diciembre de 2009, de <http://www.dlib.org/dlib/july96/lagoze/07lagoze.html>
- Lazinger, S.** (2001). *Digital preservation and metadata: history, theory, practice*. Englewood, Colorado: Libraries Unlimited
- Law, D.G.** (1996). *Electronic data archiving and acces*. Joint ICSU Press/UNESCO Expert Conference on Electronic Publishing In Science. Unesco, Paris, 19-23 February 1996
Recuperado el 12 de agosto de 2009, de <http://www.library.uiuc.edu/icsu/law.htm>
- Lyons, P.** (1997). Managing access to digital information: some basic terminology issues. *International Information & Library Review*, 29 , 207-208
- MARC Standards.** (2009). Recuperado el 16 de diciembre de 2009, del sitio Web de la Library of Congress <http://www.loc.gov/marc/>

- Martínez Arellano, F.F.** & Escalona Ríos, L. (Comps.) (2000). *Internet, metadatos y acceso a la información en bibliotecas y redes en la era electrónica*. México: Universidad Nacional Autónoma.
- Méndez Rodríguez, E.** (2002). *Metadatos y recuperación de información*. Gijón: Trea.
- Milstead, J.** & Feldman, S. (1999). Metadata: cataloging by any other name. *Online*, 23 (1), 24-31
- MODS.** Metadata Object Description Schema official Web site (2009). Recuperado el 29 de octubre de 2009, del sitio Web de la Library of Congress <http://www.loc.gov/standards/mods/>
- NISO Standard Z39.85-2007** (2007): The Dublin Core Metadata Element Set . Recuperado el 15 de diciembre de 2009, del sitio Web de NISO <http://www.niso.org/standards/resources/Z39-85-2007.pdf>
- Pinto Molina, M.,** García Marco, F.J. & Agustín Lacruz, M. (2002). *Indización y resumen de documentos digitales y multimedia*. Gijón: Trea.
- RDA:** Descripción y acceso al recurso. (2009). Recuperado el 18 de diciembre de 2009 del sitio Web del Joint Steering Comité for Development of RDA <http://www.rda-jsc.org/docs/rdabrochure-spa.pdf>
- Real Academia Española.** (1991). *Diccionario de la lengua española*. (20ª. ed.). Madrid: Real Academia Española.
- Reglas de catalogación angloamericanas** (2ª. Ed, rev. 2002, act. 2003) (2004), preparadas bajo la dirección de Joint Steering Committee for Revision of AACR.. Bogotá: Rojas Eberhard
- Sha, V.T.** (1995). Cataloguing Internet resources: the library approach. *The Electronic Library*, 13 (5), 467-476
- Sperberg-McQueen, C.M.** & Burnard, L. (Eds.) (2004). *TEI P4: Guidelines for Electronic Text Encoding and Interchange XML-compatible edition*. The TEI Consortium. Recuperado el 31 de octubre de 2009 del sitio Web oficial de TEI Consortium <http://www.tei-c.org/release/doc/tei-p4-doc/html/>
- Swick, R.** (2002). *Metadata Activity Statement*. Recuperado el 24 de enero de 2009 del sitio Web del W3 Consortium <http://www.w3.org/Metadata/Activity.html>
- Taylor, A.** (2004). *The Organization of Information*. (2nd ed.) Englewood, CO: Libraries Unlimited
- TEI Consortium,** eds. (2009). *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Version 1.5.0. Recuperado el 17 de diciembre de 2009 del sitio oficial de TEI Consortium <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>
- TEI Consortium,** eds. (2009). *The Text Encoding Initiative*. Recuperado el 25 de octubre de 2009, del sitio Web del TEI Consortium <http://www.tei-c.org/>
- Tennant, R.** (1998). 21st Century cataloguing. *Library Journal*, 123 (7), 30-31
- Understanding metadata.** (2004). Bethesda, MD: NISO

- Using Dublin Core. Dublin Core Qualifiers.** (2005). Recuperado el 12 de noviembre de 2009, de <http://dublincore.org/documents/usageguide/qualifiers.shtml>
- Vellucci, S.** (1998). Metadata. *ARIST*, 33, 187-222
- W3C** (Eds) (2004). *RDF/XML Syntax Specification (Revised)*. W3C Recommendation 10 February 2004. Recuperado el 6 de junio de 2009 del sitio oficial del W3C <http://www.w3.org/TR/rdf-syntax-grammar/>
- W3C** (Eds.) (1999). *Resource Description Framework (RDF) Model and Syntax Specification W3C Recommendation 22 February 1999*. Recuperado el 12 de marzo de 2009 del sitio oficial del W3C <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>
- Waters, D.J.** (1995) *The social organization of archiving digital information*. Recuperado el 22 de diciembre de 2009, de <http://ifla.queenslibrary.org/documents/libraries/net/waters2.htm>
- Weibel, S.** (1997). The Dublin Core: a simple content description model for electronic resources. *Asis Bulletin*, 24 (1). Recuperado el 10 de septiembre de 2009, de <http://www.asis.org/Bulletin/Oct-97/weibel.htm>
- Weibel, S.** & Dekkers, M. (2002). Dublin Core Metadata Initiative Progress Report and Workplan for 2002. *D-Lib Magazine*, 8 (2). Recuperado el 23 de octubre de 2009, de <http://dx.doi.org/10.1045/february2002-weibel>
- Wendler, R.** (2000). *Diversificación de actividades: habilidades y funciones catalográficas en la era digital*. En: F.F. Martínez Arellano & L. Escalona Ríos (Comps.). *Internet, metadatos y acceso a la información en bibliotecas y redes en la era electrónica* (pp.36-48). México: Universidad Nacional Autónoma.
- Woodward, J.** (1996) Cataloging and classifying information resources on the Internet. *ARIST*, 31 , 189-220
- Xu, A.** (1997). *Metadata conversion and the library OPAC*. Recuperado el 12 de octubre de 2007, del sitio Web de IFLA: <http://www.ifla.org/documents/libraries/cataloging/metadata/xu.pdf>
- Younger, J.** (1997). Resources description in the digital age. *Library trends* 45, 462-81

