

Vocabularios controlados y FAST: la asignación de materias en el siglo XXI

MARILYN MONTALVO MONTALVO
Universidad de Puerto Rico

VOCABULARIOS CONTROLADOS

Uno de los aspectos más fascinantes y menos comprendidos de la catalogación bibliográfica es el desarrollo de vocabularios controlados, tipificados o normalizados.¹ El propósito del lenguaje controlado es agrupar diversos recursos informativos bajo un mismo concepto o nombre propio para reducir las opciones de búsqueda en catálogos e índices y facilitar la recuperación de información. En la configuración de dichos vocabularios se conjugan, entre otras disciplinas, la lexicografía, la semántica, la lexicología, la morfología, la sintaxis, la taxonomía y la lógica.

La asignación de vocablos controlados es subjetiva. En algunos casos pueden seleccionarse términos de uso popular o términos científicos, dependiendo del tipo de usuario a quien vayan dirigidos. Además los vocablos no se seleccionan al azar. Los puntos de vista sociales, culturales y políticos también tienen un peso importante en la organización y búsqueda de información. Por otra parte, aunque los vocabularios normalizados aspiran a alcanzar un gran nivel de uniformidad, rara vez se obtiene el nivel óptimo y hay que admitir que

1 Los sustantivos 'normalizar' y 'tipificar', de los cuales derivan los adjetivos normalizado y tipificado, describen en castellano el proceso por el cual se elijen artificialmente las entradas que representan conceptos diferentes. Hemos elegido el anglicismo 'controlado' por su extenso uso y porque una de sus acepciones incluye también el concepto de 'regularización'.

ningún índice o catálogo temático es internamente tan consistente como uno desearía (OCLC 2007, p. 21). Además el nivel de especificidad de la catalogación lo determina cada unidad catalográfica.

Los epígrafes elegidos, denominados autoridades de materia, tomados de diccionarios, tesauros y de las propias obras catalogadas, pueden ser generales y abarcar varios aspectos de un tema y reducen las opciones que tendrá el usuario al momento de buscar información; o bien pueden ser específicos con el propósito de localizar subtemas, lo que aumenta tanto las opciones de búsqueda como las entradas de la lista de encabezamientos desarrollada. A su vez, el alcance de una lista de encabezamientos de materia en cuanto a las disciplinas que comprenderá y el balance que tendrá en la cobertura de los temas y su actualización, puede formar parte de una estructura preestablecida o construirse si hacen falta nuevos epígrafes.

El significado de los vocablos, simples o complejos, el número, la homonimia y la sinonimia son algunos de los elementos que intervienen en la selección, desarrollo y desambiguación de los encabezamientos de materia. En la construcción de los epígrafes también intervienen el orden de los elementos que los constituyen, ya sea natural o inverso, y el tipo de sintagma nominal, así como la especificación del tema lo cual se hace mediante el uso de paréntesis o subdivisiones. Asimismo la ubicación cronológica y geográfica del tema, así como el formato o género de los recursos, contribuyen a enmarcar la descripción del quehacer humano en sus diversos contextos. Todos los componentes que intervienen en la estructura de una lista de encabezamientos deben desarrollarse lógicamente para permitirles a los usuarios comprender y anticipar los epígrafes utilizados en el catálogo o índice consultado, y también permitirles a los catalogadores de la entidad catalográfica mantener la normalización de los epígrafes usados, por otra.

Una de las mayores fortalezas de los vocabularios controlados es su capacidad para reunir términos sinónimos en un epígrafe unívoco al cual se remiten los términos no utilizados como autoridad. Otra de sus fortalezas es la capacidad de expresar conceptos complejos mediante la construcción de cadenas de epígrafes organizadas jerárquicamente. En un esquema temático enumerativo el vocabulario

contiene encabezamientos previamente contruidos para describir tanto temas simples como complejos. En cambio en el esquema sintético solamente figuran en la lista de epígrafes algunos de los posibles temas y los restantes se construyen mediante reglas. La precoordinación permite construir encabezamientos en sentido horizontal y los hace más específicos mediante la asignación de subdivisiones. En cambio, la postcoordinación produce encabezamientos múltiples que se construyen en sentido vertical.

Según Svenonius (2000, pp.188-189), las ventajas de los encabezamientos precoordinados se pueden resumir analizando cuatro criterios básicos: la especificidad, la precisión, la contextualidad y la capacidad de sugerir resultados más específicos. Todos estos criterios apuntan hacia la fortaleza de los epígrafes precoordinados cuando se quieren obtener más aciertos en la búsqueda de información. Sin embargo, aunque no prevalecieron en su momento, visiones posteriores las han determinado como la forma más idónea para organizar un sistema de epígrafes controlados. La Clasificación Colon, de S. R. Ranganathan se proponía clasificar el conocimiento a partir de sus partes constituyentes, como si fuera un mecano, mediante el establecimiento de epígrafes discretos y cinco facetas básicas: personalidad (quién), materia (qué) energía (cómo), espacio (dónde) y tiempo (cuándo). Su objetivo era posibilitar la combinación de conceptos libremente y obviar las restricciones impuestas por un patrón preconcebido (Foskett 1981, p. 402). A su vez, PRECIS combinaba el análisis intelectual, realizado por personas, con el análisis realizado por computadoras, lo que permitía que los bibliotecarios seleccionaran, normalizaran y categorizaran los conceptos y que las computadoras asignaran la sintaxis automáticamente (Svenonius 2000, pp. 174-177). Otro sistema basado en facetas es el Bliss Classification basado en las facetas de Ranganathan, que utiliza trece facetas y se organiza de una forma progresiva que va de lo general a lo particular en un orden fijo y fácil de recordar (Bliss 2007).

Hasta finales de los años 80 la elaboración de epígrafes de materia estuvo directamente vinculada con los catálogos de tarjetas y los índices impresos. En este mundo se volvió esencial la selección del vocablo o nombre que encabezaría el epígrafe, el orden de los

subencabezamientos y la asignación de contrarreferencias. Pero era igualmente importante la economía del espacio. Cuando los usuarios no tenían otra alternativa que consultar catálogos e índices impresos, debían tratar de entender la lógica de la estructura de los epígrafes precoordinados para acceder a la información. Sin embargo, poco se sabe sobre el éxito que tenían los usuarios en sus búsquedas temáticas. Lo que sí es un secreto a voces es que muchos de los bibliotecarios que ofrecían servicios al público acogieron con complacencia la opción de búsqueda por palabra clave de los catálogos en línea, que les permitía usar un lenguaje más natural y combinar los términos de búsqueda de forma postcoordinada. Además, liberaba, tanto a los bibliotecarios como a los usuarios de una estructura que, no pocas veces, los dejaba *in albis*.

A partir de los años 90, los sistemas automatizados comenzaron a recuperar no solamente las autoridades de materia en cualquier orden, sino todas las palabras léxicas utilizadas en la descripción de los recursos informativos. Actualmente la indización automática íntegra de los documentos permite encontrar en diversos medios electrónicos, sin necesidad de pasar por procesos de normalización, todo un universo bibliográfico utilizando una palabra, varias palabras o enunciados completos. Sin embargo, la falta de normalización y de análisis de los contenidos da al traste con la síntesis y la desambiguación lograda por los vocabularios controlados y reduce la precisión de los resultados. A pesar de su agilidad, los motores de búsqueda no permiten recuperar sistemáticamente los mismos recursos, ni pueden categorizar las expresiones y manifestaciones de una obra. Ha sido precisamente la creación y población sin límites del WWW lo que ha catapultado el término metadatos, utilizado para referirse a los datos descriptivos y temáticos que se le incorporan a un documento electrónico con la intención de mejorar los resultados obtenidos por los motores de búsqueda mediante el rastreo de su texto. Esta aparente "reinención de la rueda", desde el punto de vista bibliotecológico, surge de la necesidad de poner un poco de orden en el caos del WWW. En lo que a los epígrafes de materia se refiere, el Dublin Core Metadata Initiative, cuyos campos descriptivos se han convertido en el Estándar ANSI/NISO Z39.85 y en el Estándar IETF- RFC 5013 (agosto 2007), sugie-

re que se usen vocabularios controlados para designar los temas de los recursos electrónicos. El esquema, que al principio parecía que socavaría las reglas de catalogación y los vocabularios controlados ha ido evolucionando tanto que, como bien ha señalado Karen Coyle: "We may see that when metadata grows up, it becomes cataloging" (podemos creer que cuando los metadatos crezcan se convertirán en catalogación, Trad. del C.) (Coyle 2005). Sin embargo, la posibilidad de que pueda desarrollarse un sistema de catalogación que forme parte del entramado de Internet parece estar aún lejos; de acuerdo con un estudio reciente los motores de búsqueda de Internet no están teniendo en cuenta los metadatos en sus fórmulas de búsqueda, razón por la cual el impacto de los metadatos del Dublin Core en la recuperación de páginas electrónicas es nula (Safari 2005).

EL USUARIO COMO EJE

Dos de los teóricos más importantes del siglo XX se han expresado contundentemente sobre el propósito principal de las bibliotecas y los catálogos. Cutter (1904) establece los siguientes propósitos:

Charles A. Cutter

Reglas Del Catálogo Diccionario

- ⇒ Permitirle a una persona encontrar un libro acerca del cual conoce el autor, el título o el tema.
- ⇒ Mostrar los libros que una biblioteca tiene escritos por un determinado autor, sobre un tema o en algún tipo de género literario.
- ⇒ Ayudar a seleccionar un libro en cuanto a sus características bibliográficas y en cuanto a su contenido.

Según Cutter, la conveniencia del usuario debe prevalecer sobre la del catalogador. Aunque una regla sin excepciones debería ser fácil de establecer y de entender, en su aplicación la uniformidad estricta puede conducir a prácticas que no coincidan con la forma en que los usuarios suelen buscar información. Los principios de Ranganathan (1931), también expresan de forma llana la misma tendencia:

S.R. Ranganathan

Cinco leyes de la bibliotecología

- ⇒ Los libros son para usarse.
- ⇒ A cada lector su libro.
- ⇒ A cada libro su lector.
- ⇒ Ahórrele tiempo al lector.
- ⇒ La biblioteca es un organismo en constante crecimiento.

Estos principios que pueden aplicarse a todo tipo de recurso informativo tienen un denominador común: el usuario. Según Ranganathan, la organización de la información debe siempre tener como eje la perspectiva del usuario y el dinamismo de las bibliotecas.

El requisito de tener al usuario como punto central de la organización de la información ha formado parte de la política de catalogación de la Biblioteca del Congreso de los Estados Unidos desde sus inicios (Chan 2005, p. 7) y es el más importante de los objetivos de la elaboración de códigos de catalogación, formulados en la versión más reciente del *Statement of international cataloguing principles* (IFLA 2008).

Nuevos Principios Internacionales de Catalogación

Versión de abril de 2008

- ⇒ Conveniencia del usuario.
- ⇒ Significatividad.
- ⇒ Uso común.
- ⇒ Economía.
- ⇒ Representatividad.
- ⇒ Normalización.
- ⇒ Exactitud.
- ⇒ Integración.
- ⇒ Precisión.

La “conveniencia del usuario” y el “uso común” coexisten con otros objetivos dirigidos hacia la formulación de un sistema estructurado. Sin embargo es muy difícil conocer la lógica de los usuarios, porque aunque las personas acostumbran buscar información de forma estructurada en listas temáticas de diversa naturaleza, es difícil precisar

cómo se enfrentan a los encabezamientos de materia que no utilizan el lenguaje natural y que se organizan jerárquicamente, en cadenas de palabras.

LIBRARY OF CONGRESS SUBJECT HEADINGS

Los *LCSH* (Encabezamientos de Materia de la Biblioteca del Congreso) publicados por primera vez en 1914 como los *Subject Headings Used in the Dictionary Catalog of the Library of Congress* constituyen la lista de epígrafes más completa que se conoce. En la introducción a su cuarta edición, el jefe de la División de Catalogación Temática de la LC, David Judson Haykin, señalaba que los epígrafes no se habían basado en un esquema básico al que se le pudieran seguir añadiendo otros temas. Tampoco proveían la base teórica para desarrollar un sistema de encabezamientos de materia. Sin embargo parece que se basaban en una lista publicada por la American Library Association, dirigida a bibliotecas públicas pequeñas o medianas, que se publicaría como apéndice de las reglas de Cutter. Aun así, la influencia de Cutter en el desarrollo de los encabezamientos de materia de la LC apenas ha sido reconocida, aunque los principios que la guían: el usuario como eje, el uso general y la especificidad, coinciden con los de Cutter (Chan 2005, p. 5-7). Además de estos principios básicos, la Biblioteca del Congreso ha basado el desarrollo de sus epígrafes en los conceptos de “garantía literaria”, uniformidad y unicidad de los epígrafes, consistencia interna, estabilidad, entrada directa, precoordinación y postcoordinación (Chan 2005, p. 34-38).

A pesar de su trasfondo un tanto incierto, los *LCSH* ya cuentan con una trigésima edición que contiene 280,000 encabezamientos y contrarreferencias en cinco volúmenes impresos. Según O’Neil y Chan (2003, p. 337), los *LCSH* se distinguen por:

- ❖ Ser el vocabulario controlado más extenso en lengua inglesa.
- ❖ Tener un gran riqueza léxica y temática.
- ❖ Proveer control de sinonimia y homonimia.
- ❖ Contener referencias jerárquicas y asociativas.

III Encuentro de Catalogación y Metadatos

- ❖ Usarse extensamente en otras instituciones, nacionales e internacionales.
- ❖ Disponer de numerosas traducciones y adaptaciones.
- ❖ Tener una historia documentada.
- ❖ Tener el apoyo de la Biblioteca del Congreso de los E.U.

Los *LCSH* se utilizan no solamente en Estados Unidos sino en muchos países de habla inglesa. Además su extensa difusión demostró en una encuesta realizada con los auspicios de IFLA, que 24 bibliotecas nacionales, principalmente anglófonas, usaban los *LCSH* y 12 usaban traducciones o adaptaciones (Heiner-Freiling 2000). Los encabezamientos de LC han sido usados ampliamente como modelo de las listas de encabezamientos iberoamericanas. En una muestra en la que se examinaron 1,543 epígrafes en inglés encontrados en libros publicados en español en 1992, extraída de la base de datos de OCLC y comparada con los epígrafes incluidos en la edición de 1978 de la *Lista de encabezamientos de materia*, de Gloria Escamilla González, se encontró que el 60% de ellos tenía un equivalente exacto o casi exacto en la lista de Escamilla y el 2% tenía una versión adaptada (Álvarez 1993, p. 29). La Dirección General de Bibliotecas de la Universidad Nacional Autónoma de México, que tiene a su cargo la catalogación centralizada de más de cien bibliotecas departamentales, utiliza, tanto las traducciones de los términos del *LCSH* como la lista de Escamilla. Además, la lista de Escamilla ha sido utilizada en Cuba y Costa Rica como fuente para desarrollar encabezamientos de materia nacionales y también ha sido utilizada para elaborar BILINDEX (Martínez Arellano 2004). Otra obra, también basada en los *LCSH*, que ha sido extensamente usada es la *Lista de encabezamientos de materia para bibliotecas*, compilada por Rovira y Aguayo en 1967, revisada por el Instituto Colombiano para el Fomento de la Educación Superior (ICFES) y, más recientemente, por la Biblioteca Luis Ángel Arango del Banco de la República de Colombia. Los *LCSH* también han sido utilizados como obra de referencia fundamental para compilar encabezamientos de materia en España. Muestra de ello es la *Lista de encabezamientos de materia de la Red de Bibliotecas del CSIC*, publicada en el 2000. Otro ejemplo es la Biblioteca de la Universidad

de Alicante, que utiliza como base para indizar sus registros bibliográficos los *Encabezamientos de Materia de la Biblioteca Universitaria de Sevilla*, basados, a su vez, en los *LCSH* (SIBID 2003).

Tanto los *LCSH*, como sus adaptaciones, comparten dos de los rasgos que los caracterizan: su naturaleza sintética y la precoordinación. Los bibliotecarios que utilizan los *LCSH* tienen que consultar numerosas herramientas para determinar la sintaxis más adecuada para cada epígrafe, pues las reglas establecidas para la especificación de los epígrafes por medio de subdivisiones no son uniformes.² Además, no todos los encabezamientos están incluidos, ya sea en la propia lista, o en listas ajenas. Tal es el caso de los epígrafes denominados “múltiples”, en los que se ofrece un ejemplo como patrón que puede ser sustituido por cualquier vocablo de la misma clase (O’Neill y Chan 2003, p. 337).

Uno de los aspectos de la construcción de epígrafes que más defienda la Biblioteca del Congreso (2007), es la precoordinación porque permite jerarquizar y establecer relaciones entre los temas, provee un orden estandarizado y mejora la relevancia de los resultados. Sin embargo tanto los usuarios como los bibliotecarios que no son catalogadores tienen que interpretar el orden jerárquico de los epígrafes y sus correspondientes subdivisiones, así como adivinar el tipo de sintaxis elegida por los catalogadores. Entre otros autores, Ballard (1998, p. 58) sostiene que los usuarios no han logrado aprender la sintaxis de los *LCSH*. Con base en sus observaciones estima que la mitad de las

2 Actualmente los *LCSH* disponen de las siguientes herramientas adicionales que les son necesarias para asignar epígrafes y subdivisiones: *Free-floating subdivisions: an alphabetical index*, 20th edition (2008), *Subdivisiones de Libre Flotamiento: eun índice alfabético*, edición 20 (2008), *LC period subdivisions under names of places*, 5th edition (1994), (*Subdivisiones de periodo de la LC bajo nombres de lugares*, 5a. edición 1994), *Subject cataloging manual: subject headings, updates; Subject cataloging manual: subject headings (2008 update); LC subject headings principles of structure and policies for application* (1990); *LC subject headings weekly lists on CPSO Web Page; MARC Distribution Service: subject authorities Manual de Catalogación Temática: encabezamientos de materias, actualizaciones; Manual de catalogación temática: encabezamientos de materias (actualización 2008); Principios, estructura y políticas de aplicación de la catalogación temática de la L.C. (1990); Listas semanales de encabezamientos de materia de la LC en la pág. Web de la CPSO; el Servicio de Distribución MARC de autoridades temáticas, T. del C.)* .

búsquedas que se hacen en los catálogos en línea no arrojan ningún resultado. Según Chan, los *LCSH* resultan muy engorrosos e incluyen muchas reglas complicadas para formar cadenas de encabezamientos (Chan 2005, p. 13). En un estudio en el que participaron 144 niños y 144 adultos estadounidenses, realizado con el propósito de determinar en qué medida entendían los epígrafes con subdivisiones, solamente el 31% de los niños y el 39% de los adultos interpretó correctamente los epígrafes (Drabenstott, Simcox y Fenton 1999). Por otra parte en una investigación similar en la que participaron 137 bibliotecarios que ofrecen servicios al público y 135 que laboran en servicios técnicos, el 52% de los bibliotecarios del primer grupo y el 55% de los bibliotecarios del segundo grupo interpretaron correctamente el significado de los epígrafes (Drabenstott, Simcox y Williams 1999). Estos resultados llevaron a los autores a sugerir, o bien que se estandarizara el orden de los epígrafes, o que se deconstruyeran los epígrafes muy largos, como medidas para simplificar la sintaxis de los *LCSH*. Ya Cochrane (1986, p. 62) había planteado la necesidad de que las cadenas de encabezamientos formulados por los *LCSH* se pudieran permutar porque la mayoría de los usuarios y algunos catalogadores no podían seguir la lógica de la construcción de los epígrafes. Tanto las críticas de los expertos como el resultado de estas investigaciones, demuestran que los *LCSH* no han estado apoyando adecuadamente uno de sus preceptos: la conveniencia del usuario.

Actualmente existen en Internet numerosas herramientas que les permiten a las personas formular temas en diversas instancias de su interacción con la Red. A este tipo de catalogación temática esta vez se la ha denominado “social tagging” (“etiquetación social”). Con el propósito de sondear las preferencias de las personas que les asignan temas a sus páginas favoritas, utilizando la plataforma social de *Delicious*, seleccionamos tres páginas de Internet citadas por los usuarios (una en español, una en inglés y una en francés) y examinamos, al azar, trescientos epígrafes de cada una. Más del 80% de los encabezamientos asignados a cada una de las páginas de la muestra de *Delicious* fueron epígrafes sencillos que se pueden combinar libremente con otros. En muy pocas ocasiones se eligió la opción de formular epígrafes complejos o con subdivisiones. Si bien la lógica y la

conveniencia de las personas a la hora de organizar y buscar información no es fácil de determinar y probablemente cambia de generación en generación, esta pequeña cala nos muestra una tendencia a la simplificación (Véase *Tabla 1*).

Tabla 1

TAGS´ (Etiquetas) simples y complejas o con subdivisión en una muestra de Delicious

Página	Tags´ (etiquetas) examinadas	Tags´ sencillas	Tags´ complejas o con subdivisión	% Tags´ sencillas
Comunidades de wikis libres para aprender (citada 1,177 veces)	300	286	14	95.33%
Country profiles (perfiles del país) (citada 343 veces)	300	272	28	90.67%
Joconde (citada 140 veces)	300	251	49	83.67%
TOTAL	900		91	89.89%

Muestra tomada el 21 de septiembre de 2008.

FAST

El desarrollo y la utilización de vocabularios controlados que se dan en una época de cambios en los principios internacionales de catalogación, en los requisitos funcionales de los registros bibliográficos y en las reglas de catalogación angloamericanas, y en la que las propuestas para el desarrollo de sistemas de “metadatos” convergen cada vez más con los estándares catalográficos, estos cambios quedan también enmarcados por las siguientes realidades:

- ❖ El crecimiento de todo tipo de recursos informativos.
- ❖ El crecimiento exponencial de los recursos electrónicos y su carácter ecléctico, inestable y escurridizo.
- ❖ La exigencia, por parte de los usuarios, de contar con sistemas de recuperación de información cada vez más rápidos.
- ❖ La necesidad de incorporarles esquemas de metadatos a los recursos disponibles en Internet para agilizar su localización.

III Encuentro de Catalogación y Metadatos

- ❖ La escasez de bibliotecarios catalogadores, debido a la reducción del currículo dedicado a la enseñanza de la catalogación en las escuelas de biblioteconomía.
- ❖ Las dificultades presupuestarias.

En este contexto, el Ohio Computer Library Center (OCLC) comenzó a explorar en 1998 un sistema de encabezamientos de materia que pudiera utilizarse en los registros descritos mediante los campos diseñados por el Dublin Core Metadata Initiative. A pesar de que el DCMI sugería la utilización de vocabularios normalizados, no especificaba qué tipo de listas de encabezamientos o tesauros se recomendaban. El OCLC determinó que para que una lista de epígrafes de materia fuera apropiada para proveer acceso a los recursos electrónicos debería:

- ❖ Tener una estructura simple de asignar, utilizar y mantener.
- ❖ Poder ser manejada por personas que no sean catalogadores y en otros ambientes, además de las bibliotecas.
- ❖ Proveer puntos de acceso óptimos.
- ❖ Ser flexible y compatible con diferentes disciplinas y bases de datos.
- ❖ Ser compatible con MARC, Dublin Core y otros esquemas de descripción bibliográfica.
- ❖ Ser fácil de mantener y manejar, desde el punto de vista tecnológico.

El desarrollo de un vocabulario controlado de esta naturaleza podría lograrse de dos formas: creándolo o usando una lista existente. El producto de la investigación realizada por OCLC ha sido FAST, un vocabulario normalizado, extraído de los *LCSH*, aunque con algunas modificaciones, que categoriza los encabezamientos según diversas facetas manteniendo la riqueza de los *LCSH*, a la vez que utiliza una sintaxis más sencilla (OCLC 2007). FAST, aun en etapa experimental, contiene rasgos de precoordinación y de postcoordinación. Es precoordinado porque contiene encabezamientos con subdivisiones, siempre y cuando representen la misma faceta. Es postcoordinado porque en los casos en que no se han desarrollado encabezamientos complejos o en los que los encabezamientos no pertenecen a la misma faceta se establecen encabezamientos adicionales. FAST dispone

actualmente de una base de datos de autoridades (OCLC 2008) y de un manual preliminar (OCLC 2007). La base de datos contiene dos archivos en los que figuran los registros de Asuntos,³ Lugares, Tiempo, Personas, Entidades, Títulos uniformes y Acontecimientos,⁴ por una parte, y los de Forma o género, por otra. FAST adopta un formato modular en el que cada faceta forma grupos discretos de epígrafes. La inclusión de los epígrafes en su base de datos se basa en el concepto de garantía literaria (*literary warrant*), ya que se incluyen aquellos que han sido usados en *Worldcat* como tema.

Los encabezamientos **Temáticos sobre Asuntos** de FAST constan de un epígrafe simple o complejo, directo o invertido y sus correspondientes subdivisiones, por ejemplo:

Global warming—Health aspects	Astronauts—Psychology
Acting—Study and teaching	Adult education—Administration
Soy bran—Marketing	Biography—Bibliography
Chemistry, Organic	Willy (Dolphin)
Civil War (Spain : 1936-1939) ⁵	Scarlett O'Hara (Fictitious character)
Special Olympics ⁶	Macondo (Imaginary place)
Cenozoic Geologic Period	Abolitionists—Political activity

-
- 3 Hemos traducido 'Topics' por 'Asuntos'. En el manual de FAST, los encabezamientos de asuntos se subdividen en: *Concepts, Objects, Events, Form/Genre as subjects, Animals, Imaginary Persons, Places, etc.* y *Geologic Periods*.
 - 4 En el manual de FAST los **Eventos** forman parte de los encabezamientos temáticos sobre asuntos. En los *LCSH* algunas subdivisiones cronológicas contienen el nombre del acontecimiento. FAST separa los componentes temáticos y cronológicos porque pertenecen a dos facetas diferentes. En aquellos casos en que el nombre del acontecimiento es genérico, se especifica el lugar y fecha por medio de un delimitador.
 - 5 Durante el proceso de adaptación de los epígrafes de los *LCSH* se han modificado algunos epígrafes y se han hecho excepciones a las reglas establecidas para los acontecimientos históricos. Por ejemplo, el encabezamiento *Civil War*, en el caso de la Guerra Civil de los Estados Unidos, se ha convertido en *American Civil War (1861-1865)*.
 - 6 Los acontecimientos deportivos, reuniones, conferencias, etcétera, no se delimitan por lugar o tiempo, sólo por tipo, de ser necesario.

III Encuentro de Catalogación y Metadatos

Los encabezamientos **Temáticos Geográficos**⁷ pueden ser Jurisdiccionales o no Jurisdiccionales y se pueden dividir en hasta tres niveles de especificidad:

Epígrafes Jurisdiccionales	Epígrafes No Jurisdiccionales
Puerto Rico—San Juan ⁸	Puerto Rico—Yunque Mountain
México	Costa Rica—Lake Arenal Region
Spain	Caribbean Sea ⁹
Massachusetts	Colombia—Bogotá—Ciudad Bolívar
Congo (Democratic Republic)	

Los encabezamientos **Temáticos Cronológicos** constan de fechas de inicio, término o de inicio y término. Los encabezamientos que representan siglos se expresan en números. FAST no crea encabezamientos cronológicos sistemáticamente sino sólo cuando se necesitan como referencia, por ejemplo:

1980	1900-1999
Since 1959	From 10 thousand to 2 million years ago
To 1989	2004 (March 11)
2004-2008	146 B.C.-323 A.D.

Los encabezamientos **Temáticos Personales** contienen nombres, apellidos, numeración, títulos, delimitadores y fechas. También incluyen los nombres de familias, dinastías, etcétera, pero no se subdividen, por ejemplo:

-
- 7 El manual de FAST incluye en sus anejos los nombres geográficos del primer nivel, derivados del *MARC Code List for Geographic Areas (Lista de Códigos para las Áreas Geográficas)*, aunque con ciertas modificaciones, y los delimitadores geográficos. A diferencia de los *LCSH*, FAST les asigna el primer nivel a ciertas jurisdicciones secundarias. Tal es el caso de algunas ciudades de Estados Unidos, Canadá, Gran Bretaña y Australia.
 - 8 Los nombres de lugares locales se asignan en orden indirecto, es decir colocando la jurisdicción mayor en primer lugar.
 - 9 Hay encabezamientos que se subdividen y otros que no. En el caso de *Caribbean Area* y *Caribbean Sea*, que representan una zona similar, FAST solamente subdivide *Caribbean Sea*.

Los encabezamientos **Temáticos de Entidades**¹⁰ contienen el nombre de la entidad principal, los nombres de entidades subordinadas y los delimitadores, por ejemplo:

Biblioteca Nacional de México	Puerto Rico. Legislature. Senate
World Bank	Biblioteca Agrícola Nacional (Perú)
United Nations. Committee on Disarmament	

Los encabezamientos **Temáticos de Títulos** se componen del título uniforme de las obras. Si la obra es anónima solamente se incluye el título; si tiene autor, se incluye el autor. Los epígrafes de título no contienen el idioma, versión o fecha. Los epígrafes de títulos incluyen obras anónimas, así como obras de autores personales, de entidades u obras pertenecientes a conferencias, congresos, etcétera, por ejemplo:

Cantar de mio Cid	Bible. N.T. Epistles
Coronel no tiene quien le escriba (García Márquez, Gabriel)	General act (Berlin West Africa Conference)
Education for All Handicapped Children Act (United States)	Bible. N.T. Epistles

Los encabezamientos **de Forma o Género**¹¹ constan de un encabezamiento principal y se pueden subdividir, por ejemplo:

10 Las autoridades onomásticas de FAST han sido extraídas de NACO (Name Authority Cooperative Program of the PCC).

11 Los epígrafes de **Forma** aparece en la base de datos de autoridades de FAST como encabezamiento principal y como subdivisión temática, tal y como se usa en los *LCSH*. Están incluidos, tanto en el *Fast authority file* como en los *Fast forms*. El manual de FAST no aclara que coexistan como subdivisiones temáticas y como facetas de forma. Solamente ofrece un ejemplo con *periodical*, en el que indica que puede ser un encabezamiento temático, si se refiere a obras sobre revistas, y de forma, para una obra que es una revista. Jin (2008, p. 204) menciona que *history* se usa como forma cuando sigue a un encabezamiento geográfico, pero tampoco indica si se usa como subdivisión. Aparentemente, la conversión de epígrafes de los *LCSH* a FAST ha producido algunos conflictos que aún no se han aclarado. Esto podría obedecer a que muchas subdivisiones temáticas con subcampo \$x, que debían haberse cambiado a un subcampo \$v, pasaron a FAST como subdivisiones temáticas y también se incluyeron como facetas de forma. Lo cierto es que esta clasificación es confusa.

III Encuentro de Catalogación y Metadatos

En ocasiones es necesario asignar un término delimitador¹² entre paréntesis para:

Resolver problemas de ambigüedad.	Indicar el género de un nombre propio.
Aclarar términos técnicos.	Delimitar un concepto general.
Indicar el género de un nombre propio	Especificar el tipo de ejecución en un encabezamiento musical.
Resolver problemas de ambigüedad.	Distinguir nombres iguales de lugares diferentes.
Aclarar términos técnicos.	Distinguir nombres de acontecimientos.
Identificar términos de otras lenguas que no sea el inglés	

En el manual de FAST se establece que la selección de los epígrafes debe guiarse por el principio de especificidad. A las obras que tratan de un tema solamente se les asigna el encabezamiento más específico. A las obras que contienen más de un tema se les asignan tantos temas como sean necesarios. A las obras que contienen temas complejos se les asignan encabezamientos formados por frases o con subdivisiones si están disponibles en la base de datos de forma precoordinada; de lo contrario se les asignan encabezamientos independientes.

FAST es la culminación de los deseos de diversos sectores del campo de la biblioteconomía y la información por establecer un vocabulario controlado que utilice una sintaxis sencilla, que sea enumerativo y que sea relativamente fácil de aplicar, lo que les permite a los usuarios construir un número de encabezamientos de forma postcoordinada. Hasta julio de 2008 FAST tenía un total de 1,669,375 autoridades de asuntos, personas, entidades, lugares, periodos, acontecimientos, títulos y formas (Childress 2008). Según O'Neil y Chan (2006), la simplificación de su estructura ha permitido reducir los encabezamientos de asuntos, aproximadamente, de 9,000,000 en *LCSH*, a 400,000 en FAST. Además, reduce la cantidad de encabezamientos poco usados.

Se han llevado a cabo varios proyectos para determinar la viabilidad de FAST. Uno de ellos ha sido un estudio realizado para determinar la posibilidad de añadirle epígrafes de FAST a los encabezamientos de materia del *Ulrich's Periodicals Directory*, para ofrecerles a sus usua-

12 Los delimitadores no son enumerativos, el catalogador los tiene que añadir.

rios la riqueza de los LCSH. En el estudio, 96 de los 100 temas analizados al azar tuvieron un equivalente en FAST, lo que demuestra que los epígrafes de *Ulrich's* podrían sustituirse por los de FAST (Mitchell y Hsieh-Yee 2007). Por otra parte el Subject Analysis Committee de la Association of Library Collections and Technical Services (ALCTS) examinó en 2005 una muestra de 5,000 epígrafes de LC convertidos a FAST para determinar si los encabezamientos organizados en facetas eran comprensibles por los usuarios y si el desarrollo de FAST iba en la dirección correcta. El Comité llegó a la conclusión de que la mayoría de los encabezamientos temáticos reflejaba el tema del recurso catalogado. Sin embargo, al ser deconstruidos algunos epígrafes tenían un sentido muy general y podían perder el significado específico que tenían cuando estaban precoordinados (Jin, Quiang 2008).

El cambio mayor que plantea FAST es la organización de los LCSH por facetas. Pero evidentemente mantiene la estructura interna de los sintagmas nominales establecidos en los LCSH con todas sus fortalezas y debilidades (Chan 2005, p. 13). Muchos de sus vocablos no son totalmente discretos, lo que implica que su uso requiere conocer cómo se construyen dichos epígrafes. Por eso Anderson y Hofmann (2006, p. 19-21) han propuesto que se les aplique a los LCSH el sistema de facetas de Bliss, y que se les ofrezcan a los usuarios las reglas combinatorias necesarias para hacer una construcción postcoordinada de los encabezamientos de materia. Paralelamente ya se están desarrollando prototipos de bases de datos y catálogos en línea, cuyas interfaces de búsqueda arrojan resultados agrupados en facetas, como lo demuestran los *Worldcat identities* (OCLC 2008) y muy especialmente el catálogo en línea de la Oakville Public Library (2008).

FAST Y LAS COLECCIONES DIGITALES DEL SISTEMA DE BIBLIOTECAS DE LA UPRRP

El Sistema de Bibliotecas del Recinto de Río Piedras de la Universidad de Puerto Rico ha estado catalogando e indizando diversas colecciones de imágenes durante más de una década. Algunas colecciones se han catalogado usando los LCSH en inglés, otras se han indizado

usando epígrafes en español, tomados de las diversas ediciones de la *Lista de encabezamientos de materia para bibliotecas*, de Rovira y Aguayo o creados localmente. Actualmente dichas colecciones han sido incorporadas a una base de datos utilizando la plataforma de *ContentDM*. Aunque es necesario normalizar los epígrafes no se ha tomado una decisión en cuanto al tipo de vocabulario controlado más conveniente, tanto para los usuarios como para que la catalogación de las colecciones de imágenes pueda agilizarse.

Quisimos auscultar los tipos de encabezamientos más usados en la catalogación o indización de las colecciones de imágenes del SB y la viabilidad de sustituir los encabezamientos utilizados actualmente por los epígrafes de FAST. Con este propósito seleccionamos 8 de sus 17 colecciones y elegimos los primeros 25 epígrafes utilizados en cada una de ellas. Según el análisis de las muestras, en 4 de las 8 colecciones, la mayoría de los epígrafes se componen de vocablos simples. En dos de las colecciones hay la misma cantidad de epígrafes simples que de epígrafes complejos o complejos con subdivisión. En una se utilizan mayormente epígrafes con subdivisión y en otra, epígrafes complejos con subdivisión (Véase la *Tabla 2*). Al examinar la construcción de los epígrafes en las cuatro colecciones que obtuvieron 50% o menos de epígrafes sencillos, encontramos lo siguiente: en la colección de la PRRA prevaleció el epígrafe complejo compuesto por sintagmas adjetivales; en la colección de Tarjetas Postales abundaron los sintagmas adjetivales con subdivisión geográfica; en la colección de Mapas Raros se utilizaron subdivisiones temáticas, geográficas o cronológicas; en la colección de Fotos se utilizaron las subdivisiones geográficas.

Posteriormente buscamos todos los epígrafes en la base de autoridades de FAST para determinar si podían parearse y luego los clasificamos por tipo de epígrafe. Todos los epígrafes parearon. Las mayores dificultades de pareo las presentaron los epígrafes que habían sido asignados incorrectamente. En 5 de las 8 colecciones la mayoría de los epígrafes se convirtieron en epígrafes sencillos. En dos de las colecciones la mayoría de los epígrafes se formuló con subdivisiones. En una de las colecciones hubo la misma cantidad de epígrafes complejos y sencillos (Véase *Tabla 3*). Al examinar la construcción de los epígrafes en las colecciones que obtuvieron menos del 50% de epígrafes senci-

Vocabularios controlados y FAST: la asignación de materias en el siglo XXI

llos, encontramos lo siguiente: en las colecciones Moscioni y Tarjetas postales abundaron los epígrafes geográficos con subdivisión, y en la colección de Caricaturas hubo una representación de diversos tipos de sintagma nominal.

<i>Tabla 2</i> Epígrafes usados actualmente en ocho colecciones de imágenes del sistema de bibliotecas						
Colección	A	B	C	D	E	F
Carteles	14	8	2	1	25	56%
Mario Brau	20	1	3	1	25	80%
Moscioni	13	3	4	5	25	52%
Caricaturas	15	9	0	1	25	60%
PRRA	11	11	0	3	25	44%
Tarjetas postales	7	0	1	17	25	28%
Mapas Raros	1	0	15	9	25	4%
Fotos	11	2	1	11	25	44%
A = Epígrafes sencillos (incluyendo nombres propios)			B = Epígrafes complejos			
C = Epígrafes con subdivisión			D = Epígrafes complejos con subdivisión			
E = Total			F = % Epígrafes sencillos			

<i>Tabla 3</i> Epígrafes de ocho colecciones de imágenes del sistema de bibliotecas convertidos a FAST						
Colección	A	B	C	D	E	F
Carteles	14	9	1	2	26	54%
Mario Brau	20	2	2	2	26	77%
Moscioni	5	7	16	0	28	18%
Caricaturas	13	13	2	0	28	46%
PRRA	13	5	7	1	26	50%
Tarjetas postales	2	3	24	0	29	7%
Mapas Raros	40	6	4	0	50	80%
Fotos	19	10	6	2	37	51%
A = Epígrafes sencillos (incluyendo nombres propios)			B = Epígrafes complejos			
C = Epígrafes con subdivisión			D = Epígrafes complejos con subdivisión			
E = Total			F = % Epígrafes sencillos			

III Encuentro de Catalogación y Metadatos

Aunque la muestra realizada ha sido pequeña, los resultados apuntan hacia la simplificación de los tipos de epígrafe y de su sintaxis. El siguiente ejemplo, tomado de la Colección de Mapas Raros nos muestra cómo la utilización de FAST eliminó la redundancia de los epígrafes de *LC* en un mismo registro y cómo simplificó la estructura precoordinada, al reagrupar los epígrafes según sus diversas facetas:

Título

A map of the West-Indies or the islands of America in the North Sea; with ye adjacent

countries; explaining [sic] what belongs to Spain, England, France, Holland and c. also

ye trade winds, and ye several tracts made by ye galeons and flota from place to place.

According to ye newest and most exact observations by Herman Moll, geographer.

Epígrafes con la estructura de los LCSH

Treasure-trove — Caribbean Area — Maps — Early works to 1800.

Treasure trove — West Indies — Maps — Early works to 1800.

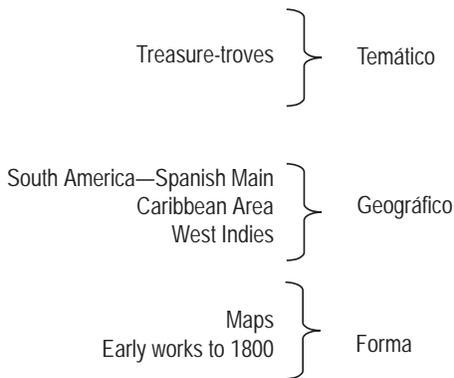
Treasure trove — Spanish Main — Maps — Early works to 1800.

Caribbean Area — Maps — Early works to 1800.

West Indies — Maps — Early works to 1800.

Spanish Main — Maps — Early works to 1800.

Epígrafes con la estructura de FAST



Durante el ejercicio, además, pudimos constatar que la base de datos de autoridades de FAST provee muchas opciones de búsqueda, lo que agiliza la consulta una vez se conoce su organización. Un catalogador o bibliotecario acostumbrado a formular o usar epígrafes de los *LCSH* o de los que provienen de listas basadas en los *LCSH* podría realizar el pareo con facilidad, pero una persona que no tuviera conocimientos de catalogación tendría que familiarizarse con la estructura de los encabezamientos de los *LCSH* y la estructura de FAST para poder localizar los epígrafes y traducirlos al español, si éste es el idioma elegido.

CONCLUSIÓN

No hay duda de que los vocabularios controlados tienen una gran capacidad para sintetizar y reunir términos sinónimos en epígrafes unívocos, lo que los hace idóneos para recuperar todo tipo de información. También es cierto que la complejidad de algunos vocabularios precoordinados limita su utilización y les resta agilidad a los procesos de sistematización de la información. FAST es un proyecto valiosísimo que si se consolida, permitirá mantener la riqueza léxica de los *LCSH* y simplificar su sintaxis al mismo tiempo que ofrecerá diversos componentes del mundo de la información tangible y virtual las ventajas de los vocabularios normalizados y hará efectivo el compromiso de los bibliotecarios con sus usuarios.

REFERENCIAS

- Álvarez, Fernando, 1993, *Comparative studies between the Lista de encabezamientos de materia by Gloria Escamilla and the Library of Congress Subject Headings List*. [s.l: s.n].
- Anderson, James D. y Hofmann, Melissa A., 2006, "A fully faceted syntax for Library of Congress Subject Headings", *en Cataloging and Classification Quarterly*, 43 (1), 7-37.

III Encuentro de Catalogación y Metadatos

Ballard, Terry., 1998, "Keyword/subject: finding a middle path", *Information Today*, 15 (6), 58.

The Bliss bibliographic classification: history & description, 2007, consultado el 28 de septiembre de 2008, tomado de <http://www.blissclassification.org.uk/bchist.htm>

Chan, Lois Mai., 2005, *Library of Congress subject headings: principles and application*, 4a. ed. Westport, Conn., Libraries Unlimited.

Childress, Eric., 2008, *FAST (Faceted Application of Subject Terminology): a vocabulary to facilitate faceted browsing*, Montreal, Université de Montréal, Présentation.

Cochrane, Pauline A., 1986, *Improving LCSH for use in online catalogs: exercises for self-help with a selection of background readings*, Littleton, Colo., Libraries Unlimited.

Coyle, Karen, 2005, "Understanding metadata and its purpose", en *Journal of Academic Librarianship*, 31, (2), 160-163, disponible en: <http://www.kcoyle.net/jal-31-2.html>

Cutter, Charles A., 1904, *Rules for the dictionary catalog*, 4a.ed. [Washington, D.C., GPO].

Drabenstott, Karen M., Schelle Simcox, and Eileen G. Fenton, 1999, "End-user understanding of subject headings in library catalogs", en *Library Resources & Technical Services*, 43 (3), 140-160.

—, Schelle Simcox, and Marie Williams, 1999, "Do librarians understand the subject headings in library catalogs?", en *Reference & User Services Quarterly*, 38 (4), 369-387.

Foskett, Anthony Charles, 1981, *The subject approach to information*, Hamden, Conn., Linnet Books.

Heiner-Freiling, Magda, 2000, "Survey on subject heading languages used in national libraries and bibliographies", en *Cataloging and Classification Quarterly*, 29 (1-2), 189-198.

- IFLA, 2008, *Statement of international cataloguing principles*, April 10, 2008 version, consultado el 15 de septiembre de 2008, tomado de http://www.ifla.org/VII/s13/icc/imeicc-statement_of_principles-2008.pdf
- Jin, Quiang, 2008, "Is FAST the right direction for a new system of subject cataloging and metadata?" en *Cataloging and Classification Quarterly* 45 (3), 91-110.
- Library of Congress. Cataloging Policy and Support Office, 2007, *Library of Congress subject headings: pre- vs. post-coordination and related issues report*, Washington, D.C., Library of Congress.
- Martínez Arellano, Filiberto Felipe, 2004, *Desarrollo de una lista de encabezamientos de materia en español*, Buenos Aires, IFLA, consultado el 10 de septiembre de 2008, tomado de http://www.ifla.org/IV/ifla70/papers/039s_trans-Arellano.pdf
- Oakville Public Library, 2008, consultado el 13 de octubre de 2008, tomado de <http://www.opl.on.ca>
- OCLC, 2007, FAST (Faceted application of subject terminology): applications guide and documentation, version 9, consultado el 9 de agosto, tomado de <http://www.oclc.org/info/research/wip/fast/manual-20070112.pdf>
- OCLC, 2008, *OCLC FAST test databases*, consultado el 13 de octubre de 2008, tomado de <http://fast.oclc.org/>
- OCLC, 2008, Worldcat identities, consultado el 13 de octubre de 2008, tomado de <http://orlabs.oclc.org/identities>
- O'Neil, Edward T. y Lois Mai Chan, 2003, "FAST (Faceted application of subject terminology): a simplified vocabulary based on the Library of Congress Subject Headings", en *IFLA Journal*, 29 (4), 336-342.
- . 2006, *A faceted LCSH based subject vocabulary*, New Orleans, ALA, Presentación.

III Encuentro de Catalogación y Metadatos

Ranganathan, S. R. 1931, *The five laws of library science*, Madras, Madras Library Association, London, Goldston.

Svenonius, Elaine, 2000, *The intellectual foundation of information*, Cambridge, Mass., MIT Press.

Universidad de Alicante, 2003, *Encabezamientos de materia en la UAL*, Alicante, Universidad de Alicante, SIBID, consultado el 30 de septiembre de 2008, tomado de www.ua.es/va/servicios/syf/selecfor/promocion_interna/2005/pi-01/documentacion/encab_mat_ual.doc