

Metadatos para la codificación y transferencia de datos bibliográficos

POR IGNACIO LÓPEZ GUILLAMÓN
Doctor y Ayudante de Archivos y Bibliotecas
Universidad de Extremadura

RESUMEN: Se expone una presentación didáctica de los nuevos modelos de ensamblaje y transferencia de los datos bibliográficos en la Era digital. Principalmente, se trata de creaciones de la Library of Congress: MARCXML, MODS, MADS, METS o de otras entidades para DC, Indecs, DIDL, OAI-PMH o SRU/W.

Se trata de ofrecer una visión introductoria, acompañada de una bibliografía relevante, que sirva a los profesionales de la catalogación como panorámica en lengua española sobre unos conceptos y procedimientos que se van a imponer a corto plazo en la formación y práctica catalográficas.

Se sigue una metodología cualitativa, desde un posicionamiento de análisis crítico y desde la idea de no ruptura en la evolución reciente de la catalogación de las dos últimas décadas. Se aboga por resaltar la necesidad de asumir los cambios tecnológicos relativos al procesamiento y transferencia de la información bibliográfica, exigidos por la tecnología digital.

Se concluye apostando por conocer y aplicar los retos de asumir la catalogación digital, en tanto modelo que sustituye el modelo de la catalogación automatizada.

PALABRAS CLAVE: Indecs, MADS, METS, MODS

Metadata for encoding and bibliographic data transfer

ABSTRACT: It presents a didactic presentation of new models of assembly and transfer of bibliographic data in the digital age. Primarily, these are creations of the Library of Congress: MARCXML, MODS, MADS, METS or other entities to DC, Indecs, DIDL, the OAI-PMH and SRU / W.

It is offering an introductory overview, accompanied by relevant literature, to serve as professionals panoramic cataloging in Spanish on concepts and procedures will impose short-term training and cataloging practice.

It is a qualitative methodology, from a position of critical analysis and from the idea of no break in the recent developments in the cataloging of the last two decades. It advocates emphasize the need to embrace the technological

changes related to the processing and transfer of bibliographic information, required by digital technology.

Betting is concluded about and using the challenges of taking digital cataloging, while the model it replaces the model of automated cataloging.

Key words: Indecs, MADS, METS, MODS

1. INTRODUCCIÓN Y OBJETIVOS

El concepto de dato bibliográfico, usado en el título, exige una reflexión previa. La sociedad desarrollada del presente se la conoce con denominaciones simbólicas relativas a una caracterización técnica o sapiencial como «Era digital» o «Sociedad del conocimiento», por citar sólo dos que guardan estrecha relación con el devenir de las últimas décadas sobre la significación de «dato bibliográfico».

En el contexto del modelo de catalogación automatizada, vigente entre las décadas de los '60 a los '00, dato bibliográfico es cualquiera de los elementos informativos de naturaleza bibliográfica que sirven para representar una obra de creación, de innovación..., en otras palabras, el nombre del autor, la ciudad de edición, el título, el título uniforme..., la numeración o volúmenes de una obra o los nombres de los conceptos que contiene la obra o la denominación de otras responsabilidades intelectuales o artísticas distintas de la responsabilidad intelectual primera...

Frente al dato bibliográfico, que tiene, además, una neta funcionalidad discriminativa en la formulación de las peticiones de acceso y recuperación de la información, el registro de tales datos presenta otros relativos a la gestión biblioteconómica de la obra. Es el caso del número de registro, número de ejemplar o bibliográfico, signatura...

Si convenimos en hacer del dato bibliográfico una unidad básica de información con este potencial significativo, se está confiriendo a este significante un valor clave en el trasvase de las informaciones bibliográficas. En virtud de esto, el dato bibliográfico forma parte de un lenguaje, que, como tal, exige que los actores de esa comunicación, los usen con un significado convencional, que no admite interpretaciones de connotación o de denotación. A es A; a es a; A no es a..., véase A.

En la llamada Sociedad del conocimiento hay una sobreabundancia de datos y de información. Se exige, pues, en este contexto que tales datos se entiendan conforme a una sistematización de relevancias, consensuadas en los distintos ámbitos de saber. Se trata de que un dato pueda significar una información, o una serie de éstas, con tal de que se haya/n convenido previamente. Sólo de este modo, el dato se transforma en agente de conocimiento. El sujeto contrasta las informaciones que le aportan los datos, une lo sabido a lo aprendido y es capaz de establecer conexiones con otras informaciones hasta llegar a procesar y presentar un nuevo conocimiento.

Daniel Innerarity relaciona la acción de conocer lo nuevo a partir de relacionar una información con la experiencia, juicio, intuición y valores de un sujeto. En otras palabras, el saber es información con valor derivado de un cierto nivel de reflexividad (Innerarity, 2011, p. 25-27).

De este modo, se ha querido enlazar los conceptos de dato bibliográfico, modelo de catalogación automatizada, comunicación de la información bibliográfica y Sociedad del conocimiento, convencidos de que el fenómeno de la codificación y transferencia de datos bibliográficos presenta en nuestros días una realidad que lo hace más determinante en el proceso de fijación y comunicación del saber. El paso del tiempo parece dar la razón a Elaine Svenonius (2000, pp. 31-50), quien años atrás abogó por que la catalogación bibliográfica constituye un lenguaje específico.

Los debates que han conducido hacia los Requisitos funcionales de los registros bibliográficos (FRBR) o los Requisitos funcionales de los datos de autoridad (FRAD) han respondido a las exigencias de la tecnología digital en los procesos de acceder, recuperar y comunicar los conocimientos. En este contexto ha sido necesario abordar el replanteamiento de unos Principios internacionales de catalogación (2009) y la tradición reglamentaria de catalogar, para la que se ha diseñado un marco normativo de aspiración global, denominada RDA: descripción y acceso a recursos (2010), de aplicación experimental en 2011 y en fase de readaptación en 2012.

Estos cambios han sobrevenido en el transcurso de dos décadas y han sido propiciados por la comunidad de expertos en catalogación ante la irrupción, implantación y globalización de la tecnología digital en procesos tales como fijación, difusión y recuperación de la información en un medio no analógico. En la medida en que las sociedades participan de la tecnología digital, la urgencia de los cambios se acelera.

Valga lo antes indicado como pórtico introductorio. Planteamos, seguidamente, los objetivos de este ensayo. En primer lugar se trata de presentar de manera explicativa los modelos y herramientas tecnológicas, surgidas en el transcurso de los últimos veinte años. En segundo lugar, se trata de contribuir a la familiarización con la nomenclatura y conceptos relativos al procesamiento de los datos bibliográficos, según el modelo de la catalogación digital.

Es abrumador el número de siglas y de conceptos especializados exigidos por las TICs en renovación permanente. Se pretende, pues, una finalidad propedéutica que oriente, si fuera necesario, sobre las mecánicas más técnicas en relación con las directrices e instrucciones, que conforman las RDA en su versión más formal.

Un tercer objetivo es establecer la conexión conceptual y convergencia técnica entre la operación de catalogar, tan asumida por muchos de los profesionales de bibliotecas y la elaboración de metadatos.

Los estándares que se citarán son el presente de la catalogación digital, como lo fueron en otra época la serie de las ISBD y del MARC. Sin duda, se asiste a un cambio de modelos, que no supone, de suyo, hecatombe alguna.

Lo sucedido hasta ahora, apuesta por asumir las TICs que van a permitir comunicar cualquier conocimiento, de forma inteligible, en cualquier sociedad innovadora.

2. RECURSOS Y METODOLOGÍA

Las fuentes de conocimiento que nos han servido en la elaboración del siguiente ensayo han sido los documentos que sobre los productos y modelos operacionales ofrecen las webs de las bibliotecas nacionales europeas, la Library of Congress, Bibliothèque National de Canada, OCLC, IFLA, algunos autores puntuales. De todos ellos se hace referencia oportuna y se recogen en las referencias bibliográficas finales.

Se trata de presentar un trabajo eminentemente propedéutico, que divulgue en la comunidad de catalogadores españoles los estándares y nueva nomenclatura exigida por el nuevo modelo de catalogación digital, frente al, todavía, vigente de catalogación automatizada. En la medida de lo posible, se rehuye el tecnicismo más especializado, conscientes de que, siendo la operación de catalogar eminentemente práctica, los aspectos teóricos deben ser relegados para otro tipo de trabajo.

En este contexto, los análisis realizados son de naturaleza cualitativa. A efectos de reforzar el fin didáctico de este trabajo se emplearán tablas con nociones muy esquematizadas. Como soporte teórico, a nivel de fundamentos del trabajo, nuestras premisas han partido del racionalismo crítico de K. Popper (1902-1994) y la teoría de paradigmas de T. S. Kuhn (1922-1996), que estimamos crucial para comprender la trayectoria especulativa de la operación de catalogar (Kuhn, 1962, p. 10; 1977, p. 319). Igualmente, ha constituido un acicate positivo el «no saber desconocido» (unknown unknowns), del que habla Daniel Innerarity (1959-), característico de la Sociedad del conocimiento en tanto que rasgo descriptivo y estímulo operativo para avanzar en la creación de nuevos saberes (Innerarity, 2011, p. 63-67).

Igualmente, hay que destacar el valor de la teoría de sistemas y la interdisciplinariedad que implican las apuestas teóricas de los últimos años sobre el concepto de universo bibliográfico en el marco conceptual de los FRBR, FRAD y las RDA (Rodríguez García, 2009, pp. 33, 37, 40, 44, 48-50).

3. ANÁLISIS Y DISCUSIÓN

Como la operación de catalogar participa activamente en el proceso de comunicar los saberes del presente y del pasado, en correspondencia lógica ésta asume las transformaciones que implica esta funcionalidad comunicativa. Se asiste a una metamorfosis de concepto, de práctica y de nomenclatura de la operación de catalogar.

Los metadatos, como agrupación de datos bibliográficos, no son ajenos al registro bibliográfico automatizado. En nuestros días, Barbara B. Tillet en su

trabajo por divulgar las RDA, expone que éstas son un estándar de catalogación con el que elaborar metadatos. Igualmente expresa que catalogar, en su sentido genuino y práctico, es elaborar agrupaciones de datos bibliográficos para manifestaciones y expresiones de las obras (Tillet, 2010, p. 10, 16, 23, 29).

Siendo esto así, se pueden plantear distintos interrogantes: ¿las novedades implican cambios que afectan al concepto mismo de catalogación?, ¿cómo va a afectar a la práctica catalográfica disponer del nuevo modelo catalográfico, basado en FRBR y FRAD?, ¿cómo afecta a la praxis catalográfica los nuevos principios internacionales de catalogación?, la nomenclatura surgida ¿expresa nuevos conceptos, nuevas prácticas o nuevos productos?

Lo óptimo es asumir pragmáticamente los cambios que se operan en el sujeto y objeto de la catalogación bibliográfica. En otro plano deben situarse las modificaciones de los procedimientos y aceptar la supeditación de éstos al fin comunicativo. Las transformaciones afectan por igual a otros fenómenos culturales, como la biblioteca o el libro. Sobre cualquiera de ellos pesa una evolución varias veces centenaria.

La catalogación bibliográfica, junto con los conceptos de biblioteca y libro, se ligan culturalmente en la transmisión fidedigna de los saberes, independientemente de sus conceptualizaciones técnicas o históricas o, meramente, especulativas. En lógica con lo expuesto, entendemos que conceptos como biblioteca digital, recurso electrónico, metadatos u otras expresiones que pudieran citarse, sirven para expresar con propiedad nuevas realizaciones de nuestros días. En cualquier caso, éstas poseen una trayectoria cultural por las que fluye el conocimiento desde el pasado hacia el presente.

Si se pasa de un nivel de expresión más generalista a otro más especializado, sin duda, debería emplearse una expresión acorde con el contexto de comunicación. En virtud de este fin comunicativo, se ha estimado que era conveniente presentar de forma conjunta algunos de los conceptos y sus expresiones, principalmente siglas, que constituyen la forma presente de materializar lo que hasta ahora se conoce por operación de catalogar.

3.1. *MARCXML, DCMI e Indecs*

MARCXML

La nueva realidad tecnológica hizo que en la primera década del siglo XXI, el formato MARC tuviera su versión en XML, toda vez que ello contribuía a una mejor gestión de los datos bibliográficos. XML expresa la estructura del formato MARC21, de forma que cada zona del formato MARC se transforma en un único elemento que incluye los atributos de la etiqueta y sus indicadores. Los campos variables de control MARC (00X) y sus campos variables de datos (01X a 8XX) constituyen los elementos específicos y son tratados como cadenas de caracteres.

Esto permite la conversión de datos MARC, según la norma ISO 2709 Información y documentación –Formato para el intercambio de información, en MARCXML, y su reconversión, sin pérdida de información. Éste ha sido elaborado por la norma ISO 25577 Información y documentación–MarcXchange (2008), que se presenta como alternativa a la norma ISO 2709, antes citada.

El MARCXML permite expresar en XML metadatos descriptivos con independencia de su presentación en MARC, como sucede con el protocolo OAI-PMH o con la extensión del esquema METS para los metadatos descriptivos. MARCXML está disponible en la Library of Congress Network Development, que conjuntamente con MARC Standards Office, han sido los responsables finales de este esquema de trabajo con el formato MARC en entorno de XML (MARCXML..., 2011).

TABLA I: UTILIDADES DE MARC EN XML

Utilidades de MARC en XML
Presentar los datos de maneras distintas
Gestionar la coherencia y corrección de las informaciones
Convertir los datos codificados en MARC en otros formatos XML, como DC o MODS
Ofrecer e intercambiar datos codificados en MARC en otros protocolos derivados de XML, como OAI-PMH o SRU

DCMI

DCMI es el acrónimo de Dublin Core Metadata Initiative (DC) (Hilman, 2005). Se trata de un formato integrado por quince elementos, agrupados en tres categorías:

- (1) sobre el contenido,
- (2) sobre los derechos de autor, y
- (3) sobre el recurso en particular.

Sus orígenes están en OCLC y NCSA (1995) (Lamarca, 2006). La actual versión fue validado en 2007 por la norma ANSI/NISO Standard Z39.50-2007 y por la ISO 15836 (2003, 2009). En la actualidad es gestionada por la Agencia DCMI (Dublin Core..., 1995-2011).

DC, en forma abreviada, ofrece un conjunto común de elementos descriptivos con los que mejorar la codificación y búsqueda de recursos, con independencia de si es un recurso de archivo, biblioteca o museos u otras comunidades de información afines, e igualmente, sin tener en cuenta sus formatos descriptivos propios.

DC permite la interoperabilidad entre sistematizaciones independientes de datos descriptivos, que respeta las reglamentaciones y estándares relativos

a la información bibliográfica, principalmente, en torno a los elementos de título, responsables, materias y descripción (Dublin Core en castellano..., 2008).

Es significativo indicar que Dublin Core es distinto a los formatos de codificación y almacenamiento de la información y que debe aplicarse conforme a un método que garantice su aplicación uniforme en su texto completo, en los recursos HTML y XHTML, en los documentos XML y en el entorno RDF. Dublin Core tiene especial utilidad en el protocolo OAI-PMH como formato de base (Directrices Driver 2.0..., 2009).

Recientemente se ha desarrollado un Dublin Core cualificado (Daudinot, 2006). Es semejante al Dublin Core Simple más algunas particularidades, a saber:

- (1) datos relativos a la audiencia, origen y derechos de autor;
- (2) cualificadores que concretan los elementos generales de Dublin Core y
- (3) modelos de codificación que relacionan un elemento y un vocabulario (clasificaciones de Dewey o decimal universal) y sintaxis controlados (Norma ISO 639-2 sobre códigos de lenguas).

Indecs

Indecs responde al acrónimo Interoperabilidad de Datos en Sistemas de E-Comercio (Rust, 2000). Surgió entre 1998 y 2000, con patrocinio de la Unión Europea, a instancias de organizaciones de música, de derechos de autor, editoriales, autores o bibliotecas a fin de alcanzar un marco de estándares de metadatos en los que sustentar el comercio de recursos digitales respetando los derechos de autor.

Se trata de un modelo que precisa las características de los metadatos relativos al comercio electrónico sobre derechos de autor en entornos digitales, de modo que los creadores de recursos digitales pueden elaborar sus obras, ofrecerlas por Internet y que distintos compradores puedan adquirirlas por el mismo medio.

Los metadatos Indecs se integran por tres tipos de elementos: entidad, atributo y valor o función. Sigue para ello la terminología sobre sistemas de información definida por la ISO 11179, que es propia del lenguaje extensible de marcas (XML) y del marco de descripción de recursos (RDF).

Por entidad se entiende aquello que identifica un concepto; por atributo, las características o propiedades de una entidad y son etiquetas, cantidades, calidades, tipos y funciones; y valor, que es la función de un atributo. Estas funciones expresan las relaciones entre dos, o más, entidades que encarnan distintos valores, como entidad de entrada, entidad de salida, entidad de contexto, entidad de tiempo o de lugar.

Contempla, también, unos identificadores que se asignan a cada elemento de los metadatos Indecs. Coinciden con las secuencias alfanuméricas de las distintas categorías de documentos, a saber: Identificador de Objetos

Digitales (DOI), código de barras EAN, ISAN, ISBN, ISMN, ISSN..., con ellos se facilita su relación a efectos de relaciones lógicas y de intercambio.

Las relaciones afectan a las entidades y se estructuran en tres niveles de complejidad: eventos, que definen las relaciones cuando algo cambia y se expresan en verbos activos; las situaciones, que definen las relaciones por las que algo sigue siendo eso mismo y se expresan por los verbos posesivos tener o poseer; y atributos, definidos por cinco modalidades (etiquetas, cantidades, calidades tipos y funciones) que determinan la estructura general de desarrollo e interoperabilidad de los metadatos.

Se expone ahora qué se entiende por obra, expresión, manifestación y ejemplar, que son las entidades clave del modelo entidad-relación definidas en FRBR, en los metadatos indecs. Obra es una abstracción de una realidad concreta, que puede ser creación intelectual o artística. Tal obra se materializa a partir de una expresión, susceptible de tener derechos de autor diferenciados, ya se trate de un video sobre la Alhambra con fines turísticos, que se expresa en alemán si se destina a este público; en inglés si se destina al mundo en general o en francés, si se destina a la comunidad francoparlante, o si es el caso de una guía turística impresa en distintos idiomas o un web multilingüe que presenta este lugar de España.

La obra y la expresión, por cuanto está sujeta a derechos de autor, tienen garantizada la integridad de sus elementos integradores: autor, presentación, formato, características técnicas...

Manifestación sería la guía turística de la Alhambra en alemán o la web de la Alhambra, en tanto que el ejemplar sería un libro concreto de la tirada hecha de esa edición o una versión concreta del sitio web sobre la Alhambra. El metadato se hará sobre la manifestación y el ejemplar. Indecs es un metadato orientado a aplicaciones en relación con el comercio de obras digitales en un medio digital.

La disponibilidad de los metadatos genera situaciones de conflicto de derechos de autor semejantes, por ello Indecs fija la información identificativa de los titulares de derechos de autor, los objetos a los que afectan tales derechos y su alcance en tanto que son susceptibles de una transacción comercial digital. Define al metadato como una relación entre dos entidades que así lo expresan.

TABLA II. PRINCIPIOS DE INDECS

Principios de Indecs
Única identificación, por la que cada entidad debe ser reconocida
Granularidad funcional, por la que es posible identificar una entidad de manera permanente
Designación de autoridad, del elemento del metadato, debe ser concreta y definida
Adecuación de acceso, por el se garantiza el acceso de forma confidencial y eficiente a los metadatos deseados

En Indecs se valoran especialmente la identificación unívoca de las entidades, lo atributos y sus relaciones. Se aplica a RDA/ONIX, marco para la categorización de recursos; a ISO/IEC 21.000-6 (MPEG), Intercambio digital de datos (DDEX, Digital Data Exchange), a ONIX (intercambio de información entre editores en la distribución digital de metadatos) y como identificador en el sistema de esquemas de metadatos de objeto digital.

El principio de granularidad funcional merece atención. La granularidad sirve para expresar particularidades de interés informativo potencial. Es, en la práctica, útil en documentos que exigen una realización sectorizada y secuencial: proceso de creación de una película, un kit multimedia, un álbum fotográfico... Al plantear la granularidad de los eventos, Indecs establece una sintaxis sobre cinco puntos:

- (1) cada entidad en un evento posee una función que relaciona entidad y elemento;
- (2) cada acontecimiento tiene un agente;
- (3) las entidades pueden tener más de una función en un único evento;
- (4) dos o más entidades pueden tener una misma función en un único evento;
- (5) las funciones de no-agente deben reclamar a las funciones agente.

Todo lo cual se ejemplifica de la siguiente forma: Siendo A un escritor y B un ilustrador, ambos preparan la edición de un cuento ilustrado C, en Zaragoza, primavera de 2012. La cadena expresiva del evento podría ser A + B + C + Zaragoza + primavera de 2012. Ahora bien, podríamos aportar más información y especificar que A escribió en la primavera de 2012 y B ilustró en Valencia en el otoño de 2012, en cuyo caso la secuencia expresiva del evento sería doble: (1) A + X + Zaragoza + primavera 2012; (2) B + X + Valencia + otoño 2012.

Indecs ha generado un modelo de metadatos general con el que describir de forma unívoca los interlocutores que median en la compra-venta de recursos digitales, precisar los derechos de autor y relacionar las personas que participan en el comercio electrónico, así como sus acuerdos transaccionales.

3.2. MODS, MADS, METS

MODS

Modelo de descripción de objetos en metadatos (MODS) podría ser la traducción de Metadata object description schema (MODS..., 2011). Se trata de una modalidad de tratamiento de datos bibliográficos en contexto bibliotecario, toda vez que permite hacer referencias a las normas de contenido de los documentos. Se creó en 2002 por la Library of Congress para intercambiar y compartir datos bibliográficos en formato MARC21 sin pérdida de información.

Puede aplicarse a documentos digitales y analógicos en cualquier soporte físico (papel, impreso o manuscrito; registros sonoros, documentos iconográficos...) (Coburn, E. y otros, 2010, pp. 15-29). Por su flexibilidad es útil en bibliotecas y también, para investigadores y editores. MODS puede presentar los datos bibliográficos en XML en bases de datos de una biblioteca o de una red de éstas (MODS Schema, 2010).

Viene a ser un formato intermediario que sirve de nexo entre registros bibliográficos elaborados en formatos diferentes, con grandes posibilidades en un portal web OAI, toda vez que puede rediseñar la descripción bibliográfica de una versión digital. Por esta causa, es capaz de convertirse en un componente de los metadatos de recursos digitales que usan el esquema METS. Está extendido por Alemania, Canadá, Estados Unidos, Francia, Suiza, etc.

Se constituye por doce categorías de informaciones bibliográficas relativas al título, autoría, tipo de recurso, información sobre la edición y publicación, lengua y codificación, descripción física, resumen, sumario, público al que se dirige, notas y materias tratadas que sirven de acceso y recuperación posterior (MODS user guidelines..., 2009). Puede ahondarse sobre MODS en Bernhard Haslhofer (2010) o en Jane Greenberg (2009).

MADS

Se trata del modelo de descripción de autoridad en metadatos (MADS). Ésta podría ser la traducción de Metadata authority description schema. MADS se ha desarrollado, al unísono con MODS, por Library of Congress para dotar a los datos de autoridad un modelo semejante al dado a los datos bibliográficos (MADS, 2011). Se creó como complemento de SKOS/RDF, desarrollado, también, por la Library of Congress como modelo de datos común para compartir y enlazar sistemas de organización del conocimiento a través de la tecnología web. En cualquier caso, MADS es el instrumento terminológico específico con el que representar un vocabulario de autoridad bibliográfica en su contextualización semántica.

Se ha diseñado para los datos sobre autoridades bibliográficas y la terminología propia de las ciencias de la información y las bibliotecas. Se enfoca hacia el control convencional del uso de nombres personales, corporativos, geográficos...; de tesauros y taxonomías conceptuales. De este modo, MADS viene a ser el medio para registrar los datos de autoridad bibliográfica de recursos culturales y bibliográficos en el formato MARC 21. En consecuencia, no se destina a la descripción de objetos del mundo real y es ajeno a las reglas de creación y uso de etiquetas (RDA/RDF..., 2011).

Consta de una autoridad y aquellos otros con los que ésta se relaciona o son sus variantes. Estos datos principales pueden contener uno, o más, subelementos. Además, el registro MADS consta de elementos descriptores, tales como nombre, información sobre títulos, tema, datos cronológicos, géneros,

conceptos geográficos...; y de elementos menores, tales como afiliación, clasificación, área de competencia, identificador, idioma, notas, URL, extensión... Más información sobre MADS se encuentra en Alicia Sellès Carot (2010).

METS

Norma de codificación y transmisión de metadatos (METS) podría ser la traducción de Metadata encoding and transmission standard. Se trata de un formato creado en 2001 por la Digital Library Foundation y gestionado por la Library of Congress para presentar distintos metadatos de un documento digital, ya sea para facilitar su intercambio, su gestión o su conservación (METS, 2011).

METS tiene una configuración en entorno XML y se integra por siete categorías de metadatos, recogidas en la Tabla III.

TABLA III: SECCIONES DE METADATOS METS

Secciones de metadatos METS
Cabecera METS, para la descripción del documento
Metadatos descriptivos, ya se trate de un registro MARC o un documento EAD, otros metadatos descriptivos o una mezcla de metadatos internos y externos
Metadatos administrativos, que informan de la trayectoria seguida por el registro de datos que forman el objeto digital
Sección de archivo, que muestra el listado de archivos con contenidos que forman parte del objeto digital
Mapa estructural, que contiene la estructura jerárquica del objeto digital, los enlaces a los diferentes archivos de contenido y sus metadatos correspondientes
Enlaces estructurales, que registran la existencia de hiperenlaces entre las secciones del mapa estructural, como sucede en el caso de los web
Comportamientos, usados para vincular éstos con los programas que permiten la lectura del documento METS

METS posee algunas particularidades (Hilman, 2010). Así, la información de los metadatos descriptivos y de las subsecciones de los metadatos administrativos (información sobre la fuente del documento digital, información técnica sobre el documento en su conjunto, información sobre el documento digital, y estado jurídico del documento) puede ir encapsulada en XML o, también, puede hacer referencia a una descripción.

En el primer caso, la información bibliográfica del metadato descriptivo puede presentarse en DC, en DC cualificado, en EAD, en MARCXML, etc.,

según se necesite. En el segundo caso, la información bibliográfica se presenta a partir de un identificador, inequívoco, que lleva hasta un catálogo.

Por otra parte, es significativo el número de valores expresados en la sección de archivo y en el mapa estructural, cuyo uso se deja al criterio del profesional. Tal sucede con el listado de los archivos con contenido que forman parte del objeto digital. Otro tanto puede decirse de los mapas físicos, o lógicos, del mapa estructural. Así, una estructuración de tipo físico establece una manera lineal de usar el documento digital.

Por el contrario, en una estructuración de tipo lógico, se permite un uso multilineal e independiente de la forma en que el documento ha sido digitalizado. Igualmente, hay que destacar la libertad de acción que el profesional tiene para especificar la gradualidad de la que se compone el documento. Hace referencia a la gradualidad entre un título de una obra multiparte, a una obra concreta de la misma, a una página y a un archivo digital.

Todo lo cual está en relación por conexiones internas y por este medio se puede reunir la información de un metadato descriptivo o una subsección del metadato administrativo con su respectivo nivel de graduación. Así, una descripción bibliográfica se hace sobre el documento conjunto, en tanto que la información sobre la página de título sólo corresponderá a un elemento particular.

También, reseñar que bastantes elementos de los METS son identificados internamente en el archivo, ya se trate de secciones, ficheros, etc. La forma en que los identificadores se estructuran la decide el usuario.

De todo esto deriva, la flexibilidad que caracteriza la norma METS tanto para expresar la información, como para adaptarla a necesidades concretas y también, el hecho de requerir que se escoja por parte del usuario un perfil concreto, ya sea en la presentación de la información de los metadatos descriptivos, en los metadatos administrativos, en el nombramiento de identificadores, los listados de archivo y de mapa estructural con los tipos de estructura y los niveles de gradualidad, arriba expresados.

Para ampliar información sobre estos modelos de transcripción y de transferencia de metadatos, la obra de Mary Beth Weber y Austin Angela Fay (2010) ofrece información puntual y contextual de interés o en Ioan Marius Bilasco, y otros (2010).

3.3. *DIDL*, *OAI-PMH* y *SRU/W*

DIDL

En 2003 surgió Digital ítem declaration language (*DIDL*) de la mano de Motions Pictures Experts Group y normalizados por la ISO/IEC – 21000. En 2005 fue revisada. *DIDL* facilita la interoperabilidad entre objetos digitales complejos y su gestión en entorno XML; abarca tanto los aspectos de descripción como los derechos de autor, la gestión de identificadores y de útiles de referencia. También se le conoce por la forma MPEG-21 *DIDL* (MPEG-21 Part 2, 2004).

Con relación a la primera, DIDL contempla que haya hasta cuatro niveles (conjunto de contenedores, conjunto de componentes, división abstracta de un conjunto de componentes y archivo digital individual). Los tres primeros son conceptualizaciones abstractas frente al nivel de concreción del archivo.

El conjunto de componentes posee el mayor grado de coherencia intelectual y es el nivel de referencia de DIDL. Es usado para los intercambios y para la manipulación de la información, toda vez que suele coincidir con el documento digital concreto. El conjunto de contenedores viene a ser una agrupación de conjunto de componentes.

La división abstracta de un conjunto de componentes posee una unidad tal que equivale a un archivo digital concreto. En el caso de una monografía digitalizada, el conjunto de componentes se identifica con el documento; cada archivo digital individual es un archivo digital concreto.

En DIDL se pueden conectar una o varias descripciones con cada nivel, a partir de un elemento descriptor. Éste se compone de una declaración, si la descripción es textual; y de un componente que incluye un archivo digital individual si la descripción es un registro digital. Por su parte, la información sobre la gestión que configura al documento digital es necesaria para gestionar un documento cuyas partes están sujetas a distintas situaciones de derechos de autor; y sobre la captación del formato adecuado de lectura en base al archivo digital puntual usado en la digitalización de la información.

Entre los protocolos para el intercambio de datos bibliográficos están la Norma ISO 23950 – Recuperación de la información: definición de servicios de aplicaciones y protocolo de especificaciones. Es más conocida por la denominación: protocolo Z39.50; se creó en los años '80 y ha estado actualizándose hasta 1998 (Lynch, 1997).

TABLA IV: FUNCIONALIDADES DE DIDL

Funcionalidades de DIDL
Expresar informaciones de estructura, que expresan distintos niveles de gradualidad en un documento
Reflejar informaciones descriptivas representativas de todo el documento o una parte
Presentar informaciones de gestión que hacen posible la configuración del documento digital

OAI-PMH

Con posterioridad, se han desarrollado otros protocolos más acordes con el avance de la tecnología digital. Entre éstos se encuentran Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH) (1999) (Barrueco, 2003) y Search/Retrieval via URL (SRU) (2002) (SRU, 2012).

OAI-PMH es un protocolo de intercambio de metadatos en el marco de Internet y como tal, facilita el acceso a los documentos digitales. Su web es de uso libre (Open Archive Initiative, 2012) y cuenta con proveedores de datos que difunden sus metadatos en un depósito y proveedores de servicios que reúnen tales datos y los presentan en sus bibliotecas digitales. Cumplen, pues, una doble función de identificación y de difusión, de descripción y acceso, que se sustenta en la filosofía cliente-servidor, de forma tal que a una demanda en http por parte del cliente, el servidor responde con resultados en XML.

TABLA V: FUNCIONALIDADES DE OAI-PMH

Funcionalidades de OAI-PMH
Aumentar la visibilidad de éstos
Constituir conjuntos de recursos digitales
Aportar recursos a los portales temáticos

SRU/W

SRU/W es un protocolo de intercambio de metadatos, muy semejante al protocolo Z39.50, adaptado a la tecnología digital (SRU, 2012). Ha sido creado por el Grupo ZING para Z29.50 International Next Generation en 2002 como un protocolo doble para la búsqueda y recuperación de la información vía URL (SRU) y búsqueda y recuperación de la información vía servicios web (SRW). Se actualizó en 2007 y para 2011 se prepara, de nuevo, otra versión. Posee un gran alcance en la exploración y análisis de los metadatos de un archivo de éstos.

Como se recoge en la Tabla VI, este protocolo establece hasta tres tipos de comunicación entre un servidor y un cliente por los protocolos HTTP (Hypertext Transfer Protocol) o SOAP (Simple Object Access Protocol) (SOAP..., 2007).

TABLA VI: COMUNICACIONES CLIENTE-SERVIDOR POR HTTP

Comunicaciones cliente-servidor por HTTP
Comunicación explicativa, que muestra el nombre del servidor de los recursos e incluso, las direcciones de otros posibles servidores, además de un listado con los conceptos sobre los que formular las demandas de información, así como los modelos de metadatos (DC, MODS, MARCXML...);
Comunicación de búsqueda y recuperación de los registros según el planteamiento hecho;
Comunicación a partir de un listado de entradas en relación con un índice, de forma que a un lado consta el listado de las entradas y a otro el número de resultados de cada entrada. La utilidad de esta aplicación es servir en la selección de los términos más significativos

Si cualquiera de estas dos últimas aplicaciones no ofreciera resultados, el servidor, según la versión 1.2 de SRU (2007), debe responder desde la aplicación explicativa. El protocolo SRU/W contempla una comunicación entre el cliente y el servidor basada en mensajes de error y en la posibilidad de añadir más información sobre la aplicación y sobre la respuesta. Igualmente, el administrador del servidor SRU/W puede especificar correspondencias entre los índices o restringir, si es el caso, los índices a ciertos contextos.

4. CONCLUSIONES

1. El dato bibliográfico constituye un elemento básico en un lenguaje específico: la catalogación bibliográfica. Éste posee una significación convencional y es tratado en el proceso de la comunicación del saber conforme a los desarrollos tecnológicos coetáneos.
2. El dato bibliográfico se reviste de la formalidad y nomenclatura tecnológicas de la sociedad en que se crean y difunden los recursos informativos. Sirven para la representación de tales recursos y se transfieren de manera geoméricamente creciente junto al mismo.
3. El análisis de la operación de catalogar desde presupuestos teóricos contribuye a separar sus fundamentos conceptuales, sus funcionalidades y los procedimientos puramente técnicos. Esta disección origina que de la misma se comience a percibir una imagen de disciplina de conocimiento dinámica y versátil, estructurada y sólida, innovadora y secular.
4. En la práctica catalográfica se acusa una renovación intensa, que es paralela a la obrada en su objeto: el recurso informativo; y en su fin: difusión del saber y que afecta notoriamente al lenguaje con el que se expresa. Se impone abordar tal nomenclatura y asumir sus aportaciones, teóricas y prácticas, en la certeza de que la sustancia de la fijación y transferencia de los conocimientos, objeto y fin de la catalogación, se enriquece a partir de su adecuación tecnológica permanente.
5. El cambio del modelo de catalogación automatizada al digital implica abrir los límites conceptuales y empíricos de la operación de catalogar hacia la abstracción «universo bibliográfico» y conectar con la transformación técnica que afecta a los procesos mismos de creación y difusión del saber.

5. REFERENCIAS BIBLIOGRÁFICAS

- BARRUECO, J. M. y SUBIRATS COLL, I. *OAI-PMH: protocolo para la transmisión de contenidos en Internet*, 2003 [en línea]. [Fecha de acceso 15 febrero 2012]. Disponible en Internet: <http://www.uv.es/=barrueco/cardedeu.doc>
- BILASCO, I. M., y otros. *Semantics for intelligent delivery of multimedia context. SAC '10 Proceedings of the 2010 ACM Symposium on Applied Computing*. New York: ACM, 2010. ISBN 978-1-60558-639-7.

- COBURN, E., y otros. The cataloging cultural objects experience: codifying practice for the cultural heritage community [en línea]. *IFLA journal*, 2010, 36, n. 2, pp. 15-29. [Fecha de acceso 1 febrero 2012]. Disponible en Internet: <http://ifl.sagepub.com/content/36/1/16.full.pdf+html>
- DAUDINOT FOURNIER, M. C. Isabel. Descripción de los recursos de información en Internet: formato Dublin Core [en línea]. *Acimed*, 2006, 14, n. 4. [Fecha de acceso 4 enero 2012]. Disponible en Internet: http://bvs.sld.cu/revistas/aci/vol14_4_06/aci09406.htm
- DCMI, *Dublin Core Metadata Initiative*, 1995-2011 [en línea]. Fecha de acceso 26 diciembre 2011]. Disponible en Internet: http://www.dublincore.org/Directrices_Driver_2.0, creado 4 septiembre 2009 [en línea]. [Fecha de acceso 05 diciembre 2011]. Disponible en Internet: http://www.driver-support.eu/documents/DRIVER_2_0_Guidelines_Spanish.pdf
- Dublin Core en castellano: elementos del conjunto de metadatos de Dublin Core: descripción de referencia*, actualizado a 01.12.2008 [en línea]. [Fecha de acceso 04 febrero 2012]. Disponible en Internet: <http://www.rediris.es/search/dces/>
- GREENBERG, J. Metadata and digital information [en línea]. En: BALES, M. J.; MAACK, M. N. y DRAKE, M. (eds.). *Encyclopedia of library and information Science*. New York: Marcel Dekker, imp. 2009. [Fecha de acceso 3 febrero 2012]. Disponible en Internet: http://wiki.datadryad.org/wg/dryad/images/9/95/Metadata_ELIS.pdf
- HALSHOFER, B. y KLAS, W. A survey of techniques for achieving metadata interoperability [en línea]. *Journal ACM computing survey (CSUR)*, Vol. 42, n. 2, February 2010, pp. 7.1-7.37. [Fecha de acceso 6 febrero 2012]. Disponible en Internet: <http://dl.acm.org/citation.cfm?id=1667064>
- HILMANN, D. Guía de uso del Dublin Core, 2003 [en línea]. [Fecha de acceso 4 febrero 2012]. Disponible en Internet: http://www.sedic.es/usando_dublin_core.pdf
- Ídem y otros. RDA Vocabularies: procesos, outcome, use. *D-Libre Magazine*, 2010, Vol. 16, n. 1-2.
- LAMARCA LAPUENTE, María Jesús. *Hipertexto: el nuevo concepto de documento en la cultura de la imagen* [en línea], modificado 19 noviembre 2011. Madrid : UCM, Facultad de Ciencias de la Información, 2006. [Fecha de acceso 30 enero 2011]. Disponible en Internet: http://hipertexto.info/documentos/dublin_core.htm
- LYNCH, CA. The Z39.50 Information retrieval standard [en línea]. *D.Lib Magazine*, April 1997. [Fecha de acceso 1 febrero 2012]. Disponible en Internet: <http://www.dlib.org/dlib/april97/04lynch.html>
- MADS: *metadata authority description schema* [en línea], October 18, 2011. [Fecha de acceso 30 diciembre 2011]. Disponible en Internet: <http://www.loc.gov/standards/mads/>
- MADS/RDF *primer* [en línea], March 28, 2011. [Fecha de acceso 10 enero 2012]. Disponible en Internet <http://www.loc.gov/standards/mads/rdf/>
- MARXML *MARC21 XML Schema* [en línea]. [Fecha de acceso 20 diciembre 2011]. Disponible en Internet: <http://www.loc.gov/standards/marxml/>

- METS: metadatos encoding and transmission standard* [en línea], July 1, 2011. [Fecha de acceso 19 diciembre 2011]. Disponible en Internet: http://www.loc.gov/standards/mets/METSOverview_spa.html
- MODS: metadata object description schema* [en línea], October 18, 2011. [Fecha de acceso 28 diciembre 2011]. Disponible en Internet: <http://www.loc.gov/standards/mods/>
- MODS Schema* [en línea], October 18, 2010. [Fecha de acceso 28 diciembre 2011]. Disponible en Internet: <http://www.loc.gov/standards/mods/mods-schemas.html>
- MODS user guidelines versión 3* [en línea], July 28, 2009. [Fecha de acceso 29 diciembre 2011]. Disponible en Internet: <http://www.loc.gov/standards/mods/v3/mods-userguide-elements.html>
- MPEG-21 Part 2: Digital Item Declaration Language (DIDL)* [en línea], February 17, 2004. [Fecha de acceso 30 enero 2012]. Disponible en Internet: <http://xml.coverpages.org/mpeg21-didl.html>
- Open Archives Initiative* [en línea], February 1, 2012. [Fecha de acceso 31 enero 2012]. Disponible en Internet: <http://www.openarchives.org/>
- RODRÍGUEZ GARCÍA, A. A. Los objetos bibliográficos confirmados en la integración compleja de la descripción y acceso a recursos [en línea]. *Investigaciones bibliotecológicas*, 2009, 23, n. 48, p. 33-59. [Fecha de acceso 1 de febrero 2012]. Disponible en Internet: <http://www.ejournal.unam.mx/ibi/vol23-48/IBI002304803.pdf>
- RUST, G. y BIDE, M. *The «indecs» metadata framework: principles, model and data dictionary. June 2000, WP1a-006-2.0* [en línea]. [Fecha de acceso 6 febrero 2012]. Disponible en Internet: http://www.doi.org/topics/indecs/indecs_framework_2000.pdf
- SELLÈS CAROT, Alicia, y otros. Proposal of a goal-oriented shared catalog model [en línea]. *2010 IEEE Fourth International Conference on Semantic Computing*. Pittsburg : IEEE, 2010, pp. 502-507. ISBN 978-0- 7695-4154-9. [Fecha de consulta 20 enero 2012]. Disponible en Internet: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5629022>
- SOAP Version 1.2 Part: Messaging Framework* [en línea] (2nd ed.), 27 April, 2007. [Fecha de acceso 8 enero 2012]. Disponible en Internet: <http://www.w3.org/TR/soap12-part1/>
- SRU Search/Retrieval via URL* [en línea], January 13, 2012. [Fecha de acceso 2 febrero 2012]. Disponible en Internet: <http://www.loc.gov/standards/sru/>
- TILLET, B. B. *RDA: antecedentes y aspectos de su implementación. Manual para instructor* [en línea], actualizado en enero 2010. Washington: Library of Congress, 2009. [Fecha de acceso 03 febrero 2012]. Disponible en Internet: http://www.loc.gov/catdir/cpso/RDA/RDAantecedentes_instructor.pdf
- WEBER, M. B. y FAY, A.A. *Describing electronic, digital, and other media using AACR2 and RDA*. New York: Neal-Schuman Publishers, 2011. ISBN 1-55570-668-5.