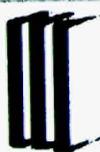




**Facultad de Humanidades y Ciencias de la Educación
Universidad Nacional de La Plata**

34



BIBHUMA
Biblioteca de Humanidades
"Prof. Guillermo Obiols"

<http://www.bibhuma.fahce.unlp.edu.ar>
bibhuma@fahce.unlp.edu.ar
Tel / Fax: 54-0221-423 5745
Calle 48 entre 6 y 7 - 1er. subsuelo



Facultad de Humanidades y
Ciencias de la Educación
Universidad Nacional de La Plata

en Línea

Martínez y Tamayo

La búsqueda por materia y la descripción de contenido en el catálogo en línea

Nº 34

Año 1998

COMITÉ EDITORIAL:

TITULARES:

DRA. MARÍA ELENA INFESTA

DRA. MARÍA JULIA BERTOMEU

DRA. MARÍA MALBRÁN

PROF. CAROLINA SANCHOLUZ

SRTA. BARBARA ROSSI

ALTERNOS:

DR. MIGUEL DALMARONI

PROF. LUISA GRANATO

PROF. GONZALO DE AMÉZOLA

PROF. VERÓNICA DELGADO

SRTA. NANCY DELLA ROSSA

SECRETARIA DE EXTENSIÓN:

PROF. ADRIANA BOFFI

DISEÑO DE TAPA:

ARQS. RUBÉN PUENTE / ADRIANA ROMERO

PAGINACIÓN ELECTRÓNICA:

PROF. MARÍA DEL ROSARIO MARTÍNEZ

E-MAIL: MARTINEZ@ISIS.UNLP.EDU.AR

DIAGRAMACIÓN:

JANE AVRIL COMUNICACIÓN EDITORIAL

530 NRO. 1160 "2", TEL. 225718, (1900) LA PLATA

E-MAIL: HF361@KANCA.INS.CWRU.EDU

Facultad de Humanidades y Ciencias de la Educación

Decano

Prof. José Luis de Diego

Vicedecano

Prof. Luis Viguera

Secretaría de Asuntos Académicos

Prof. Ana María Barletta

Secretario de Investigación y Posgrado

Dr. Julio César Moran

Secretaría de Extensión Universitaria

Prof. Adriana Boffi

Area de Asuntos Estudiantiles

Prof. César Arrondo

Area Capacitación Docente

Prof. Alicia Villa

Consejo Académico

Claustro Docente

Prof. Telma Piacente

Prof. Carlos Carballo

Prof. María Celia Agudo de Córscico

Dr. Fernando Enrique Barba

Prof. Rosa Pisarello

Prof. Alicia Alliau

Claustro de Graduados

Prof. Osvaldo Rodríguez

Prof. Claudio Suasnavi

Claustro Estudiantil

Miguel Nahon

Nancy Della Rosa

Silvia Guardia

Bernardo Raimondi

SERIE ESTUDIOS E INVESTIGACIONES

- Nº 1 FRONTERA Y JUSTICIA COLONIALES
- Nº 2 MERCADO DE TRABAJO Y PARO FORZOSO I
- Nº 3 MERCADO DE TRABAJO Y PARO FORZOSO II
- Nº 4 ESTUDIOS DE LÍRICA CONTEMPORÁNEA
- Nº 5 XII CONGRESO INTERAMERICANO DE FILOSOFÍA
- Nº 6 CUESTIONES AGRARIAS REGIONALES
- Nº 7 ESTUDIOS DE HISTORIA RURAL I
- Nº 8 LA PROBLEMÁTICA AGROALIMENTARIA EN LA ARGENTINA (1970-1988) T. I
- Nº 9 ESTUDIOS SOBRE BORGES
- Nº 10 TERRITORIO Y PRODUCCIÓN. CASOS EN LA REGIÓN METROPOLITANA EN BUENOS AIRES
- Nº 11 ESTUDIOS HISTORIA RURAL II
- Nº 12 MITOS, ALTARES Y FANTASMAS
- Nº 13 ESTUDIOS DE HISTORIA COLONIAL
- Nº 14 TRANSPORTE. ESPACIOS PERIURBANOS
- Nº 15 ESTUDIOS DE HISTORIA RURAL III
- Nº 16 TEMAS DE HISTORIA ARGENTINA I
- Nº 17 EL NUDO CORONADO. ESTUDIO DE CUATRO CUARTETOS.
- Nº 18 ESTUDIOS DE LÍRICA LATINA
- Nº 19 HISTORIA Y HUMANIDADES
- Nº 20 MERCADO DE TRABAJO Y CONSUMO ALIMENTICIO EN LA ARGENTINA AGROEXPORTADORA
- Nº 21 HOMENAJE A MANUEL PUIG
- Nº 22 IGLESIA, SOCIEDAD Y ECONOMÍA COLONIAL
- Nº 23 PSICOLOGÍA: DOCENCIA E INVESTIGACIÓN
- Nº 24 LITERATURA ARGENTINA Y NACIONALISMO
- Nº 25 FRONTERA GANADERA Y GUERRA CON EL INDIO DURANTE EL SIGLO XVIII
- Nº 26 HISTORIADORES DEL SIGLO XIX Y LA HISTORIA DE AMÉRICA
- Nº 27 ESTUDIOS DE HISTORIA RURAL IV
- Nº 28 ESTRUCTURA DISCURSIVA DE LA ENTREVISTA RADIAL
- Nº 29 LA MÚSICA COMO REVELADORA DEL SENTIDO DEL ARTE EN MARCEL PROUST
- Nº 30 ROMANCES. POESÍA ORAL DE LA PROVINCIA DE BUENOS AIRES
- Nº 31 TEMAS DE HISTORIA ARGENTINA II. INSTITUTO DE HISTORIA ARGENTINA RICARDO LEVENE
- Nº 32 LITERATURA POLICIAL EN LA ARGENTINA. WALEIS, BORGES, SAER.
- Nº 33 CONSUMO ALIMENTICIO EN SECTORES POBRES URBANOS DEL GRAN LA PLATA
- Nº 34 LA BÚSQUEDA POR MATERIA Y LA DESCRIPCIÓN DE CONTENIDO EN EL CATÁLOGO EN LÍNEA

Para correspondencia y canje dirigirse a: Comité Editorial
Facultad de Humanidades y Ciencias de la Educación
Calle 48 y 6 - (1900) La Plata - Buenos Aires - Argentina

UNIVERSIDAD NACIONAL DE LA PLATA
FACULTAD DE HUMANIDADES Y CIENCIAS DE LA EDUCACIÓN

La búsqueda por materia y la descripción de contenido en el catálogo en línea

Ana M. Martínez y Mariana Pichinini

Serie: Estudios/Investigaciones

Año 1998

Cátedra de Clasificación I, Departamento de Bibliotecología,
Facultad de Humanidades y Ciencias de la Educación,
Universidad Nacional de La Plata

Presentación

Un catálogo en línea, que no pretenda ser una versión del antiguo catálogo en fichas sobre otro soporte, pone de manifiesto dificultades y falencias que en el pasado quedaban en penumbras: el problema de encontrar la información pertinente y las insuficiencias en el registro de los datos.

El catálogo tradicional servía, en tiempos de una producción editorial moderada, con el auxilio de las bibliografías, los servicios de resúmenes, las revistas de alertamiento, para dar cuenta aproximada del contenido del fondo documental. Los servicios de referencia hacían el resto. En la actualidad, con una producción editorial que crece en progresión geométrica, se imponen metodologías de trabajo más eficientes para satisfacer las demandas del usuario.

Por fortuna, se cuenta con una valiosa herramienta, la omnipresente computadora, que constituye una ayuda inapreciable para registrar y recuperar la información, para participar en sistemas de catálogos colectivos y para compartir el uso de los fondos bibliográficos.

Pero, al mismo tiempo, desnuda implacablemente los problemas que pasaban inadvertidos y se disculpaban en los catálogos manuales. La computadora exige normas de descripción precisas, porque exhibe las incoherencias en su aplicación, sobre todo cuando se participa en proyectos colectivos, pero el verdadero problema se presenta en el registro del contenido, con una serie de decisiones concomitantes: ¿sistemas precoordinados o poscoordinados o ambos? ¿Lenguajes documentales o lengua natural o ambas? ¿Inclusión de resumen en el registro?

Este trabajo, que reseña la experiencia acumulada en países con mayor tradición que el nuestro en la construcción y uso de catálogos en línea, es un aporte para prevenir dificultades en la etapa de generalización de catálogos en línea que se avecina en nuestro país.

Prof. Amelia Aguado
Directora
Departamento de Bibliotecología

Prefacio

En este trabajo se intenta presentar una revisión bibliográfica sobre la búsqueda por materia y la descripción de contenido en el catálogo en línea. Se ha querido fijar la atención en los catálogos de bibliotecas generales, es decir aquéllas que cuentan con una colección universal o enciclopédica y que ofrecen servicios a una comunidad heterogénea de usuarios. Como es sabido, estos catálogos incluyen principalmente monografías, con una cobertura temática extensa, pero de menor intensidad que las colecciones especializadas. Por lo general, la descripción del contenido se ha limitado a sintetizar el tema o la materia del documento en uno o dos epígrafes y/o un número clasificatorio seleccionado de un esquema de clasificación.

Al incorporarse la computadora a las actividades bibliotecarias, la descripción del contenido se llevó a cabo siguiendo estas características tradicionales. Pero tras varias décadas de experiencia, los cambios generados por la tecnología plantean nuevas concepciones del catálogo que hacen cuestionar las habituales prácticas de descripción; los usuarios presentan diferentes hábitos y preferencias de búsqueda que requieren sistemas más amigables y programas formales de instrucción, y al mismo tiempo es posible que toda la profesión bibliotecaria necesite redefinir sus actitudes frente al *nuevo catálogo*.

En el primer capítulo de esta revisión se presentan algunas tendencias del catálogo en línea; en el capítulo 2 se plantean los problemas que afectan al usuario durante la búsqueda por materia; en los capítulos 3 y 4 se abordan los aspectos relacionados con la descripción del contenido, incluyendo respectivamente la adición y la combinación de términos en cada registro documental; en el capítulo 5 se revisan algunas opciones para la exploración y selección de términos, mientras que el capítulo 6 está dedicado a los lenguajes documentales.

Como se puede constatar en la bibliografía, la mayoría de la información publicada sobre este tema no se ha producido en Argentina y, por lo tanto, esta revisión presenta una síntesis de las tendencias identificadas en países que cuentan con mayor experiencia en relación a los catálogos en línea. En el

séptimo y último capítulo se mencionan algunos aspectos referidos a la problemática local, con la esperanza de contribuir, en la medida de lo posible, a estimular una discusión más amplia y fructífera sobre este tema en nuestro medio.

Tal vez los estudiantes de Bibliotecología y los graduados no especialistas encuentren en este trabajo una revisión y puesta al día en español, mientras que aquéllos que deseen profundizar estos conocimientos vean facilitado el acceso a una porción importante de la información primaria mediante las referencias bibliográficas aquí citadas.

1. Los catálogos en línea y la búsqueda por materia

Los catálogos de biblioteca son instrumentos formales de recuperación de información que tradicionalmente han cumplido cuatro funciones: a) identificar un documento en particular, ya sea por su autor, su título o la materia de que trata, b) reunir los documentos sobre un mismo autor o materia, c) permitir la selección de un documento entre todos los que se han registrado y d) localizar el documento físicamente (Taylor, 1992). Estos catálogos se pueden presentar en forma impresa, en fichas, en microformas o en una computadora. En este último caso, cuando son accesibles desde una terminal e n modo interactivo, se denominan *catálogos en línea*.

Las nuevas tecnologías para el procesamiento de la información y las telecomunicaciones (Clayton, 1991) se han adaptado perfectamente a las necesidades de las bibliotecas, de manera que los catálogos en línea muestran en la actualidad ciertas tendencias que han sido resumidas por Carol Mandel, Directora de Procesos Técnicos de las Bibliotecas de la Columbia University de Estados Unidos, así como por otros autores (Bates, 1989; Buckland, 1992; Culkin, 1989; Gregor y Mandel, 1991; Mandel, 1991; Su, 1994; Tyckoson, 1991):

Procesos técnicos cooperativos: esta modalidad, también conocida como procesos técnicos por copia, se ha convertido en el punto central de las operaciones bibliotecarias eficientes, dado que la catalogación y la clasificación originales en cada biblioteca son extremadamente costosas; entre las empresas que brindan estos servicios pueden mencionarse las conocidas OCLC, RLIN y WLN de Estados Unidos, BLMCP, SWALCAP, LASER y SCOLCAP de Gran Bretaña, UTLAS en Canadá (Carpenter, 1988; Taylor, 1992; Tedd, 1987), así como las bibliotecas nacionales de diversos países (Holley, 1993; Jouguelet, 1989; Lopes, 1993).

Comportamiento de los usuarios durante la búsqueda en línea: la computadora ha permitido conocer mejor el comportamiento y las necesidades reales

del usuario, mediante los estudios de las operaciones de búsqueda. Los usuarios han demostrado que requieren considerable ayuda para recuperar la información del catálogo en línea, en especial para seleccionar sus términos y definir sus estrategias.

El catálogo universal: el catálogo en línea de una biblioteca individual puede considerarse como el subconjunto de un catálogo nacional virtual y cada vez más de un catálogo internacional o universal. En algunas bibliotecas es posible realizar una búsqueda multicatálogos, mediante interfaces que permiten a los usuarios moverse de un catálogo a otro fácilmente. En este sentido, también se ha llamado la atención sobre la necesidad de crear catálogos multilingües como ha ocurrido en Canadá (Rolland-Thomas y Mercure, 1989) o en Suiza (Satija, 1988).

El supercatálogo: existe una continua expansión de catálogos en línea que incluyen datos producidos externamente, ya sean resúmenes y/o referencias importados de diversas fuentes, de manera que los catálogos ya no consisten solamente en registros creados y controlados por la biblioteca. En este contexto, las diferencias entre los procesos técnicos de las bibliotecas y aquéllos de los servicios comerciales de indización, aunque parezcan insignificantes, pueden producir inconvenientes en la búsqueda. Para añadir complejidad, se comienzan a incluir documentos de texto completo y archivos no bibliográficos a los que se accede a través del catálogo.

Explosión de formas: aunque la impresión en papel sigue siendo la forma de mayor prevalencia para el conocimiento registrado, las nuevas formas electrónicas y magnéticas abundan y deben ser procesadas. Al mismo tiempo, las colecciones no bibliográficas se utilizan cada vez más, con el requisito concomitante de incluir estas colecciones en los catálogos de biblioteca.

Explosión de versiones: la introducción de nuevos medios, cambios en los patrones de publicación, reproducción y preservación, así como la preparación de documentos electrónicos, son algunos de los factores que conducen a una variedad cada vez más amplia de ediciones y versiones en las que puede aparecer una obra en particular.

Control de inventario: en muchas bibliotecas el catálogo en línea es el módulo de un sistema integrado que soporta una variedad de procesos y

funciones asociados al catálogo. Los registros de este catálogo están vinculados con información detallada de cada documento individual, identificado por un código de barras. De manera que mediante el acceso a un solo sistema es posible consultar varias fuentes y obtener distintas referencias de documentos que luego son localizados físicamente a través del catálogo.

Culkin (1989) ilustra en parte estas tendencias con el siguiente ejemplo: un usuario se conecta al catálogo de una biblioteca y solicita una búsqueda por materia, recuperando un cierto número de registros. El usuario se da cuenta de que varios de estos registros contienen un término cuyo significado desconoce; se mueve entonces, dentro del mismo sistema y comunicación, a un diccionario en línea que le brinda la definición correspondiente; luego a una enciclopedia que le amplía la información con un artículo de síntesis; a continuación consulta un diccionario ideológico, también en línea, para obtener los sinónimos y por último una bibliografía que le muestra las referencias de documentos que han sido publicados. Retoma entonces al catálogo y ahora sí selecciona los documentos que le interesan. Si este usuario hipotético descrito por Culkin tuviera además la posibilidad de moverse hacia los catálogos en línea de otras bibliotecas -en la misma comunicación-, sus posibilidades se multiplicarían infinitamente.

Como puede observarse, una de las características principales del catálogo en línea es el *acceso público*, es decir la posibilidad de que los usuarios realicen sus propias búsquedas sin intermediación del bibliotecario. Para que este acceso público sea eficiente, se necesita un acceso remoto al catálogo, que puede ser de dos tipos (Clayton, 1991; Duval y Main, 1992; Sloan, 1991):

Acceso remoto local: requiere una red de área local (*local area network* o LAN), que puede definirse como un conjunto de computadoras y periféricos conectados, dentro de un área geográfica limitada quizás a unos pocos kilómetros (por ejemplo varios edificios). Las LAN pueden transmitir datos, imágenes y voz. Este tipo de red permite, entre otras cosas, que un usuario acuda a la biblioteca, utilice una terminal de la sala de lectura o de catálogo y realice su consulta simultáneamente con otros usuarios, mientras que el personal de la biblioteca almacena información en la oficina de procesos técnicos, o tramita un préstamo en el mostrador de circulación.

Acceso remoto telemático: requiere una comunicación telemática, para lo cual las computadoras se conectan a la línea telefónica utilizando un módem que traduce las señales digitales de la computadora en señales analógicas para el teléfono y vice versa. La cobertura geográfica es amplísima, dado que la comunicación puede realizarse inclusive entre computadoras de distintos países. Este tipo de comunicación permite que un usuario consulte el catálogo de una biblioteca desde su oficina, escuela u hogar, en la misma ciudad o en otra. Una comunicación telemática puede efectuarse dentro de una red, por ejemplo Internet, o por teledisco directo.

Se puede afirmar entonces que el futuro plantea una biblioteca cada vez más virtual: sin papel porque los documentos tienden a ser electrónicos o magnéticos, sin paredes porque los usuarios podrán consultarla desde lugares remotos, sin horarios porque las computadoras nunca duermen. Pero hay un aspecto en el que aún no se ha podido dar una respuesta satisfactoria a los usuarios: la búsqueda por materia.

En un interesante artículo, Birger Hjørland (1992) reflexiona sobre las distintas concepciones epistemológicas del concepto *materia* en Bibliotecología; aunque diferentes entre sí, cada una de estas concepciones evidencia el mismo dilema: hacer coincidir el contenido del documento con la descripción hecha por el bibliotecario y la consulta del usuario.

Ya en 1954 Oliver L. Lilley solicitó a 340 estudiantes que eligieran distintos encabezamientos de materia para localizar seis libros en el catálogo; el promedio de sugerencias fue de 62 encabezamientos diferentes para cada libro. En una experiencia similar realizada por Marcia Bates en 1977, 71 estudiantes sugirieron 46 encabezamientos diferentes para un solo libro, con un máximo de seis personas recomendando un mismo epígrafe. Al resumir estos dos estudios con los de otros autores, Bates ha calculado que el promedio de posibilidades de que dos personas (o un usuario y un catálogo) utilicen el mismo término para el mismo concepto es de 10-20% (Bates, 1989).

Según McCarthy (1989), el bibliotecario tiene un documento en la mano que intenta describir asignándole una etiqueta o término de indización; lo que busca es precisión. El usuario en cambio tiene una idea en mente y espera encontrar reunidos todos y cada uno de los documentos que respondan a esa

idea; lo que busca es precisión más *recall* (en este trabajo se utiliza el término *recall*, por no conocerse una traducción satisfactoria en español).

La búsqueda por materia es por naturaleza inexacta y exploratoria y requiere múltiples rutas de acceso (Mandel, 1991). Si bien los catálogos en línea han puesto en evidencia con mayor intensidad la falta de coincidencia entre el vocabulario del usuario y el del sistema, es justamente el enorme potencial de la computadora para procesar y recuperar información el recurso que puede ofrecer las soluciones que parecían imposibles en el catálogo de fichas, pero es necesario ante todo conocer las características de los usuarios y su desempeño en este tipo de catálogos.

2. Los usuarios y la búsqueda por materia

Idealmente, se ha postulado que los usuarios deberían tener una participación protagónica en los sistemas de información, tanto en las etapas iniciales de diseño e implementación (consultas, estudios de opinión, determinación de necesidades), como en aquéllas intermedias y finales de evaluación y corrección (encuestas sobre servicios, estudios de usuarios, etc.). Sin embargo, no todos los sistemas toman en cuenta al usuario de la misma manera; Cochrane (1984) distingue tres enfoques distintos:

Enfoque orientado al sistema: el acceso por materia es considerado en términos de producto final (catálogos y bibliografías) y se pone énfasis en las que se consideran operaciones primarias del sistema: la descripción bibliográfica y la descripción del contenido. La búsqueda es un servicio que se brinda a los usuarios, pero éstos quedan fuera del sistema. Este enfoque toma como factores importantes la indización profunda, el tamaño de la colección, el vocabulario del sistema y las reglas y procedimientos requeridos para registrar y recuperar información de una base de datos particular. Son los catalogadores, indizadores y analistas de sistemas quienes determinan qué áreas del sistema bajo control se pueden mejorar. La producción de la base de datos es el foco central de esta visión.

Enfoque orientado al usuario: en contraposición, este enfoque pone énfasis en los procesos humanos de la información, la respuesta del usuario, los resultados de la búsqueda, la interfaz usuario-máquina y el rol del bibliotecario como intermediario. Se ignora la influencia del diseño del catálogo sobre la búsqueda por materia, incluyendo la complejidad del sistema, el formato de la base de datos y la naturaleza de los documentos. El usuario debe tomar decisiones entre la solicitud y la respuesta final, como explorar, refinar, limitar, solicitar ayuda y elegir puntos de acceso adicionales.

Enfoque usuario-sistema: Esta tendencia involucra al usuario (sus habilida-

des, conocimiento, experiencia, entrenamiento, actitudes, aspiraciones, necesidades de información y tareas a desarrollar) y su comunicación con el sistema (bases de datos propias y externas, telecomunicaciones, redes, etc.); el foco de atención es la interfaz que debe ser dinámica y explícita, incluyendo funciones, operaciones, comandos, menús, mensajes de ayuda y error, exhibición de términos, formatos, etc.

Como señalan algunos autores, las operaciones del sistema (descripción del contenido) y los requerimientos del usuario (búsqueda por materia) son dos caras de una misma moneda, es decir se debe describir para facilitar el acceso (Bishop, 1983; Gregor y Mandel 1991; Mandel, 1991). Sin embargo, con frecuencia el conocimiento o la información no se encuentran a disposición de los usuarios porque los sistemas que deben estar a su servicio no logran hacerse compatibles con ellos ni con sus necesidades.

Uno de los aportes importantes que ha hecho la informática a las bibliotecas ha sido la posibilidad de conocer mejor el desempeño de los usuarios, mediante el estudio de los registros de las operaciones de búsqueda (*transaction logs*) (Gregor y Mandel, 1991; Mandel, 1991), que ha permitido complementar los datos surgidos de las tradicionales encuestas. Los resultados de ambos tipos de estudio demuestran que los usuarios enfrentan más problemas y requieren mayor ayuda en la búsqueda por materia en los catálogos en línea que en aquéllas por referencias conocidas (Klugman, 1989; Ríos García, 1991).

Es obvio que para evitar estos problemas y mejorar los sistemas, en primer lugar se deben tener en cuenta las características de los usuarios, ¿pero cuáles son éstas?

Expectativas del usuario

Ante el catálogo en línea, las expectativas de los usuarios, en especial sobre el acceso por materia, son muchas. Las estadísticas indican que, en promedio, 59% de las búsquedas en línea son temáticas (Larson, 1991b; Su, 1994).

Los usuarios, que durante largo tiempo han enfrentado las limitaciones de los catálogos en fichas, desean que la magia de la tecnología haga la diferencia; esperan sentarse ante el teclado, pedir los temas específicos y que el mago

invisible en la caja negra frente a ellos les dé las respuestas que nunca recibieron (Dykstra, 1989; Klugman, 1989). Es evidente que estas expectativas se ven alimentadas por experiencias con sistemas computarizados en otros contextos, como los bancos y los comercios, y les es difícil comprender que los catálogos en línea de las bibliotecas sean diferentes.

En general, les agrada el catálogo en línea y lo prefieren al de fichas porque les brinda nuevos servicios. Además, desean sentir que controlan lo que ocurre y que pueden aprender y recordar el uso del sistema fácilmente, así como obtener éxito en sus búsquedas. Aunque parecen amar todo lo que es en línea, el entusiasmo inicial dura un tiempo demasiado corto. El catálogo puede ser poco comunicativo y poco útil y hay que saber cómo descifrar y usar el sistema en forma correcta para lograr cierto éxito (Klugman, 1989).

Comportamiento del usuario

Cuando una persona se comunica con el catálogo en línea, la interfaz es vista como otro individuo con quien conversar. El usuario humaniza el sistema y demuestra reacciones emocionales ante sus respuestas. Como señala Ríos García (1991), la expresión *amigable* refleja el aspecto social y antropomórfico de esta interacción. Se reacciona mal a un tiempo de respuesta demasiado largo, porque distrae y hace perder la concentración en lo que se está haciendo; también a un tiempo demasiado corto, porque impone un ritmo acelerado que genera tensión y obliga a trabajar bajo presión. La falta de mensajes oportunos y precisos o el exceso de instrucciones pueden perturbar la sensibilidad del usuario y hacerlo sentir molesto e incluso agredido.

En algunos casos, la satisfacción con una respuesta inicial considerada positiva por el usuario no implica éxito real en la búsqueda. Algunos estudios muestran una alta satisfacción de los usuarios; por ejemplo Cherry (1992) señala una proporción de 80-85%, Olsen et al. (1989) de 67%, Ensor (1991) de 93%. En todos estos casos los datos se obtuvieron de encuestas, que incluían preguntas con respuestas SI/NO. Ankeny (1991) de la Ohio State University, realizó dos estudios de satisfacción: en el primero la pregunta *¿Obtuvo la información deseada?* recibió una respuesta afirmativa del 78%; en el segundo estableció tres criterios de satisfacción para los cuales el usuario encuestado debía emitir su

propio juicio en una respuesta subjetiva. Solo 39% de las respuestas reunieron los tres criterios, lo que coincide -según el mismo Ankeny- con los trabajos de otros autores que postulan que el promedio de satisfacción está en el 50% o menos.

Los fracasos reiterados sin duda producen estrés y frustración, provocando un rechazo al catálogo en línea y en general a la biblioteca (Borgman, 1996; Klugman, 1989; Ríos García, 1991). Se ha detectado incluso una cierta declinación de la búsqueda por materia a lo largo del tiempo, en favor del acceso por autor y por título (Larson, 1991b).

Problemas que enfrenta el usuario

Ríos García (1991) reconoce dos tipos de problemas: los mecánicos y los conceptuales.

Entre los problemas mecánicos, es decir aquéllos que se originan en las dificultades para comprender e interactuar con el sistema, se pueden mencionar:

Transición y variedad: la transición del catálogo tradicional al catálogo en línea, que a veces se extiende demasiado con muchas fases intermedias, así como la coexistencia de diferentes sistemas en la misma biblioteca, con frecuencia desorientan al usuario debido a lo que se ha denominado un *estado de continua mutabilidad* (Klugman, 1989).

Interfaz usuario-sistema: los usuarios tienden a usar solo las técnicas básicas de selección y combinación de términos. Rara vez utilizan más de un limitado conjunto de las características del sistema. La consulta en línea les obliga a tener un conocimiento mínimo sobre la computadora. Por simple que sea un menú, deben leer y seguir instrucciones a cada momento, así como analizar sus pensamientos y traducirlos a una serie de órdenes y comandos que no siempre resultan fáciles de construir (Borgman, 1996; Klugman, 1989; Millsap y Ferl, 1993; Ríos García, 1991).

Normalización: la variedad de órdenes, comandos, menús, así como de interfaces y formatos de salida entre distintos catálogos hace más difícil aún la tarea de los usuarios. Aquéllos con experiencia en catálogos distintos suelen asumir analogías que no existen realmente -aplicando procedimien-

tos de un sistema a otro- y que solo aumentan la confusión. Con el fin de normalizar el proceso de búsqueda, en 1988 el American National Standard Institute desarrolló la norma Z39.50 para la recuperación de información, que ha sido revisada en 1992 y 1994. La Z39.50 le permite al usuario utilizar los mismos comandos de búsqueda en diferentes bases de datos, aún cuando éstas tengan distintas estructuras. El usuario crea la estrategia de búsqueda en un protocolo Z39.50; el server de la base de datos recibe este protocolo, lo traduce a sus propios comandos, recupera la información, retraduce al protocolo Z39.50 y lo envía al usuario, que recibe así un formato normalizado. Esta norma Z39.50 ha permitido diseñar nuevas interfaces que reducen los requisitos de conocimiento sintáctico y semántico para realizar búsquedas en línea (Beheshti, 1997), aunque el costo de su implementación aún no ha permitido su uso masivo en bibliotecas pequeñas o medianas (Mulder, 1995).

Por otro lado, los problemas conceptuales, aquéllos que se originan en una falta de conocimiento del proceso de búsqueda, incluyen principalmente los siguientes:

1. Dificultad en hacer coincidir los términos del usuario con el vocabulario del sistema

La búsqueda por materia es la que mayor índice de fracasos presenta (40% de respuestas con cero registros recuperados) (Hancock, 1987 y 1990; Kern-Simirenko, 1983; Larson, 1991a y 1991b; Su, 1994). Este alto índice de fracasos ha sido relacionado con la falta de coincidencia entre el vocabulario del usuario y el vocabulario del sistema.

En un trabajo realizado por Alyson Carlyle (1989) de la University of California at Los Angeles, los términos utilizados por los usuarios coincidieron total o parcialmente con el 74% de los encabezamientos simples de LCSH y el 21% de los encabezamientos compuestos; al comparar sus resultados con los datos de otros autores, Carlyle pudo establecer que el promedio de coincidencia es de aproximadamente un 50%.

Por su parte, Cousins (1992) en Gran Bretaña obtuvo 62% para LCSH y

30% para CDD en una biblioteca politécnica y 73% y 36% respectivamente en una biblioteca pública.

Esta falta de coincidencia ha sido atribuida a dos causas principales: a) pocos términos descriptivos de materia en cada registro y b) limitaciones de los lenguajes documentales.

2. Dificultad en construir búsquedas booleanas

A los usuarios se les hace muy difícil entender cómo operan sobre el sistema los operadores booleanos. Thomas Peters (1989), de la University of Missouri, analizó más de trece mil búsquedas realizadas en el catálogo Luis de esa universidad entre 1987 y 1988 y encontró que sólo el 1% de los usuarios había utilizado los operadores booleanos.

Millsap y Ferl (1993), de la University of California at Santa Cruz, sobre los registros de 1045 búsquedas realizadas en el sistema Melvyl, detectaron que solo un 10% incluía estrategias booleanas.

En una encuesta realizada por Olsen et al. (1989), de la Brigham Young University, el 31% de los usuarios manifestó haber comprendido el uso de estos operadores, pero solo el 24% los había utilizado.

Pat Ensor (1991), de la Indiana State University, llevó a cabo otra encuesta para medir el grado de satisfacción de los usuarios en la búsqueda por materia; respecto a la dificultad para aprender el uso de los operadores booleanos, el 40% de los encuestados ni siquiera contestó la pregunta y muchos de los que contestaron que *era fácil*, en la práctica demostraron no conocerlos.

Investigaciones realizadas en el campo de la psicología cognitiva demuestran que las personas no utilizan la lógica en su vida cotidiana, de modo que tienden a confundir el Y lógico (exclusivo) con el Y lingüístico (inclusivo), así como el O lógico (inclusivo) con el O lingüístico (exclusivo) (Borgman, 1996).

3. Dificultad para truncar y limitar

A la mayoría de los usuarios tampoco les es fácil truncar efectivamente para aumentar la *recall*, o deshacer un truncamiento implícito o utilizar límites para reducir el resultado de una búsqueda (Ríos García, 1991; Kluqman, 1989).

Los trabajos de Kern-Simirenko (1983) y Peters (1989) coinciden en señalar el poco uso del truncamiento y el hecho de que rara vez se inicie una búsqueda por esta técnica. En la encuesta de Olsen et al. (1989), el 58% de los usuarios manifestaba comprender la técnica del truncamiento, pero solo el 52% la había utilizado realmente.

El uso incorrecto del truncamiento y de los límites produce sobrerecuperación o sobrecarga de información, es decir recuperaciones con demasiados registros. Diversos estudios demuestran que un 22-53% de las búsquedas por materia obtienen más de 50 registros y que alrededor del 40% de ellas son revisadas por los usuarios sin haber sido reducidas (Millsap y Ferl, 1993; Kern-Simirenko, 1983).

Por su parte Larson (1991b), de la University of California at Berkeley, ha calculado que las búsquedas por materia en el sistema Melvyl tienen un promedio de 200 registros recuperados, aunque los usuarios solo revisan un promedio de 9 registros por búsqueda. El 37% de las búsquedas con más de 100 registros recuperados y el 52% de aquéllas con más de 1000 registros recuperados son búsquedas por materia.

Micco (1991) ha informado que las búsquedas por materia en la Indiana University of Pennsylvania tienen un promedio de 170 registros recuperados.

4. Dificultad para modificar el resultado de una búsqueda

Raya Fidel (1991) de la University of Washington at Seattle, analizó 1246 movimientos de conversión de 281 búsquedas realizadas por 47 usuarios; sus resultados indican que en la mayoría de los casos se convirtió para ampliar la búsqueda, es decir para mejorar la *recall* (62%), no para limitarla o mejorar la precisión (32%); el 6% restante intentó mejorar tanto la *recall* como la precisión.

Hancock (1990), de la City University de Londres, analizó 204 búsquedas en línea, de las cuales 60% fueron por materia. En un primer movimiento, el 58% de los usuarios inició su búsqueda en la computadora utilizando un término exacto respecto del tema expresado en la solicitud; 27% utilizó un término más específico (mayor precisión) y 15% un tema más amplio (mayor *recall*). Después de varios movimientos, el resultado final demostró que la preferencia por el término exacto se redujo al 21%, por el tema más específico aumentó ligeramen-

te al 35% y por el término más amplio a un 44%.

En este último estudio, un tercio de las búsquedas que se iniciaron por autor o por título se reconvirtieron a materia, indicando que cuando los usuarios buscan información sobre una persona como materia, lo intentan en el índice de autores, según afirma también Peters (1989) y que la búsqueda por palabras claves de título tiene una doble utilidad: localizar un título conocido o bien realizar una búsqueda por materia mediante las palabras claves en lengua natural.

Con respecto al uso del título como materia, los usuarios intentaron reconversiones similares al caso anterior: la búsqueda por el término exacto disminuyó de 52% a 14%, mientras que la búsqueda por un término más específico aumentó de 30% a 42% y por un término más amplio de 18% a 44%.

Cherry (1992), de la University of Toronto, encontró que el 50% de las búsquedas por materia debieron reintentarse; en aquéllas que utilizaron los encabezamientos de LCSH y que no obtuvieron ningún registro recuperado, la conversión más útil (71% de éxito) resultó ser la utilización de palabras claves de título; sin embargo, solo el 25% de los usuarios intentó esta conversión.

Usuario local, usuario remoto

Por otro lado, el acceso remoto ha producido dos tipos de usuarios: el usuario local (en la biblioteca) y el usuario remoto o *invisible* (lejos de la biblioteca) (Sloan, 1991).

Para consultar el catálogo en línea desde la terminal de una biblioteca, el usuario requiere solo un entrenamiento sencillo, se le pueden proporcionar guías o instructivos impresos y se comunica con el bibliotecario en forma personal, ya sea para recibir ayuda o sugerir mejoras en el servicio; una vez que ha seleccionado un documento en el catálogo, tiene la posibilidad de revisarlo físicamente y decidir si le sirve o no para su propósito; si la biblioteca no ha completado la reconversión de su catálogo en fichas, puede consultar este último en la biblioteca.

En cambio, cuando se accede al catálogo desde fuera de la biblioteca, se requiere conocer el hardware y el software de comunicación y los aspectos operativos de la red, antes de poder consultar el catálogo; no puede comunicarse con el bibliotecario y éste nunca está seguro de si el usuario ha recibido un buen

servicio, qué documentación posee sobre el sistema o qué problemas ha tenido durante la búsqueda; el usuario está imposibilitado de consultar el documento físicamente y en consecuencia su solicitud puede estar equivocada; si la biblioteca no ha completado la reconversión de su catálogo en fichas, se pierde la información que éste registra.

Algunas recomendaciones

Para solucionar los problemas relacionados con la búsqueda por materia se han señalado tres caminos: mejorar la interfaz usuario-sistema, mejorar los programas de instrucción del usuario y mejorar las bases de datos, particularmente la descripción de contenido (Duval y Main, 1992; Larson, 1991b; Su, 1994), ya sea añadiendo más puntos de acceso por materia, combinando los distintos tipos de puntos de acceso por materia y ofreciéndole al usuario la posibilidad de explorar y seleccionar sus términos de una lista.

3. Descripción de contenido: adición de términos

Charles Hildreth ha clasificado los catálogos en línea en tres generaciones (Su, 1994). Aquéllos de primera generación no son más que catálogos en fichas reciclados en un ambiente electrónico (Bates, 1989; Culkin, 1989; Lancaster, 1995; Su, 1994) o como los denomina Mary Dykstra (1989) *catálogos en fichas sobre ruedas*.

Cuando los bibliotecarios tuvieron a su disposición la tecnología para diseñar catálogos en línea, lo que sabían hacer -y lo habían hecho durante más de un siglo- era fichas; por lo tanto, la decisión fue muy sencilla: pasar la ficha de 7,5 x 12,5 cm a la computadora y probar qué pasaba. Esta decisión también se vio influida por la prioridad de las grandes bibliotecas, en especial las universitarias, de informatizar el servicio de circulación, para lo cual no se necesitaba más información que la contenida en la ficha tradicional.

Sin embargo, cuando estos catálogos se pusieron a disposición del usuario para la búsqueda por materia, convirtiéndose así en verdaderos instrumentos de recuperación de información y no solo de datos (Larson, 1991b), el problema se hizo evidente sobre todo en las bibliotecas generales: la cantidad de puntos de acceso por materia que se ofrece al usuario durante la búsqueda es insuficiente para lograr la coincidencia entre el vocabulario del usuario y el del sistema. En los Estados Unidos se ha podido determinar que el promedio de epígrafes por registro en los catálogos en línea se limita a 1,8, con lo cual no es posible siquiera una combinación booleana; además carecen de números clasificatorios y de recuperación por lengua natural, a excepción de algunos pocos sistemas que permiten acceder por palabras claves de título (Su, 1994).

Como solución a este problema, Pauline Cochrane, de la Syracuse University de Estados Unidos, propuso en 1977 una metodología consistente en aumentar los puntos de acceso por materia en los registros de monografías, incluyendo los términos de sumarios e índices analíticos en el campo de resumen, para ser recuperados por lengua natural (Seille y Cochrane, 1982). Aunque el

resumen tenía ya una larga historia en la elaboración de bibliografías, su inclusión en un catálogo de biblioteca resultó novedosa. Con el fin de evaluar esta propuesta, se comparó un catálogo de 1979 libros en humanidades y ciencias sociales cuya descripción se hizo con el formato MARC por un lado y con el *formato aumentado* de Settle y Cochrane por el otro; esta última base de datos se denominó BOOKS. En 90 búsquedas de prueba la recuperación con MARC fue de 56 registros en 8 minutos, mientras que con BOOKS fue de 130 registros en 4 minutos (233% de los registros MARC en 50% del tiempo).

Este formato aumentado constituye la base de los llamados catálogos de segunda generación y ha sido utilizado y recomendado ampliamente por diversos autores (Byrne y Micco, 1988; Drabenstott et al., 1990; Dwyer, 1991; Micco, 1991; Stephens, 1991; Studwell, 1991; Syracuse y Poyer, 1991).

Algunos consideran que el formato aumentado, si bien tiene méritos teóricos, puede producir problemas prácticos en cuanto al costo y tiempo de la indización, así como el manejo de grandes cantidades de palabras claves por los índices invertidos (Holley y Killheffer, 1982; Larson, 1991b).

Según Mandel (1985) la indización con formato MARC tenía un costo para Library of Congress de U\$S15,00 por título, cifra que aumentaría notablemente al incluirse las palabras claves de sumarios e índices; sin embargo, Settlet y Cochrane (1982) informaron un costo de solo U\$S5,00 por título.

Byrne y Micco (1988) obtuvieron un tiempo de indización de 15 minutos, para un promedio de 53,6 términos por registro, mientras que en la experiencia de Settlet y Cochrane (1982) el tiempo fue de 10 minutos por título con un promedio de 32,4 términos por registro. De manera que se hace necesario evaluar estas posibilidades, antes de su implementación.

No obstante, Shirley A. Cousins (1992), del University College of Wales de Gran Bretaña, propuso traducir los términos de sumarios e índices analíticos a un lenguaje documental, en este caso PRECIS. Realizó tres pruebas con distintos resultados: formato MARC sin sumarios ni índices 13% de *recall*; formato aumentado recuperado por lengua natural 76% de *recall*; formato aumentado recuperado por PRECIS 81% de *recall*. La diferencia de *recall* entre el formato MARC y el formato aumentado es concluyente; sin embargo, aunque Cousins opina lo contrario, la diferencia entre la lengua natural y PRECIS (ambas con formato aumentado) no parece ser significativa y quizás no justifique el costo de

la traducción de los términos de sumarios e índices a un lenguaje documental.

En forma similar, Gunnar Knutson (1991) llevó a cabo una investigación en la University of Illinois at Chicago con el propósito de establecer la relación entre el formato aumentado y el uso de la colección, evaluado a través de la circulación. Dividió su colección experimental de 291 documentos en tres grupos: a) formato MARC aumentado con más encabezamientos de materia (promedio de 7,5 encabezamientos por registro) y una nota de contenido según AACR2, b) formato MARC común y c) formato MARC aumentado sólo con la nota de contenido. La circulación para el primer grupo fue de 46%, mientras que para los otros dos fue de 26% y 28% respectivamente. Estos resultados no son concluyentes debido al pequeño tamaño de la colección experimental, por lo que Knutson recomienda continuar con este tipo de investigaciones para determinar si el lenguaje documental es tan decisivo en el uso de la colección como lo indican las cifras mencionadas.

El formato aumentado también parece satisfacer mejor a los usuarios, según lo han revelado distintas encuestas. En un trabajo realizado para OCLC, Kaske y Sanders (1980) entrevistaron a 200 usuarios pidiendo sugerencias con el fin de mejorar la búsqueda por materia; la principal respuesta que obtuvieron fue la recuperación por lengua natural de las palabras claves del título, del sumario y de los índices analíticos. La misma respuesta obtuvo Bessant (1982) en una encuesta realizada para el Council of Library Resources y Frost (1987) entre estudiantes avanzados de la University of Huston. Por su parte, Piascik (1993) de la Kent State University realizó una encuesta entre indizadores de las bibliotecas públicas del estado de Ohio sobre la conveniencia del formato aumentado; la respuesta general de estos bibliotecarios fue que era de gran utilidad para los usuarios.

En síntesis, los catálogos de segunda generación se caracterizan por ofrecer al usuario una mayor cantidad de puntos de acceso por materia en cada registro, incluyendo los siguientes (Duval y Main, 1992; Larson, 1991b; Su, 1994):

Más términos controlados verbales (epígrafes o descriptores).

Un número clasificador para recuperar por materia, con el mismo sentido que el catálogo sistemático.

Los sinónimos verbales de los números clasificatorios, incluyendo la

denominación de la clase, las entradas en el índice alfabético de la clasificación y la vinculación de la notación con epígrafes o descriptores.
La recuperación por lengua natural de las palabras claves de título.
La recuperación por lengua natural de las palabras claves de resumen.

En particular el resumen, que generalmente tiene una extensión de 200-250 palabras y que ofrece la posibilidad de complementar el lenguaje documental con sinónimos y las propias palabras del autor, marca la gran diferencia entre los catálogos de primera y segunda generación. Sin embargo, su implementación tiene un costo adicional que muchas bibliotecas no están en condiciones de afrontar. No obstante, se ha demostrado que las restantes características del formato aumentado mejoran la recuperación.

Los catálogos de tercera generación también se basan en el formato aumentado, pero cuentan con una interfaz inteligente. Como afirma Hildreth, las tres generaciones de catálogos en línea coexisten en este momento, pero mientras que los de primera generación son la gran mayoría, los de tercera generación son aún excepcionales, incluso en los países altamente desarrollados, debido al costo de la interfaz inteligente. Mientras tanto, para aquellas bibliotecas que cuentan con recursos informáticos convencionales la solución parece ser la reconversión de los catálogos en línea de primera generación a catálogos con formato aumentado (Su, 1994).

Sin embargo, como aseguran Chan (1990) y Larson (1991b), no basta con incrementar el número de términos por registro, sino que además deben combinarse distintos tipos de términos, que ofrezcan diferentes opciones para el usuario.

4. Descripción de contenido: combinación de términos

Tradicionalmente, cuando se organizaba un catálogo, el bibliotecario se enfrentaba a una serie de alternativas, debiendo elegir solo una de entre dos opciones posibles:

Lengua natural o lenguaje controlado.

Acceso alfabético o acceso sistemático.

Precoordinación o poscoordinación.

Por otra parte, estaban los catálogos públicos y los internos. Estos últimos incluían el catálogo topográfico y el de autoritativos; los bibliotecarios hacían uso de ellos como apoyo a sus tareas de procesos técnicos, pero los usuarios no tenían acceso a estos catálogos internos.

Uno de los grandes aportes de la computadora ha sido precisamente la eliminación de estas encrucijadas, permitiendo que en un mismo catálogo se puedan ofrecer todas las opciones en forma simultánea, compensando las desventajas de unas con las ventajas de otras, tanto para los usuarios como para los bibliotecarios.

Lengua natural y lenguajes documentales

Como es sabido, el uso de la lengua natural en la indización tiene importantes ventajas (Tabla 1), pero presenta serios inconvenientes motivados principalmente por la falta de control del vocabulario (semántica, morfología, etc.) y de relaciones explícitas entre los términos de ese vocabulario. Para compensar en parte estas desventajas, se recomiendan las técnicas del truncamiento y la adyacencia, o los tesauros de búsqueda (Aitchison et al., 1997; Bates, 1988; Harris, 1986; Lancaster, 1986 y 1995; Strong y Drott, 1986).

Otra opción es traducir los términos o palabras claves de la lengua natural a un lenguaje documental (lista de encabezamientos, tesoro o clasificación).

Tabla 1. Lengua natural vs. lenguajes documentales
(Aitchison et al., 1997; Dubois, 1986; Lancaster, 1995)

LENGUA NATURAL	LENGUAJE DOCUMENTAL
<i>Ventajas</i>	<i>Desventajas</i>
- Alta especificidad y precisión. Excelente para recuperar términos individuales, nombres de personas, organizaciones, etc.	- Menor especificidad, aún en sistemas muy detallados.
- Mayor exhaustividad.	- Menor exhaustividad.
- Actualización inmediata.	- La actualización no es inmediata.
- Utiliza las mismas palabras del autor, no existen malas interpretaciones del indizador.	- Los errores del indizador pueden causar pérdidas de información
- El usuario utiliza las palabras en lengua natural.	- El usuario tiene que aprender el lenguaje documental.
- Bajos costos en el almacenamiento.	- Altos costos en el almacenamiento.
- Compatibilidad entre distintos catálogos.	- Incompatibilidad entre distintos catálogos.
<i>Desventajas</i>	<i>Ventajas</i>
- Presenta problemas por la falta de control de vocabulario: sinonimia, homonimia, polisemia y antonimia, singular y plural, etc.	- Facilita la búsqueda mediante el control del vocabulario, así como por las relaciones paradigmáticas y sintagmáticas
- Puede perderse información implícita, pero no explícita en el texto.	- Expresa conceptos que están implícitos, pero no explícitos.
- Produce problemas de sintaxis. Existe la posibilidad de recuperaciones falsas por una asociación incorrecta de términos.	- Evita los problemas de sintaxis mediante los términos compuestos y otros elementos como vínculos y roles.
	- A niveles normales de indización, evita la sobreexhaustividad y puede facilitar la tarea de catálogos multilingües.

Los lenguajes documentales, también llamados de indización o controlados, son lenguas artificiales consistentes en un conjunto de términos (epígrafes, descriptores, notaciones) que permiten describir en forma resumida el contenido de los documentos y que se utilizan con el fin de indizar, almacenar y recuperar la información (ISO 2788, 1986).

Poseen un vocabulario unívoco y un sistema explícito de relaciones paradigmáticas y sintagmáticas. Sus ventajas y desventajas se presentan también en la Tabla 1.

La controversia entre el uso de la lengua natural y los lenguajes controlados tiene una larga historia. Elaine Svenonius (1986) la ha dividido en tres eras. La primera tuvo lugar desde los tiempos en que Samson Low elaboró en 1854 el *British catalogue of books* sobre la base de las palabras claves del título y que su asistente Andrea Crestadoro hacía lo propio con el *Catalogue of books of the Manchester Free Library*, mientras que por el contrario, Charles A. Cutter, al argumentar en favor de sus reglas para encajeamientos de materia en 1876 aseguraba tratar con la muerte de la indización por palabras claves, vaticinio que casi se convirtió en realidad hasta mediados de este siglo.

En esta era aparecieron LCSH, LCC, CDD, CDU, BBC, CC, por lo cual se puede considerar como la época de oro de los lenguajes documentales universales. Sin embargo, en la década de los cincuenta el sistema Uniterm de Mortimer Taube y los índices KWIC de Hans P. Luhn dieron lugar a que se revaloraran las posibilidades de la lengua natural, reabriendo la polémica con los lenguajes documentales y estimulando, paradójicamente, el afianzamiento del recién aparecido tesoro en el campo de la información (Lancaster, 1995; Roberts, 1984).

En la segunda era tuvieron gran influencia los dos estudios de Cranfield, Gran Bretaña, y otras investigaciones similares. En particular los resultados de Cranfield 2 señalaban que la lengua natural, sin control o con un mínimo de control del vocabulario, permitía una recuperación tan eficiente como un vocabulario totalmente controlado. Aunque la metodología de estas investigaciones fue criticada, los defensores de la lengua natural se apoyaron en estos informes, teniendo en cuenta sobre todo los considerables costos de construir un lenguaje documental (Dubois, 1987; Lancaster, 1995; Svenonius, 1986).

La tercera y última era puede resumirse en las palabras de Russell J. Rowlett

Jr., ex director de *Chemical Abstracts*: «las búsquedas con lenguaje controlado o con lengua natural [...] ofrecen resultados *complementarios*, no idénticos» (Svenonius, 1986).

Los experimentos de Cranfield introdujeron dos variables operativas para medir el desempeño de los sistemas de recuperación (Aitchison et al., 1997; Dubois, 1987; Foskett, 1996; Lancaster, 1986 y 1995; Svenonius, 1986):

Recall: es la proporción de documentos relevantes recuperados, es decir el porcentaje de documentos relevantes que se han recuperado, en relación con el total de documentos relevantes disponibles en la base de datos. Dicho de otra forma, la capacidad de ampliar una búsqueda o de reunir todos los documentos que tratan sobre un mismo tema.

Precisión: es la proporción de documentos recuperados relevantes, es decir del total de documentos que se recuperaron, aquéllos que son relevantes. Dicho de otra forma, la capacidad de reducir o limitar una búsqueda.

Entre *recall* y precisión existe una relación inversamente proporcional: a mayor *recall* menor precisión y vice versa. Una búsqueda muy amplia tiende a recuperar mayor cantidad de documentos (alta *recall*, baja precisión), mientras que una búsqueda más restringida recupera pocos documentos (alta precisión, baja *recall*) (Foskett, 1996; Lancaster, 199).

Algunos estudios posteriores a Cranfield 2, realizados sobre bibliografías en línea, han ido aportando nuevos conocimientos (Dubois, 1987; Lancaster, 1995; Svenonius, 1986). En 1978, una evaluación de la base de datos CANCERNET concluyó que el lenguaje documental podía mejorar la *recall*, pero no en temas nuevos, para los cuales se desempeñaba mejor la lengua natural (en este caso solo el título del documento); el lenguaje controlado aumentaba la precisión, pero no en temas puntuales, que se veían favorecidos por el uso de la lengua natural.

En 1980 se realizaron dos estudios sobre la base de datos ERIC (título y resumen): para el primero la lengua natural aumentaba la *recall* y el lenguaje documental mejoraba la precisión; el segundo detectó mayor *recall* para el lenguaje controlado y una precisión similar para ambos vocabularios.

En 1985, una investigación sobre la base de datos Harvard Business Review

Online encontró una *recall* para la lengua natural (texto completo) del 73,9% y una precisión de 18%; para la lengua natural (título y resumen) 19,3% y 35,6% respectivamente y para el lenguaje controlado 28% y 34%; sin embargo, cuando se utilizaban en forma simultánea la lengua natural y el lenguaje controlado la precisión aumentaba al 78,6%.

En consecuencia, la conclusión a la que se ha llegado es que la combinación de ambas modalidades garantiza la calidad de la recuperación, ya que las dos pueden mejorar tanto la *recall* como la precisión, aunque de diferente manera (Aitchison et al., 1997; Dubois, 1987; Lancaster, 1995; Svenonius, 1986):

La lengua natural aumenta la *recall* cuando se incluyen resúmenes y texto completo, por la exhaustividad de la cobertura.

El lenguaje documental aumenta la *recall* mediante el control de sinónimos y las relaciones asociativas.

La lengua natural aumenta la precisión por el nivel de especificidad y actualización de sus términos.

El lenguaje documental aumenta la precisión por el uso de términos compuestos, el control de homónimos y polisemas y la utilización de vínculos y roles.

En adición a lo anterior, también se ha demostrado que los usuarios tienen distintas preferencias respecto del vocabulario. Según Pat Ensor (1992), los estudiantes de la Indiana State University mostraron una preferencia por el lenguaje documental (62% habían utilizado LCSH), aunque una importante proporción (48%) había buscado también por palabras claves de título; solo el 16% de estos usuarios hicieron búsquedas por autor. En cambio, entre los docentes e investigadores el 36% había utilizado la búsqueda por autor, 44% el lenguaje documental y solo 14% las palabras claves de título.

Raya Fidel (1991a y 1992) también estudió las preferencias respecto a la lengua natural y el lenguaje documental. Sus resultados fueron:

En todas las bases de datos estudiadas, 44% de las búsquedas se hicieron por lenguaje documental y 56% por lengua natural. Al eliminar del estudio las

bases de datos sin lenguaje documental en línea, la proporción era del 50% para cada tipo de lengua.

El 75% de los usuarios consultaron el lenguaje documental antes de la búsqueda, aunque solo el 50% seleccionó un término controlado. Es decir, un 25% ni siquiera consultó el lenguaje documental y otro 25% no encontró un término controlado apropiado.

Si bien el 56% de todas las búsquedas se hizo por lengua natural, este porcentaje se redujo al 39% en ciencias sociales y humanidades y aumentó al 76% en ciencia y técnica. Los bibliotecarios están conscientes de que esto se debe a que la terminología en ciencia y técnica es más precisa, pero los usuarios dieron las siguientes razones: a) el 22% tenía la creencia de que era mejor utilizar los términos en lengua natural para obtener referencias más relevantes; b) el 57% manifestó razones relacionadas con las bases de datos, por ejemplo que el lenguaje documental no era confiable (16%), el término no estaba incluido en el lenguaje documental (13%), la búsqueda se hizo en varias bases de datos con distintos lenguajes documentales (15%), o no se tenía acceso al lenguaje documental en línea (13%).

Acceso alfabético y acceso sistemático

El catálogo de una biblioteca puede ordenarse en forma sistemática utilizando un esquema de clasificación, o en forma alfabética mediante un lenguaje documental de representación verbal como las listas de encabezamientos, los tesauros o la lengua natural. Las ventajas y desventajas de ambos sistemas se presentan en la Tabla 2.

Nancy Williamson (1989b), Profesora de la University of Toronto, Presidente del Comité de Investigación en Clasificación de FID y Directora del proyecto de LCC en línea, considera que los sistemas de clasificación han recibido poca atención por parte de los diseñadores de bases de datos, debido a las desventajas mencionadas en la Tabla 2 y que es necesario que su potencial para el almacenamiento y recuperación de información sea estudiado en profundidad, tanto en los catálogos como en las bibliografías en línea y las bases de texto completo.

No obstante, se han desarrollado algunas experiencias en diversos países

que vinculan el acceso alfabético (lengua natural o vocabulario controlado) con el acceso sistemático (Koh, 1995; Shoham y Yitzhaki, 1994).

En los Estados Unidos se ha implementado una modalidad de clasificación para el catálogo en línea que consiste en asignar a cada registro documental una única notación tomada de LCC o de CDD, con el propósito de agrupar los documentos dispersos por el orden alfabético de los encabezamientos (LCSH) bajo una misma clase.

Ray R. Larson (1991a), de la University of California at Berkeley, ha desarrollado el sistema Cheshire que funciona de la siguiente manera: 1) los registros documentales son copiados de un servicio de procesos técnicos cooperativos; 2) la signatura topográfica asignada por Library of Congress (LCC) es reprocesada, eliminando la signatura librística; 3) la signatura de clase, convertida así en una notación, es exportada a un registro recolector, que reúne todos los encabezamientos de materia y los títulos de los documentos que responden a esa notación. En el momento de la búsqueda se procede en dos etapas: cuando el usuario ingresa sus términos, el sistema le presenta en pantalla el registro recolector; entonces el usuario selecciona el título del documento o el encabezamiento de materia y el sistema, automáticamente, recupera y exhibe los registros documentales.

Por su parte, Mary Micco (1991) de la Indiana University of Pennsylvania ha implementado otro sistema en el que los términos del usuario (lengua natural) son vinculados con los encabezamientos de materia (LCSH) más relevantes y la notación de LCC. Esta notación agrupa diversos epígrafes, representando distintos enfoques de un mismo tema.

Como se ha señalado, estas experiencias han tenido el propósito de reunir, bajo una sola notación, los documentos dispersos por los encabezamientos de materia y en este sentido los autores se muestran conformes con los resultados. En el trabajo de Larson (1991a) 30.471 registros documentales que contenían 221.042 términos de indización fueron reunidos bajo 8.435 notaciones; Larson opina que esta modalidad permite suplir las relaciones equivalentes y asociativas de un tesoro.

Por otra parte, al utilizar la signatura de clase ya asignada por Library of Congress y copiada de los servicios de procesos técnicos cooperativos, es posible lograr que el procesamiento para convertirla en una notación y agrupar los

Tabla 2. Ventajas y desventajas del acceso alfabético y sistemático
(Bates, 1988; Foskett, 1996; Taylor, 1992)

ACCESO ALFABETICO	ACCESO SISTEMATICO
<i>Ventajas</i>	<i>Desventajas</i>
- Orden evidente por sí mismo (acceso directo)	- El orden no es evidente por sí mismo y requiere un índice alfabético (acceso indirecto)
- Mayor hospitalidad.	- Menor hospitalidad
- Mayor flexibilidad.	- Menor flexibilidad
<i>Desventajas</i>	<i>Ventajas</i>
- Dispersión de los temas que deben estar juntos, debido al orden alfabético.	- Orden útil: reúne los temas que lógicamente deben estar juntos.
- Las búsquedas genéricas son tediosas.	- Simplifica las búsquedas genéricas.
- La terminología se desactualiza. Presenta conflictos con la lengua natural.	- No requiere cambios continuos para la actualización de la terminología.
- No compatible entre varias lenguas.	- Compatible entre varias lenguas.

registros se haga automáticamente sin intervención de un clasificador, reduciendo de esta forma los costos. Sin embargo, tanto Larson como Micco coinciden en afirmar que en colecciones pequeñas la cantidad de encabezamientos y títulos incluidos bajo una sola clase es mínima, de manera que esta modalidad resulta apropiada solo para colecciones grandes.

Elaine Broadbent (1989) de la University of Utah llevó a cabo un trabajo experimental con el propósito de conectar los encabezamientos de LCSH con las notaciones de LCC; aunque los recursos destinados a su investigación no le permitieron aplicarlo a un catálogo en línea, llegó a la conclusión de que es posible utilizar los encabezamientos como acceso verbal por materia del catálogo y al mismo tiempo como índice del esquema de clasificación.

El Dewey Decimal Classification Online Project fue una investigación realizada en los años ochenta y financiada por el Council of Library Resources, OCLC y Forest Press, para evaluar la capacidad de CDD como instrumento de recuperación por materia en el catálogo en línea. Los resultados han sido informados en diversos trabajos (Markey y Demeyer, 1985; Markey, 1987; Drabenstott et al., 1990). Para este estudio se desarrolló un catálogo experimental en el que se asignó a cada registro una única notación de CDD. Con el fin de facilitar el acceso verbal por parte del usuario, se incluyeron en los registros documentales los siguientes datos: a) la denominación de la clase tomada de los esquemas de CDD, b) el término de entrada del índice relativo de CDD y sus respectivos sinónimos, c) las notas de alcance de CDD, d) el encabezamiento de LCSH más relevante, que con ese fin se registró en primer lugar. Una vez que el usuario ingresaba sus términos de búsqueda (lengua natural), el sistema buscaba la coincidencia con la denominación de la clase, los términos del índice relativo, las notas o el primer encabezamiento y a partir de esta coincidencia agrupaba las notaciones de CDD, exhibiéndolas en pantalla.

Aprovechando que CDD es un esquema jerárquico, se ofreció al usuario la opción de explorar clases más genéricas o más específicas. También podía revisar las notas para una determinada clase u otros términos verbales adicionales. Este catálogo experimental fue consultado por usuarios y bibliotecarios de cuatro grandes bibliotecas de los Estados Unidos, incluyendo Library of Congress. Basándose en las opiniones de los propios usuarios, Markey ha señalado algunos beneficios de la búsqueda en línea utilizando CDD: estimulación de la terminología (ahorro de términos, mayor cantidad de ideas sobre un tema), ubicación del tema en un contexto, organización lógica de los temas, así como mayor capacidad de exploración y exhibición de las disciplinas que tratan los temas de interés para el usuario.

Una característica de esta experiencia fue la asignación automática de los términos verbales por medio de la notación, logrando una indización menos costosa. El procedimiento propuesto es el siguiente: 1) el registro documental se copia de un servicio de procesos técnicos cooperativos que incluye la notación de CDD y el encabezamiento de LCSH; 2) la biblioteca posee una versión de CDD en línea que incluye la notación y los términos verbales equivalentes (denominación de la clase, términos del índice relativo, notas, etc.); 3) el sistema

hace coincidir la notación del registro documental con la notación de CDD en línea y los términos verbales equivalentes de esta última se copian automáticamente en el registro documental. El promedio de términos de indización por registro fue de 19,6. La denominación de la clase y los términos del índice relativo han resultado tan adecuados para la recuperación como LCSH y las palabras claves de título (Drabenstott et al., 1990).

Otras experiencias con CDD que se han llevado a cabo en Estados Unidos incluyen los trabajos de Liu y Svenonius (1991) de la University of California at Los Angeles y de Micco y Popp (1994) de la Indiana University of Pennsylvania.

Mucho antes que las experiencias anteriores -en 1968- el American Institute of Physics de Estados Unidos desarrolló un proyecto de investigación sobre el uso de CDU (edición especializada para ciencia y tecnología nuclear) en sistemas automatizados, incluyendo un catálogo en línea denominado AUDACIOUS. En este catálogo, el usuario ingresa sus términos de búsqueda en lengua natural y el sistema le muestra en pantalla todas las clases de CDU cuya denominación incluya dichos términos. Una vez que el usuario selecciona una clase, el sistema exhibe el esquema jerárquico de CDU, permitiendo la exploración de lo general a lo particular y vice versa (Buxton, 1990). Más adelante, en la sección de Precoordinación y Poscoordinación, se comentan otros aspectos de este sistema.

Un desarrollo similar es el de ETHICS en Zurich, Suiza. Este sistema consta de un componente jerárquico de conceptos; a cada uno de estos conceptos le corresponde una única notación de CDU y numerosos términos verbales, el principal de los cuales es considerado como un descriptor y el resto como sinónimos, formando entre todos un vocabulario controlado. En adición, los términos verbales se encuentran en tres idiomas (inglés, francés y alemán). Cuando el usuario ingresa sus términos de búsqueda puede optar por la lengua natural o por el vocabulario controlado; luego el sistema le muestra en pantalla todas las clases cuya denominación incluya los términos ingresados, permitiéndole acceder a una exhibición sistemática y explorar las clases más genéricas o más específicas. El usuario también puede solicitar su búsqueda por la notación de CDU, en cuyo caso el sistema exhibe directamente el orden sistemático. ETHICS cuenta con más de 50.000 clases de CDU y 1,3 millones de descriptores (Buxton, 1990; Satiya, 1988).

Teniendo en cuenta el tratamiento del componente verbal de ETHICS, Buxton (1990) considera que esta modalidad de vincular los descriptores y sus sinónimos con el esquema jerárquico de CDU equivale a disponer de un tesoro universal. Es decir, el acceso alfabético o verbal del esquema de clasificación puede desarrollarse de acuerdo con las pautas de control de vocabulario de la norma ISO 2788 para la construcción de tesauros, mientras que las relaciones jerárquicas y asociativas son aportadas por el esquema de clasificación.

También se ha señalado que los tesauros, a pesar de su naturaleza sistemática, por lo general se utilizan solamente como instrumentos de acceso verbal, cuando en realidad sus relaciones equivalentes, jerárquicas y asociativas deben convertirse en una ayuda disponible para el usuario en la etapa de recuperación (Weinberg, 1995; Williamson, 1989b).

Precoordinación y poscoordinación

Tanto la lengua natural como los lenguajes documentales permiten coordinar sus palabras claves, descriptores, encabezamientos o notaciones, con el fin de lograr una mejor descripción del contenido del documento. Hay dos formas de lograr la coordinación: la primera consiste en coordinar los términos independientes en una cadena antes de registrar la información en el catálogo y se denomina *precoordinación*; la segunda se realiza al recuperar la información del catálogo, coordinando los términos independientes mediante un instrumento físico (en la computadora esta posibilidad la brindan los operadores booleanos AND, OR y NOT), en cuyo caso se denomina *poscoordinación* (Aitchison et al., 1997; Foskett, 1996; Lancaster, 1986 y 1995). Como resume Svenonius, la *precoordinación* constituye una indización por cadenas de términos, mientras que la *poscoordinación* es una indización por términos independientes (Holley, 1993). En la *precoordinación* el esfuerzo de combinar los términos en la cadena corresponde al indizador y por lo tanto la indización resulta costosa, mientras que en la *poscoordinación* el esfuerzo y el costo se transfieren al usuario -o al referencista- y en consecuencia la indización es más barata, ya que no requiere aplicar ninguna sintaxis.

Los lenguajes documentales universales más conocidos son *precoordinados*, por ejemplo LCSH, LCC, CDD o CDU, todos los cuales se utilizan ampliamente

en sistemas en línea (Bates, 1988; Buxton, 1990; Chan, 1990; Drabenstott, 1990; Gerhan, 1989; Holley, 1989 y 1993; Jouguelet, 1989; Koh, 1995; Larson, 1991a; Liu y Svenonius, 1991; McIlwaine, 1994 y 1995; Markey, 1987; Micco, 1991; Rolland-Thomas y Mercure, 1989; Strachan y Oomes, 1995; Trotter, 1995; Vizine-Goetz y Markey, 1989; Williamson, 1989a y 1989b). Existen también lenguajes documentales precoordinados especializados como MeSH de la National Library of Medicine de Estados Unidos, que se utiliza para indizar la base de datos Medline.

Durante cierto tiempo se postuló que los lenguajes precoordinados (LCSH, LCC, CDD, CDU) no eran aptos para la computadora. Sin embargo, la Reunión Satélite de IFLA sobre Indización, realizada en Lisboa, Portugal, en noviembre de 1993, puso de manifiesto que los sistemas precoordinados han sido preferidos para los catálogos en línea de las bibliotecas generales, no sólo en los Estados Unidos, donde se calcula que existen más de mil millones de registros documentales descritos con LCSH (Holley, 1989), sino también en muy diversos países, por ejemplo la red de bibliotecas brasileñas Bibliodata de la Fundação Getulio Vargas (LCSH en portugués), la National Library of Canada (LCSH y CSH), la Biblioteca Nacional y Universitaria de Croacia y la Biblioteca Nacional de Portugal (CDU), la red de bibliotecas de la Bibliothèque Nationale de Francia (RAMEAU, traducción al francés de LCSH), la Deutsche Bibliothek de Alemania y la Biblioteca Nacional de España que utilizan encabezamientos de materia, la Biblioteca Nacional de la República Islámica de Irán (LCSH traducido al iraní junto con LPSH) e incluso la British Library que ha reincorporado LCSH después de haber discontinuado su uso en 1987 (Holley, 1993).

Existen dos modalidades para tratar los lenguajes documentales precoordinados en la computadora. La primera consiste en mantener estrictamente esta característica, registrando la cadena precoordinada en un solo campo; la segunda en almacenar los componentes de la cadena mediante subcampos, con lo cual es posible tratarlos con un cierto grado de poscoordinación (Bates, 1988; Buckland, 1992; Buxton, 1990; McIlwaine, 1994; Taylor, 1992).

Cuando se registra la cadena precoordinada de encabezamientos de materia en un solo campo, el epígrafe principal se coordina con las subdivisiones correspondientes y se registra en un solo campo, de la misma forma que en el catálogo de fichas. Por ejemplo: BIBLIOTECAS ESCOLARES-ARGENTINA-

BIBLIOGRAFIA. Cuando se aplica a los sistemas de clasificación (para la descripción de contenido y la recuperación por materia, no para la ubicación física), la cadena precoordinada se registra también en un solo campo con todos los signos de puntuación. Por ejemplo 026.8(82)(01).

En el momento de la recuperación se tiene que hacer coincidir, exactamente y carácter por carácter, los términos de búsqueda del usuario con los términos de la cadena precoordinada. Esta coincidencia exacta con toda la cadena es casi imposible de lograr, de manera que cuando se elige esta modalidad, se debe disponer de tres características para mejorar la recuperación:

La exhibición de la cadena precoordinada en pantalla, mediante una lista de términos en la que el usuario pueda hacer su selección (véase el Capítulo 5 sobre Exploración y Selección de Términos).

La técnica del truncamiento. Los sistemas permiten por lo general un truncamiento a la derecha, con lo cual es posible recuperar el epígrafe o el número principal. Esta característica ha demostrado ser útil para las clasificaciones jerárquicas, ya que permite recuperar las clases más específicas de una jerarquía. Sin embargo, en la mayoría de los sistemas no se dispone de un truncamiento a la izquierda, de manera que las subdivisiones o los números auxiliares no serían recuperables.

La recuperación palabra por palabra, que permite buscar cada uno de los componentes de la cadena en el caso de los epígrafes; no obstante debe tenerse en cuenta que dispersa los términos compuestos.

El registro de la cadena precoordinada en subcampos consiste en asignar un subcampo al encabezamiento o al número clasificatorio principales y otros subcampos para las subdivisiones y los números auxiliares, para lo cual se recomiendan los programas compatibles con la norma ISO 2709 (Bates, 1988; Buckland, 1992; Buxton, 1990; McIlwaine, 1994).

De esta forma, se puede exhibir la cadena precoordinada en pantalla mediante una lista de términos y, al mismo tiempo, recuperar la información por cualquiera de los componentes de la cadena, manteniendo los términos compuestos y poscoordinando los subcampos mediante los operadores booleanos.

En los formatos tipo MARC se cuenta con seis campos destinados a la

descripción de contenido por términos verbales, identificados con los tags 6xx. Cada uno de estos campos puede dividirse en varios subcampos: \$a encabezamiento principal, \$x subdivisión por asunto, \$y subdivisión cronológica, \$z subdivisión geográfica (Taylor, 1992).

Por su parte, la guía para el uso de la CDU, que incluye un capítulo dedicado a las aplicaciones de CDU en sistemas computarizados, presenta distintas experiencias y recomendaciones. Entre estas últimas se destaca el tratamiento de la cadena precoordinada en subcampos, así como su referencia específica a Microsis como uno de los softwares aceptables para esta modalidad (Buxton, 1990; McIlwaine, 1994).

CDU presenta ciertas dificultades para su tratamiento en subcampos, que han sido resueltos de diferentes maneras:

La cantidad de signos de puntuación ofrece problemas en cualquier software, porque éstos pueden ser interpretados por la computadora para otras funciones. En algunos sistemas estos signos se han cambiado por indicadores de faceta literales, por ejemplo en el catálogo en línea AUDACIOUS del American Institute of Physics, según se muestra en la Tabla 3 (Buxton, 1990).

Los números auxiliares comunes (le. gua, forma, lugar, raza o nación, tiempo) no presentan mayores dificultades para ser tratados en subcampos, pero los auxiliares especiales no pueden separarse del número de clase principal porque cambian su significado.

Si bien los signos combinatorios (:, ::, +, []) pueden ser sustituidos por los operadores booleanos, la extensión (/) presenta dificultades especiales. Como es sabido, el uso de la barra implica un rango: 643/645 incluye las clases 643, 644 y 645; el período "1660/1700" abarca también todos los años desde 1661 hasta 1699. Para recuperar los números intermedios del rango se requiere una programación especial. Esta programación debe permitir la recuperación de los números del rango en forma independiente. Por ejemplo, 643/645 tendría que ser recuperable por la fórmula 644 or 645. De no ser posible la programación especial para tratar la extensión, es necesario registrar todos los números del

rango como ocurrencias repetibles y recuperarlos mediante la coordinación con el operador OR, lo cual puede resultar sumamente incómodo tanto en el almacenamiento como en la recuperación.

Tabla 3.
Indicadores de faceta para CDU en el catálogo en línea AUDACIOUS

Faceta	Indicador de faceta de CDU	Indicador de faceta de AUDACIOUS	Ejemplo
Número de clase	n	Un	U55152437
Auxiliares comunes independientes			
- Lengua	=n	En	E30
- Forma	(0n)	Fn	F843
- Lugar	(mn)	Pmn	P265
- Raza	(=n)	Rn	R30
- Tiempo	«n»	Tn	T65
Auxiliares dependientes			
- Punto de vista	.00n	Vn	U55V22
- Especiales	-n .0n	Wn Xn	U62W45 U62118X46
Conexiones			
- Síntesis	n'n	nYn	U54632Y131
- Extensión	n/n	Coordinación con operadores booleanos	U543 or U544 or U545 or U546 (o por programación)
- Relación	n:n	Idem	U543 and U546
- Adición	n+n	Idem	
- Subordinación	n[n]	Idem	

n = dígito del 0 al 9; m = dígito del 1 al 9.

Hay que recordar que la precoordinación produce una indización más costosa; sin embargo, no existen por el momento lenguajes documentales universales que permitan una indización poscoordinada en las bibliotecas generales.

Algunas alternativas como la conciliación de varios tesauros o el uso de clasificaciones facetadas se discuten en el Capítulo 6 sobre Lenguajes Documentales.

5. Exploración y selección de términos

Una de las recomendaciones que se han hecho para facilitar la coincidencia entre el vocabulario del usuario y el del sistema es la posibilidad de exhibir en pantalla una lista de términos que permita explorarlos y seleccionarlos (Chan, 1990). A continuación se presentan algunas opciones.

Índice de la base de datos

Es la opción más económica porque el sistema lo genera automáticamente, sin necesidad de adquirir otros elementos ni incorporar tareas adicionales.

Tiene la enorme ventaja de ofrecer al usuario un 100% de garantía documental, ya que cada uno de los términos del índice responde, por lo menos, a un documento de la colección, y en consecuencia la exploración y selección de términos a partir del índice de la base de datos elimina las búsquedas fracasadas (cero registros recuperados).

Sin embargo, la desventaja del índice es su desorden. En algunas bases de datos es imposible saber si el término seleccionado es una palabra clave de resumen o de título, un descriptor o un epígrafe, si un determinado nombre corresponde al autor o al tema del documento, etc. Tampoco cuenta con un sistema de referencias cruzadas que permita navegar entre los términos de una manera estructurada.

Lenguaje documental en línea

Esta opción tiene un costo adicional, ya que es necesario adquirir el lenguaje documental en una versión informática, instalarlo en el sistema y conectarlo con la base de datos del catálogo, de manera que el término seleccionado por el usuario pueda recuperar los registros documentales correspondientes.

El lenguaje documental en línea tiene la ventaja de facilitar la navegación entre los términos siguiendo la estructura de referencias cruzadas (véase, véase

además, x, xx, USE, UF, BT, NT, RT, etc.), así como las notas de alcance, todo lo cual favorece la exploración y selección de los términos de búsqueda.

La desventaja es que no brinda un 100% de garantía documental, ya que la mayoría de las bibliotecas no utilizan la totalidad de los términos controlados incluidos en el lenguaje documental y, por lo tanto, el usuario puede seleccionar un término que no responda a ningún documento de la colección.

Por otro lado, toda biblioteca tiene necesidad de crear términos que no se encuentran en el lenguaje documental, por ejemplo los identificadores personales, institucionales y geográficos, de manera que cuando el usuario explora el lenguaje documental en línea no tiene acceso a estos términos y pierde la información.

Lista de autoridades

Esta tercera opción es la más costosa, porque implica crear una base de datos para la lista de autoridades, conectarla con aquélla del catálogo en línea y mantener ambos actualizados y coherentes.

Sin embargo, es la opción más eficaz, ya que por un lado permite una exploración estructurada de acuerdo con el sistema de referencias cruzadas y por el otro, al incluir sólo los términos realmente usados o *autorizados* por la biblioteca, le brinda al usuario un 100% de garantía documental.

Existen diversas listas de autoridades que han sido compiladas en otros países; por ejemplo Library of Congress de Estados Unidos, la British Library de Gran Bretaña y la National Library of Canada firmaron en 1997 un convenio para la creación de la lista Anglo American Authority File como parte del PCC, mientras que el proyecto COBRA de las bibliotecas nacionales europeas desarrolla una actividad similar (CENL, 1997; Jáudenes Casaubón, 1996; Martínez et al., 1997; Taylor, 1992; Tillett, 1989).

Lista sistemática

Las opciones de exploración mencionadas responden a un orden alfabético, pero resulta muy útil para el usuario que el sistema exhiba en pantalla una lista ordenada de acuerdo con un sistema de clasificación o con el componente

sistemático de un tesoro, ya que le permite realizar una exploración lógica, por ejemplo de lo general a lo particular (Koh, 1995).

En este sentido, son interesantes los formatos de autoridades para datos de clasificación USMARC y UNIMARC desarrollados por Library of Congress e IFLA, respectivamente (Guenther, 1992; Woods, 1996).

Es importante también tener en cuenta la presentación de la clasificación en pantalla. Como ya se indicó, la notación no es un término de indización amigable para el usuario, porque su significado no es evidente por sí mismo. En el sistema desarrollado por Micco (1991) para la Indiana University of Pennsylvania, durante la etapa de recuperación las clases se presentan en pantalla por su denominación y no por la notación. En AUDACIOUS y ETHICS se muestran simultáneamente la denominación y la notación de cada clase (Buxton, 1990).

Signatura topográfica

La signatura topográfica ofrece la posibilidad de que el usuario *explore los estantes* de la biblioteca desde su propia computadora como una forma de complementar el acceso por materia, ya que una vez que se han elegido uno o varios documentos mediante una búsqueda convencional, se pueden revisar las signaturas topográficas y localizar otros documentos que se encuentran adyacentes en el estante porque tratan sobre la misma materia (Klugman, 1989; Koh, 1995).

Exploración y selección de términos para la descripción

Las opciones mencionadas anteriormente, en especial el lenguaje en línea y el catálogo de autoridades, brindan también al bibliotecario la posibilidad de explorar y seleccionar los términos para la descripción de contenido, con la ventaja de poder indizar en pantalla, así como de verificar los términos y realizar cambios globales en algunos o en todos los registros, incluyendo la corrección de errores de tipeado u ortografía y el reemplazo de términos obsoletos (Chan, 1990).

El lenguaje documental en línea permite asignar términos nuevos que no se han usado en el catálogo, pero como ya se mencionó, cada biblioteca tiene

necesidad de incorporar términos controlados propios que no se encuentran en el lenguaje documental en línea. Por el contrario, el catálogo de autoridades permite seleccionar los términos que ya han sido usados por la biblioteca, incluyendo aquéllos definidos internamente, pero no permite asignar términos nuevos que aún no hayan sido autorizados. De manera que el lenguaje documental en línea y el catálogo de autoridades se complementan para la descripción de contenido, siendo conveniente contar con ambos; sin embargo, es posible que el costo de esta combinación no se justifique en bibliotecas pequeñas.

6. Lenguajes documentales

Listas de epígrafes

LCSH se publicó por primera vez entre 1909 y 1914 con el título de *Subject headings used in the dictionary catalogs of the Library of Congress*. A partir de la undécima edición (LCSH11) de 1988 se decidió agilizar la actualización de la lista, de modo que la versión impresa en cuatro volúmenes tiene una frecuencia anual y un suplemento denominado *Weekly subject list*, mientras que la versión en microfichas se actualiza trimestralmente y en cinta magnética una vez por semana (en 1997 se discontinuó el CDROM); las nuevas modificaciones están disponibles en Internet. LCSH20 se publicó en 1997 con 234.000 epígrafes y referencias y un crecimiento anual de unas 10.000 entradas (Library of Congress, 1997).

Durante mucho tiempo los bibliotecarios estadounidenses expresaron sus críticas a LCSH (Bloomfield, 1993; Cochrane, 1984; Dwyer, 1987; Holley y Killheffer, 1982; Klugman, 1989; Reynolds, 1989; Studwell, 1985, 1987, 1993; Studwell y Folsom, 1993) e incluso Henige (1987) se permitió sugerir la eutanasia como solución para esta lista; las críticas aumentaron durante el período 1970-1979 y más aún en 1980-1989 (Shubert, 1991), coincidiendo con la implementación y evaluación de los catálogos en línea. Entre los defectos detectados se incluyen la desactualización terminológica, la falta de coherencia y especificidad de los epígrafes, la carencia de principios teóricos para su construcción, así como su red asistemática de referencias cruzadas.

En 1984 Pauline A. Cochrane publicó una serie de seis artículos sobre el acceso por materia en el catálogo en línea, a manera de curso de actualización. La Lección 3 fue dedicada íntegramente a LCSH, con la participación de varios bibliotecarios (Cochrane, 1984). Entre ellos, Bates, Mandel y Mischo coincidían en que los defectos de LCSH podían solucionarse de varias maneras: a) poniendo más atención en hacer de este lenguaje documental un instrumento de recuperación y no sólo de almacenamiento, b) convertirlo en un tesoro, c)

distribuirlo en versión informática para facilitar el acceso en línea, d) conciliarlo con tesauros especializados o con esquemas de clasificación para proporcionarle una estructura sistemática y e) estableciendo un sistema interbibliotecario para su continuo perfeccionamiento.

En cambio Beckman, Harper y Buckland consideraban que LCSH sería sustituido a largo plazo por el procesamiento de la lengua natural con sistemas expertos y la poscoordinación con los operadores booleanos. De manera que LCSH, complementado con el acceso por lengua natural, sería suficiente para el período de transición (Cochrane, 1984).

Otro de los requerimientos que se han hecho a Library of Congress ha sido la redacción de un código para los encabezamientos de materia, en el que coinciden diversos autores como Chan (1989), Reynolds (1989) y el persistente Studwell (1985, 1987, 1993; Studwell y Folsom, 1993), entre otros. Hay dos propuestas para el código: una declaración de principios semejante a la Conferencia de París sobre Catalogación de 1961 (Chan, 1989), o bien un sistema de principios y reglas similar a AACR2, que establezca los fundamentos conceptuales de los encabezamientos de materia (teoría) y las reglas para su asignación (práctica) y que sirva además como parámetro de evaluación del catálogo de materia (Reynolds, 1989; Studwell, 1985 y 1993).

En respuesta a estas solicitudes, Library of Congress ha implementado distintas iniciativas:

Mejoras en la lista: En 1986 LCSH cambió la famosa y centenaria estructura *sindética* inventada por Charles A. Cutter por códigos de tesauro, según se muestra en la Tabla 4. Si bien este cambio ha significado una ayuda porque permite distinguir ciertas relaciones jerárquicas y asociativas, hay que tener en cuenta que LCSH no ha sido convertida en un tesauro, porque sigue siendo un lenguaje precoordinado y porque el estudio a fondo de las relaciones entre los epígrafes queda aún pendiente, en especial si se tiene en cuenta que los cambios se hicieron en forma mecanizada (Bates, 1988; Dykstra, 1988; Rolland-Thomas, 1993). En los últimos años también se han revisado y modificado epígrafes que implicaban discriminaciones, por ejemplo hacia los afroamericanos, las mujeres, los grupos lesbigay, etc. (Rogers, 1993).

Vinculación con tesauros especializados o esquemas de clasificación: todas las versiones de LCSH vinculan sus epígrafes con las notaciones de LCC y CDD; esta conciliación se ha visto facilitada por el hecho de que Library of Congress es la responsable intelectual de los tres lenguajes, así como de los formatos, manuales y registros documentales que los incluyen.

Tabla 4. Cambios en los códigos de las referencias de LCSH (1986)

REFERENCIA ANTERIOR	REFERENCIA ACTUAL
See (véase)	USE
x	UF (used for, usado por)
See also (véase además)	BT (broader term, término genérico) o RT (related term, término relacionado)
xx	NT (narrower term, término específico) o RT (related term, término relacionado)

Código para la asignación de epígrafes: Desde 1984 se han publicado cinco ediciones del *Subject cataloging manual: subject headings*, que Library of Congress considera una herramienta suficiente para la asignación de epígrafes, pero que en opinión de otros no sustituye al código (Chan, 1989; Library of Congress, 1996; Pietris, 1987; Studwell, 1993). Sin embargo, existen muchas dificultades que impiden la redacción de este último; Reynolds (1989) opina que hay aspectos filosóficos no resueltos para establecer objetivos y definir conceptos como *materia*, *especificidad* o *relevancia* en relación con los epígrafes y en el aspecto práctico, dado que la elaboración e implementación de AACR2 tuvo un enorme costo para las bibliotecas estadounidenses, Library of Congress se ha negado a emprender una iniciativa de este tipo respecto a los epígrafes.

Cooperación interbibliotecaria: en la actualidad esta lista de epígrafes se elabora mediante un sistema de bibliotecas cooperantes, que forman parte del PCC e integran el subprograma SACO (Tabb, 1996). Además, Library of Congress recibe rutinariamente las sugerencias y comentarios en relación con estos encabezamientos de materia, incluyendo incoherencias, errores, nuevos sinónimos u otras recomendaciones sobre las que se basan después las medidas correctivas.

A pesar de las críticas, LCSH se utiliza ampliamente en bibliotecas generales y algunas especializadas de todos los países de lengua inglesa; ha sido traducida o adaptada a numerosos idiomas y es, con certeza, el lenguaje documental de acceso verbal más utilizado en todo el mundo, debido sobre todo a la inclusión de sus epígrafes en los registros MARC producidos por Library of Congress, British Library, National Library of Canada y muchas otras fuentes de procesos técnicos por copia.

Existen otras listas de epígrafes en inglés como la *Sears list of subject headings*, cuya decimoquinta edición con unos 5000 epígrafes se publicó en 1994, o aquellas especializadas en literatura infantil, por ejemplo *Library of Congress subject headings for children's literature* o *Children's subject headings* publicada por OCLC en 1997, con unos 2000 epígrafes. En general se distribuyen en forma impresa, se actualizan cada 3-4 años, sus epígrafes son compatibles con LCSH y los registros indizados con ellas se pueden adquirir de distintas empresas de servicios bibliotecarios.

En nuestro país hay dos listas de epígrafes muy difundidas, que son LEMB y Sears en español. LEMB fue publicada por primera vez en 1967, con autoría de Carmen Rovira y Jorge Aguayo y bajo la responsabilidad editorial de la OEA y en 1969-1970 se publicaron dos suplementos compilados por la bibliotecaria argentina Emma Linares. Si bien LEMB se basó en LCSH no es exactamene una traducción, sino una adaptación. Quince años después, en 1985, el ICFES de Colombia se hizo cargo de la lista y publicó su segunda edición que incluye unos 15.000 epígrafes, sin que se haya actualizado desde entonces.

En cuanto a Sears, la primera edición en español fue dirigida por Carlos Víctor Penna y publicada en Buenos Aires en 1949, por la editorial ACME; la segunda edición apareció ¡treinta y cinco años después!, en 1984, dirigida por

Carmen Rovira y publicada por la editorial H.W. Wilson de Nueva York, hecho que hacía suponer una continuidad en su publicación, que por desgracia no ha existido. Ninguna de las dos listas cuenta con una versión informática.

Otras listas de epígrafes en español son las del CIRBIC de España y el Epigrafiario de la Biblioteca Central de Ciencia y Técnica de Cuba, publicados en 1988 y 1987 respectivamente; un estudio comparativo entre estas listas y LEMB2, realizado por Rodríguez Ricard y Trejo Franqui (1989), mostraba que LEMB2 seguía siendo la mejor.

Quizás las alternativas más actualizadas en nuestro idioma sean: a) la Lista de autoridades de materia de la UNAM, publicada en México en forma impresa (UNAM, 1997) y b) la lista de autoridades de la Biblioteca Nacional de España que incluye 300.000 registros entre nombres, títulos uniformes y encabezamientos de materia, se actualiza semestralmente y que distribuye en CDROM la empresa Chadwyck-Healey (BNE, 1997), pero ninguna de estas lista cuenta con una amplia difusión en Argentina.

Sistemas de clasificación

LCC fue publicada por primera vez entre 1899 y 1940, elaborada específicamente para la colección de Library of Congress, que la utiliza en sus signaturas de clase. Se trata de un esquema de clasificación universal y jerárquico, pero enumerativo, no sintético. Está dividido en esquemas individuales, publicados en 45 volúmenes y posee un ordenamiento por clases principales, una notación alfanumérica que utiliza letras mayúsculas para clases y subclases y números arábigos para las subdivisiones, así como una tabla especial para el nombre del autor con el fin de lograr subordenamientos alfabéticos. La actualización se realiza mediante las ediciones revisadas de los esquemas individuales y las publicaciones *Library of Congress classification: additions and changes* y *Library of Congress classification schedules: a cumulation of addition and changes* (Foskett, 1996; Taylor 1992).

Nancy Williamson, de la University of Toronto, trabajó durante varios años en la informatización de LCC, tarea más que difícil por el carácter enumerativo de este sistema de clasificación, que cuenta con más de 500.000 clases. Sin

embargo, la implementación del formato USMARC para clasificación hizo posible concretar, entre 1994 y 1995, la base de datos del esquema y emprender su revisión. En 1996 Library of Congress publicó la primera edición de *Classification Plus*, un CDRom que contiene los esquemas individuales ya revisados de LCC, que se actualiza anualmente (Guenther, 1996; Williamson, 1994).

LCC ha sido traducida al español por el CUIB de la UNAM en México y es utilizada en diversos países hispanohablantes, aunque no en Argentina. No existe una versión en línea en esta lengua.

Por su parte, CDD es el esquema de clasificación más antiguo entre los existentes; se publicó por primera vez en 1876 como un folleto anónimo de 44 páginas titulado *A classification and subject index for cataloguing and arranging the books and pamphlets of a library*, que contenía menos de mil clases y a partir de la segunda edición (1885) se ha publicado bajo el nombre de su autor Melvil Dewey. La última edición es CDD21 de 1996, que incluye más de 20.000 clases en cuatro volúmenes; en el cuarto volumen se ha mantenido el manual de procedimientos que se introdujo en CDD20. Las principales revisiones se realizaron sobre las clases Administración Pública, Educación, Biología, Religión y los períodos históricos, así como la terminología empleada en las denominaciones de las clases y en el índice (Mitchell, 1995 y 1996).

Si bien la distribución comercial de CLD le sigue perteneciendo a la empresa Forest Press (desde 1988 una división de OCLC), la responsabilidad intelectual, es decir su revisión y actualización, está a cargo de la Decimal Classification Division de Library of Congress (Mitchell, 1995). La publicación *Dewey decimal classification additions, notes and decisions*, conocida popularmente como DC& y de frecuencia anual, incluye las modificaciones del esquema. Cuenta también con una versión abreviada, cuya última edición es la decimotercera de 1997, basada en CDD21 (Foskett, 1996; Sweeney, 1983; Taylor, 1992).

Como es sabido, en la actualidad CDD es un lenguaje documental precoordinado apto para ordenar tanto el estante como los catálogos y las bibliografías (Foskett, 1996; Sweeney, 1983; Taylor, 1992) y está siendo utilizado para ordenar los sitios Internet, sobre todo de numerosas bibliotecas públicas estadounidenses (Forest Press, 1998). No obstante, debido que la coordinación de las notaciones puede dar lugar a signaturas de clase muy

extensas, muchas bibliotecas han decidido reclasificar sus estantes con LCC (Chressanthis, 1995; Pattie, 1995).

En 1988 la Forest Press logró una versión en línea de CDD con fines meramente editoriales. En 1993 se publicó por primera vez *Electronic Dewey*, la versión de CDD20 en CDROM, revisada en 1994, que contaba con nueve índices y vinculaba las notaciones de CDD con los epígrafes de LCSH (Carpenter, 1995; Trotter, 1995). En 1996 y basada en CDD21, apareció la primera versión de *Dewey for Windows*, también en CDROM (Beall, 1996).

Library of Congress incluye la notación de CDD en sus registros USMARC, con lo cual este sistema de clasificación tiene garantizada la amplia difusión de la que siempre ha gozado. En Estados Unidos el 95% de las bibliotecas públicas y escolares, el 25% de las universitarias y el 20% de las especializadas utilizan CDD. Además, ha sido traducida a 30 idiomas y se utiliza en más de 200.000 bibliotecas de 135 países.

En español se han publicado tres ediciones de CDD; las dos primeras (1955 y 1980) fueron en realidad adaptaciones que contemplaron distintos aspectos propios de los países iberoamericanos, por ejemplo en las clases Religión, Derecho, Literatura, Geografía e Historia. La tercera edición, que es una mera traducción de CDD20 en inglés, se publicó en 1995, apenas unos meses antes de que apareciera CDD2 (decisión más que lamentable) y no se encuentra en versión informática.

CDU es el sistema de clasificación desarrollado por Paul Otlet y Henri LaFontaine en 1895, basándose en CDD5, por lo cual mantiene las características de ser universal, jerárquico y decimal; sin embargo, su finalidad original fue clasificar una bibliografía universal y no el estante. La primera edición del esquema contenía 33.000 clases y varias tablas auxiliares que han dado a CDU una importante capacidad de síntesis.

En la década de los ochenta, toda la estructura de gestión de CDU entró en crisis por razones económicas y de organización interna. En 1990 la FID decidió que el futuro desarrollo del esquema quedara en manos de los editores responsables de CDU en los distintos idiomas, de manera que el 1 de enero de 1992 se constituyó el Consorcio CDU, con la participación de la FID y los editores de CDU en holandés, inglés, francés, japonés y español (Strachan y Oomes, 1995).

Hasta 1990 se mantenían varias ediciones de CDU según su contenido, a saber: completa (150.000 clases), media (30% de las clases), abreviada (10%) y resumida (2-3%), todas las cuales han sido discontinuadas. En 1991, luego de varios años de discusiones y trabajo, se creó una base de datos denominada Master Reference File (MRF), con el propósito de constituir una única edición de CDU en soporte informático. El MRF cuenta con unas 60.000 clases y se basó en la edición media internacional de 1985 elaborada en Gran Bretaña (BSI 1000), adicionada con las extensiones y correcciones hasta 1992 y las ediciones medias japonesa, húngara, serbocroata y francesa. La versión en disquete de CDU se publicó por primera vez en 1993 y en ese mismo año apareció la primera guía para el uso de este esquema, que es un documento oficial del Consorcio CDU.

El uso de CDU se encuentra muy extendido en todo el mundo; ha sido traducida a 26 idiomas y las modalidades introducidas en su sistema de gestión hacen suponer una actualización más ágil y una continuidad en las mejoras; se están llevando a cabo estudios con el fin de lograr un esquema completamente facetado de CDU y de facilitar su uso como tesoro (McIlwaine, 1995; McIlwaine y Williamson, 1994; Williamson, 1996).

La notación de CDU no se encuentra en los registros USMARC de Library of Congress y tampoco ha sido tomada en cuenta para el formato USMARC de clasificación y algunos temen que esta circunstancia atente contra su difusión y uso en el futuro (McIlwaine, 1995); sin embargo, los formatos y la metodología UNIMARC de IFLA le brindan a CDU la misma relevancia, o más, que a LCC y CDD, incluyendo el desarrollo del formato UNIMARC para clasificación (Woods, 1997).

En 1995, AENOR, organismo responsable de CDU en español, publicó la nueva edición con el nombre de *séptima edición abreviada en español*, que es una denominación incorrecta y se presta a confusiones; afortunadamente se trata de la primera edición española del MRF, que ha sido acompañada de su correspondiente versión en disquete; en 1994 AENOR había traducido la guía de uso elaborada por McIlwaine, de manera que CDU se ha convertido en el lenguaje documental universal más actualizado en lengua española, con su respectiva guía y el único que cuenta con una versión informática (AENOR, 1995; McIlwaine, 1994).

Otros lenguajes documentales

Los principales avances en relación a los lenguajes documentales incluyen los tesauros, las clasificaciones facetadas, PRECIS y la conciliación de distintos lenguajes.

El tesoro es un lenguaje documental cuyos términos (descriptores, no descriptores e identificadores) conforman un vocabulario controlado y se estructuran en un sistema de relaciones equivalentes, jerárquicas y asociativas, de manera que un mismo lenguaje proporciona el acceso alfabético y el acceso sistemático. Es además poscoordinado, es decir cada uno de sus descriptores representa un único concepto, mientras que el conjunto de descriptores de un mismo registro representa el contenido global del documento (Aitchison et al., 1997; Bates, 1988), de esta forma se adapta mejor al uso de los operadores booleanos en la computadora y la indización resulta menos costosa. Desde su aparición en la década de los cincuenta se ha registrado una proliferación de tesauros especializados, pero desafortunadamente no existe un tesoro universal que pueda utilizarse en las bibliotecas generales.

En cuanto a las clasificaciones facetadas se ha demostrado su eficiencia en sistemas poscoordinados, en la medida en que las clases producidas por el análisis facetado estén representadas por una notación tal que permita poscoordinarla con los operadores booleanos y al mismo tiempo posibilite la exploración jerárquica por medio del truncamiento (Gödert, 1991). BC es una clasificación universal elaborada originalmente por Henry E. Bliss para organizar la información en catálogos y bibliografías, no para el estante. Aunque Bliss era estadounidense, BC tuvo mayor difusión en Gran Bretaña y en la década de 1970 la Bliss Classification Association decidió emprender una segunda edición, desarrollada por miembros del CRG bajo la dirección de Jack Mills, de acuerdo con los conocidos principios de Bliss y de Ranganathan, que la han convertido en una clasificación totalmente facetada. Su publicación se inició en 1977 y se encuentra aún en curso, faltando aproximadamente la mitad de los esquemas por publicar (Thomas, 1995). Aunque hace algunos años se vaticinaba que podría llegar a sustituir a CDD y CDU (Foskett, 1996), el escaso presupuesto de la Bliss Classification Association, que ha demorado la publicación del esquema completo, hace dudar de su capacidad de revisión y actualización periódica en el futuro

(Taylor, 1992). Otra clasificación facetada y universal es CC, elaborada por Shiyali R. Ranganathan en 1933 y cuya séptima edición fue publicada en 1988 por la Sarada Ranganathan Endowment; a pesar de sus sesenta años de existencia este esquema no ha logrado imponerse en las bibliotecas generales ni competir en difusión y uso con los tradicionales CDD y CDU. Ni BC2 ni CC7 han sido traducidas al español.

Por su parte PRECIS es un sistema creado en Gran Bretaña por Derek Austin, que entró en servicio en la British National Bibliography en 1971. Según Foskett (1996) mientras LCSH sigue la tradición estadounidense de Cutter, PRECIS sigue la tradición británica de Coates y del CRG. Para Dykstra (1989) este sistema reúne todos los avances logrados en relación con la descripción de contenido en la última mitad de este siglo. PRECIS no es un lenguaje documental propiamente dicho, sino un procedimiento que se basa en el concepto de vocabulario *open-ended*, es decir que los términos pueden ser admitidos en cualquier momento en que aparezcan en la literatura. Una vez que el documento ha sido analizado, los conceptos son ordenados como una frase en lengua natural, formando una cadena de términos, cada uno de los cuales está precedido por un operador determinado. Aquellos términos que aparecen en la posición líder son marcados y las estructuras semánticas se añaden en forma de referencias de *véase además* para cualquier rama de los términos de la cadena. También se incluyen los códigos de manipulación para la computadora y ésta se encarga de hacer el resto. Por lo tanto, la entrada al sistema es una cadena de términos codificados. La salida muestra un término líder, un término genérico y uno o varios términos específicos. Un número indicador de referencias vincula los términos con sus referencias cruzadas y un número indicador de materia vincula el término con LCSH, LCC y CDD (Rowley, 1988).

El formato UKMARC posee un campo para PRECIS y la British Library lo incluía en sus registros hasta 1990 (Dykstra, 1989; Foskett, 1996; Taylor, 1992). El sistema se ha difundido en Gran Bretaña, en Canadá y en Australia; aún en Estados Unidos se ha propuesto como sustituto de LCSH y en India se desarrolló un sistema similar llamado POPSI (Taylor, 1992). Lamentablemente, PRECIS resulta costoso debido al tiempo que consume el análisis sintáctico y la codificación de las cadenas; a fines de 1990, la propia British Library se vio forzada a sustituirlo por un sistema menos sofisticado denominado COMPASS

(Holley, 1993; Taylor, 1992) y por el mismo motivo Library of Congress nunca adoptó PRECIS (Holley y Killheffer, 1982).

Conciliación de lenguajes documentales

Con la proliferación de tesauros especializados, se ha producido un aumento en la incompatibilidad de los catálogos y bibliografías en línea, tanto en el nivel de especificidad, como en la exhaustividad, los términos compuestos, los sinónimos y las relaciones entre los términos. La conciliación de estos tesauros incluye las técnicas de conexión y fusión, así como los macrotesauros (Aitchison et al., 1997; Niehoff y Mack, 1985; Rada, 1987; Taylor, 1992).

Un ejemplo de conciliación es el sistema unificado de lenguajes médicos desarrollado por la National Library of Medicine de Estados Unidos (Squires, 1993). Este sistema consta de tres elementos: a) el metatesauro que no pretende reemplazar los lenguajes médicos ya existentes, sino que es una síntesis de ellos lograda por conexión y fusión; b) la red semántica que tiene el propósito de ofrecer una categorización coherente a todos los conceptos del metatesauro y proporcionar una red de relaciones útiles y c) el mapa de fuentes de información que orienta al usuario en la selección de las bases de datos más relevantes para su consulta. También la NASA ha desarrollado un sistema operacional que permite conectar los términos de indización utilizados por el U.S. Department of Defense y los términos utilizados por la NASA, mediante una tabla de traducción o léxico y un programa que determina qué reglas seguir (Silvester y Klingbiel, 1993).

La conciliación de lenguajes documentales no incluye solamente tesauros; un ejemplo es el esfuerzo de FID y Unesco para producir un esquema denominado Broad System of Ordering (BSO), que básicamente es un sistema diseñado para servir como mecanismo de conciliación entre varios lenguajes de indización, en forma de esquema de clasificación, ya que éste es considerado como el lenguaje más predecible, controlado, actualizado, universal y compatible con todos los idiomas. Los criterios que se han postulado son la flexibilidad, la simplicidad y la facilidad de uso en sistemas tanto manuales como automatizados (Foskett, 1996; Niehoff y Mack, 1985; Taylor, 1992).

Finalmente, Douglas J. Foskett (1991) ha postulado que BC2 puede ser el

mejor lenguaje de conciliación. Este esquema facetado tiene una doble característica: a) dado que sus clases principales han sido desarrolladas en esquemas individuales con un alto nivel de especificidad, es posible indizar en profundidad aún la información más específica de cualquier colección especializada y b) todos los esquemas individuales forman un único esquema universal, totalmente compatible, porque al haber sido elaborado con la misma metodología del análisis facetado, estructurado con las mismas reglas y dotado de un mismo sistema de notación, garantiza una homogeneidad global.

El futuro de los lenguajes documentales

Como se ha indicado, los lenguajes documentales tradicionales (LCSH, LCC, CDD, CDU) han sido muy criticados por su carácter precoordinado, su dificultad para procesarlos en la computadora, la obsolescencia de su terminología, su ordenamiento por disciplinas y no por conceptos, etc. Al mismo tiempo cuentan con una amplia difusión internacional, han sido traducidos a varios idiomas y administrados durante décadas por organizaciones responsables (Library of Congress, Forest Press, FID) que han garantizado su continuidad. Difusión internacional y continuidad parecen otorgarle a estos lenguajes suficiente confianza entre los bibliotecarios.

Otras opciones como PRECIS, BC2 o la conciliación de varios tesauros, si bien reconocidos como instrumentos eficaces, no han logrado la misma difusión internacional, ni cuentan todos con el respaldo de una organización que garantice su continuidad con el mismo grado de responsabilidad probada que los anteriores.

La alternativa sería entonces *eficacia versus difusión y continuidad*, pero en el medio de la opción se encuentran las dificultades económicas del último cuarto de este siglo, que en mayor o menor escala afectan a las bibliotecas de todo el mundo.

Fred E. Croxton, Director de Sistemas Automatizados de Library of Congress y Carol Mandel coinciden en afirmar que la inversión en LCSH es enorme y que debe ser protegida (Cochrane, 1984; Mandel, 1991). Como ya se mencionó, sólo en Estados Unidos existen más de mil millones de registros documentales procesados con LCSH, además de contar con una gran difusión

en otros países (Holley, 1989 y 1993); en términos de economía, redescubrir toda esta cantidad de documentos parece una tarea imposible. Para Chan (1990) dadas las enormes implicancias financieras, sería irresponsable sugerir el abandono total de los lenguajes tradicionales y el desarrollo de otros nuevos diseñados especialmente para los sistemas en línea.

Sin embargo, Foskett (1991) ha señalado que en los últimos treinta años muchas bibliotecas estadounidenses reclasificaron sus colecciones pasando de CDD a LCC («un esquema del Siglo 19 por otro esquema del Siglo 19»), con un tremendo costo y sin más beneficio que una signatura de clase más corta para el lomo de los libros; con mucha más razón debería afrontarse el costo de incorporar un lenguaje documental eficiente como BC2. Debe tenerse en cuenta que las nuevas tecnologías, especialmente los sistemas expertos, pueden facilitar la tarea de reclasificación (Williamson, 1989b Vickery y Vickery, 1993) y, por otra parte, dado que la precoordinación produce un almacenamiento más costoso, quizás un ahorro en este sentido permita financiar la implementación de un sistema poscoordinado más eficiente.

Charles Hildreth, investigador de OCLC, considera que aún con todas sus desventajas LCSH es lo más próximo a un vocabulario controlado *común* entre las bibliotecas de los Estados Unidos y sería poco realista abandonarlo (Cochrane, 1984); se podría añadir, teniendo en cuenta las múltiples traducciones y adaptaciones de la lista (Holley, 1993) y las tendencias hacia la globalización, que es lo más cercano a un vocabulario común internacional. Sin embargo, Brian Aveney, Director de Investigación y Desarrollo de Blackwell North America se pregunta por qué los bibliotecarios persisten en lograr un acceso unificado por materia y afirma que la diversidad es el modelo del futuro: la solución del problema del vocabulario puede ser la utilización de varios lenguajes controlados en un mismo catálogo en línea (Cochrane, 1984). Las técnicas de conciliación de lenguajes documentales resultarían entonces de gran utilidad para conformar ambos criterios.

Chan (1990) también hace notar que en la actualidad, aún en Estados Unidos coexisten numerosas bibliotecas con catálogos de fichas, otras con catálogos semiautomatizados y otras con catálogos en línea de diversos grados de sofisticación, de modo que uno de los grandes desafíos que deben satisfacer los lenguajes documentales es prestar servicio a los distintos sistemas en

operación y al mismo tiempo a los sistemas del futuro, muchos de los cuales ni siquiera se han proyectado.

Durante la conferencia *Classification in the Computer Age* realizada en Albany, Estados Unidos, en noviembre de 1988, Carol Mandel describía el momento como un período de transición; si bien reconocía que los lenguajes documentales tradicionales no dan la mejor respuesta a los requerimientos del usuario, resaltaba la necesidad de hacer un uso más eficiente de ellos mientras se encuentra una solución mejor. Para desarrollar este lenguaje del futuro consideraba imprescindible contar con un mayor conocimiento sobre el comportamiento de los usuarios durante la búsqueda. Al cierre de la conferencia, Williamson ratificó que los viejos lenguajes deben ser utilizados hasta que la investigación y el desarrollo tecnológico garanticen la creación de nuevos lenguajes que respondan con eficiencia a la era de la computadora y puedan reemplazar a los actuales (Williamson, 1989b).

Se han hecho también dos predicciones que pueden hacer innecesarios todos los lenguajes documentales para los catálogos de biblioteca. Una de ellas es el procesamiento de la lengua natural por los sistemas expertos como señala Beckman, o el desarrollo de tesauros de búsqueda que permitan al usuario manejar sus términos en lengua natural (Aitchison et al., 1997; Bates, 1988; Lancaster, 1986 y 1995; Larson, 1991b). La otra, postulada por Harper, es la que concibe el catálogo como un simple inventario que facilitaría la ubicación física de los documentos, mientras que la localización intelectual se realizaría por medio de las bibliografías en línea, con lo cual el acceso por materia en los catálogos no sería necesario (Cochrane, 1984).

Como quiera que sea, en el mediano plazo no se vislumbran cambios en el uso de los lenguajes documentales tradicionales para los catálogos en línea de las bibliotecas generales, sino un uso más eficiente de ellos mediante su optimización y el complemento con una recuperación por lengua natural, según se detalló en el Capítulo 4.

7. Los catálogos en línea en Argentina: un desafío para esta década

La XX Reunión Nacional de Bibliotecarios, organizada por ABGRA en Buenos Aires en septiembre de 1984, tuvo como tema central la automatización de las bibliotecas. El peso de la reunión recayó sobre el bibliotecario norteamericano Dana A. Haverkamp de OCLC, quien participó en mesas redondas y conferencias, explicando la evolución de la informática en las bibliotecas de los Estados Unidos, el formato MARC, los programas para procesos técnicos, adquisiciones, circulación y publicaciones periódicas, los catálogos y bibliografías en línea, las redes o sistemas cooperativos desarrollados en ese país y el impacto de la informática entre los bibliotecarios (Haverkamp, 1984).

En esta misma reunión, Constantino V. Menelle, de la Biblioteca de Gas del Estado, presentó un programa pionero en nuestro país, diseñado por esa institución para la gestión bibliotecaria, en el que cada registro se almacenaba en dos fichas perforadas de 80 columnas cada una, con un formato propio que incluía descripción bibliográfica, descripción de contenido, datos de adquisición y de circulación, ¡en 19 campos y 160 caracteres! (Menelle, 1984). La presentación argentina se completó con algunos profesionales que realizaban consultas remotas a bases de datos bibliográficas del extranjero, por ejemplo en el CAICYT. Estas experiencias y algunas otras no presentadas en la Reunión eran, desde luego, la excepción; en los pasillos, durante los intervalos los asistentes comentaban que aquéllo era ciencia ficción. La informática golpeaba la puerta de nuestras bibliotecas, pero faltaba un mecanismo disparador que facilitara el uso masivo de esta tecnología.

El mecanismo disparador fue presentado en Argentina, en diciembre de 1985 por Giampaolo Del Bigio, de la Unesco: el programa CDS/ISIS para microcomputadoras o Microisis. Como es sabido, este programa ha sido desarrollado por la Unesco desde 1975; es un administrador de bases de datos documentales que permite definir distintos tipos de bases, ingresar los datos,

recuperarlos y presentarlos en pantalla o imprimirlos; es además multilingüe y puede intercambiar datos con otros programas. Según el tipo de computadora que se utilice, existen tres versiones del CDS/ISIS: Macroisis para mainframes, Minisis para minicomputadoras y Microisis para microcomputadoras (Del Bigio, 1992). Con Microisis la informática en nuestras bibliotecas ha registrado un importante desarrollo, habiéndose diseñado infinidad de bases de datos, entre ellas catálogos, bibliografías, directorios o bases de texto completo, así como programas de gestión bibliotecaria integrales o en módulos separados, interfaces más amigables, utilitarios varios, hipertextos e hipermedia, etc. (CNEA 1992, 1993).

En 1985 el Subgrupo Formato del Grupo Bases de Datos del SIDCYT presentó un formato común para el ingreso de información a bases de datos bibliográficas, que habría de evolucionar hasta convertirse en el actual FOCAD (CAICYT, 1994). El formato consta de una importante cantidad de elementos descriptivos y ofrece la posibilidad de seleccionar solo aquellos campos y subcampos de interés para cada biblioteca, o bien agregar campos y subcampos adicionales según se requiera. Puede convertir la información ingresada con este formato a otros de uso generalizado a nivel internacional como el formato MARC, e intercambiar la información transfiriendo datos bibliográficos mediante un soporte computacional, independientemente del hardware y software utilizados.

Si bien en Argentina no existe una empresa como OCLC y lamentablemente la Biblioteca Nacional solo puede describirse en términos de LADRILLOS and CEMENTO, se ha logrado acumular una importante experiencia en el área de los procesos técnicos cooperativos. Mientras que el Catálogo Colectivo Nacional Universitario de Libros del SISBI-UBA intenta nuclear a todas las bibliotecas universitarias del país (UBA, 1993), numerosas bibliotecas participan en la elaboración de bibliografías internacionales coordinadas principalmente por organismos del sistema ONU/OEA como AGRIS (FAO) en agricultura, IMIS (OIEA) en energía atómica, CLADES (CEPAL) en ciencias sociales, LILACS (BIREME) en salud y muchas otras. La cooperación se manifiesta además en otros niveles como el Sistema de Bibliotecas Públicas del Partido de Gral. Pueyrredón, integrado por 29 bibliotecas, con un servicio de procesos técnicos centralizados en el Centro Cultural de Mar del Plata que elabora, a partir de una

base de datos en Microisis, las fichas catalográficas impresas que son distribuidas a las distintas bibliotecas del sistema (Vega, 1993).

También se han instalado en nuestro país distintos desarrollos en materia de telecomunicación, como la red Startel (antes Arpac), los sistemas de correo electrónico privados, las redes académicas, así como la flamante incorporación de Internet, que facilitan la transmisión de datos de una computadora a otra con grandes beneficios y bajos costos.

Sin duda, hay otras experiencias importantes que se han escapado a este breve resumen, pero queda claro que aquéllo que en 1984 parecía ciencia ficción hoy es una realidad concreta en numerosas bibliotecas argentinas.

Sobre esta base, ¿qué se puede esperar para los próximos años? Es muy probable que la respuesta a esta pregunta sea el acceso remoto a los catálogos en línea. Si 1985/1995 fue la década de la creación de bases de datos, la próxima parece ser la década del acceso público. La tecnología, el formato, las experiencias, las bases de datos con la información registrada y los recursos humanos existen en el país; de hecho, algunas bibliotecas argentinas ya están implementando estos sistemas. Quizás solo falte un nuevo mecanismo disparador que facilite el camino hacia un desarrollo masivo.

Mientras tanto, ¿cuál es nuestra situación frente a este desafío? En particular con respecto al tema central de esta revisión, es decir la búsqueda por materia y la descripción de contenido en los catálogos en línea, ¿qué se puede anticipar?

Los catálogos en línea de bibliotecas bonaerenses

Con el propósito de conocer las prácticas de descripción de contenido y la búsqueda por materia en los catálogos en línea de bibliotecas de la Provincia de Buenos Aires, la Cátedra de Clasificación I del Departamento de Bibliotecología de FAHCE-UNLP envió un formulario de encuesta a 205 bibliotecas populares (BP) y 113 bibliotecas especializadas (BE), obteniéndose respuestas de 42 BP y 24 BE (21% en ambos casos), con los siguientes resultados preliminares:

El 71% de las BP y 79% de las BE contaban con computadora y más de un tercio de ellas poseían 2 o más; sin embargo, 33% de las BP y 26% de las

BE no utilizaban la computadora ni para la referencia ni para los procesos técnicos, mientras que un 21% adicional de BE la usaban para la referencia pero no para los procesos técnicos.

Solo 4 BP y 5 BE permitían al usuario realizar sus propias búsquedas (acceso público). Podría argumentarse que la falta de equipamiento para el acceso remoto fuera la causa principal de este problema, pero no todas las bibliotecas equipadas con LAN permitían la consulta directa de sus usuarios y por el contrario, bibliotecas que poseían una sola computadora sí la facilitaban, por lo que no parece haber una relación entre el acceso público y la falta de equipamiento; tal vez habría que suponer que no existe una idea clara entre los bibliotecarios locales de que el catálogo en línea debe ser consultado directamente por el usuario.

Respecto a los procesos técnicos, el 50% de las BP y el 90% de las BE utilizaban Microisis, solo 1 BP y ninguna BE utilizaba FOCAD, mientras que 60% de las BP y 30% de las BE no utilizaban AACR2.

Por último, la descripción de contenido en los catálogos en línea de estas bibliotecas se correspondía con aquella de los catálogos de primera generación, ya que todos carecían de resumen, solo 30% de las BP y las BE recuperaban por palabras claves de título, 50% de las BP y 20% de las BE combinaban acceso alfabético y sistemático, 1 sola BP y 3 BE registraban los epígrafes en subcampos, mientras que 2 BP y 1 BE hacían lo propio con los números clasificatorios. Solo 1 BP disponía de catálogo de autoridades y 2 BE de lenguaje documental en línea.

En resumen, estos datos sugieren que no en todas las bibliotecas se planifica la incorporación de la informática, que los catálogos en línea no se han diseñado para uso del público y que de facilitarse el acceso del usuario, éste se encontraría con problemas derivados de la falta de normalización y pocos puntos de acceso por materia, con lo cual cabría esperar un alto índice de fracasos como se describió en el Capítulo 3.

FOCAD y la descripción de contenido

Los campos relacionados con la descripción de contenido en FOCAD se encuentran numerados del 60 al 69. Teniendo en cuenta los capítulos precedentes, a continuación se comparan estos campos con los del formato BIBUN elaborado por el SISBI-UBA (UBA, 1993):

Número clasificatorio: ambos formatos registran la clasificación en el campo 60, pero no se ha contemplado el registro en subcampos como recomienda por ejemplo la guía de uso de CDU (McIlwaine, 1994). De esta forma, la recuperación de cada uno de los componentes de la cadena precoordinada y su poscoordinación es imposible.

Sinónimos verbales del número clasificatorio: FOCAD no ha definido aún ningún campo para esta nueva modalidad, pero BIBUN sí ha destinado para ello el campo 62.

Encabezamientos de materia: FOCAD tampoco ha definido un campo para los epígrafes, a pesar de que posiblemente sea el tipo de lenguaje documental más utilizado en todo el mundo (Holley, 1993). Por el contrario, BIBUN ha destinado el campo 61, pero sin contemplar el registro en subcampos. En este caso, tratándose de términos verbales, puede lograrse una recuperación palabra por palabra de cada componente de la cadena precoordinada, pero los términos compuestos se separan y dispersan produciendo recuperaciones falsas, cosa que se puede evitar registrando los términos de la cadena en subcampos, como sucede con el formato MARC.

Descriptores: ambos formatos han destinado el campo 65 para el registro de descriptores. FOCAD permite incluir los identificadores en este campo o desagregarlos en los campos 66 a 68, aunque en esta segunda opción no se han tenido en cuenta los identificadores institucionales.

Resumen: tanto en FOCAD como en BIBUN es posible incluir un resumen en el campo 69.

En síntesis, FOCAD permite definir adecuadamente un catálogo en línea de segunda generación, pero necesita mejorar el tratamiento de los lenguajes precoordinados, con el propósito de facilitar su utilización en las bibliotecas generales.

8. Conclusiones

Las dificultades de los usuarios para realizar búsquedas por materia en los catálogos en línea, de acuerdo con las experiencias publicadas en otros países, se ha visto reflejada en dos grandes problemas:

El fracaso (búsquedas con cero registros recuperados).

La sobrerrecuperación (búsquedas con más de 50 registros recuperados).

El fracaso ha sido relacionado con la falta de coincidencia entre el vocabulario del usuario y el del sistema; como soluciones se han postulado las siguientes:

Adición de términos: el aumento sustancial de puntos de acceso por materia en cada registro ofrece al usuario más posibilidades de hacer coincidir sus términos de búsqueda con el vocabulario del sistema (criterio cuantitativo). En este sentido, incluir un resumen de 200-250 palabras que incluya los términos de sumarios e índices analíticos resulta de especial utilidad.

Combinación de términos: la disponibilidad de distintos tipos de puntos de acceso por materia (lengua natural y lenguajes documentales, acceso alfabético y sistemático, precoordinación y poscoordinación) permite compensar las desventajas de algunos términos con las ventajas de los otros (criterio cualitativo).

Exploración y selección de términos: la posibilidad de explorar la lista de términos incluidos en el catálogo, tanto en orden alfabético como sistemático, facilita al usuario la selección de aquellos términos que le permitirán obtener 100% de coincidencia y de garantía documental. Para ello resulta

muy útil poder navegar en la lista siguiendo la estructura de referencias cruzadas.

Por otro lado, la sobrerrecuperación ha sido relacionada con la no aplicación o la mala aplicación de las técnicas de búsqueda por parte del usuario, ya sean las formulaciones booleanas, el truncamiento o la adyacencia.

Como soluciones se ha propuesto lograr una interfaz usuario-sistema más eficiente, amigable y normalizada, así como mejorar los programas de instrucción del usuario, poniendo énfasis en los aspectos conceptuales de la búsqueda.

El catálogo en línea de acceso remoto es una realidad emergente en Argentina. En cierta forma, parece obvio recomendar que las bibliotecas de nuestro país evalúen sus catálogos informatizados, a la luz de las experiencias obtenidas y publicadas en otros países y teniendo como meta facilitar el acceso público, con el fin de mejorar los sistemas, tanto en lo que se refiere a la descripción bibliográfica y de contenido de las bases de datos, como a las interfaces de búsqueda y los programas de instrucción.

Sin embargo, debe tenerse en cuenta que los sistemas tienen éxito si se adaptan a los usuarios y a sus necesidades, si satisfacen en mayor o menor medida sus expectativas y si proporcionan servicios competentes. De la misma manera, un sistema falla cuando desconoce o no toma en cuenta a quienes pretende servir.

La investigación en Bibliotecología debe redundar en estudios serios que registren el comportamiento de los usuarios y sus necesidades. El relevamiento del estado actual de los servicios y del nivel de participación del usuario en ellos podría ser un primer paso en este sentido.

Bibliografía

- AENOR. Clasificación decimal universal. 7 ed. abreviada española. Madrid: Asociación Española de Normalización, 1995.
- Aitchison J, Gilchrist A, Bawden D. Thesaurus construction. 3 ed. London: Aslib, 1997.
- Ankeny ML. Evaluating end-user services: success or satisfaction? *J Acad Libr* 1991; 16(6):352-6.
- Bates M. How to use controlled vocabularies more effectively in online searching. *Online* 1988; 12(6):45-56.
- Bates M. Rethinking subject cataloging in the online environment. *Library resources and technical services* 1989; 33(4):400-12.
- Beall J. Dewey for Windows. *Advances in knowledge organization* 1996; 5:396-405.
- Beheshti J. The evolving OPAC. *Cataloging and classification quarterly* 1997; 24(1/2): 163-85.
- Bessant L. Users of public online catalogs want sophisticated subject access. *American libraries* 1982; 14(3):160.
- Bishop D. Technical services in health science libraries: an introduction. En: Darling L, Bishop D, Colaianni LA. *Handbook of medical library practice*. 4 ed. Chicago: MLA, 1983:1-26.
- Bloomfield M. A look at subject headings: a plea for standardization. *Cataloging and classification quarterly* 1993; 16(1):119-24.
- BNE. *Autoridades de la Biblioteca Nacional de España en CDROM*. Londres: Chadwyck-Healey, 1998.
- Borgman CL. Why are online catalogs still hard to use. *Journal of the American Society for Information Science* 1996; 47(7):493-503.
- Broadbent E. The online catalog: dictionary, classified, or both? *Cataloging and classification quarterly* 1989; 10(2):105-17.
- Buckland MK. Agenda for online catalog designers. *Information technology and libraries* 1992; 11(2):157-63.
- Buxton AB. Computer searching of UDC numbers. *Journal of documentation* 1990; 46(3):193-127.
- Byrne A, Micco M. Improving OPAC subject access: the ADFA experiment. *College and research libraries* 1988; 49(5):43-6.
- Carlyle A. Matching LCSH and user vocabulary in the library catalog. *Cataloging and classification quarterly* 1989; 10(2):37-63.
- Carpenter M, ed. *National and international bibliographic databases*. Binghamton NY: Haworth Press, 1988.
- CAICYT. FOCAD: Formato común argentino para documentos. 3 ed. Buenos Aires: CAICYT, 1991.
- CENL. COBRA: Computerised bibliographic records actions. sL: CENL, 1997. <http://portico.bl.uk/gabriel/en/projects/cobra.html>.
- Chan LM. Library of Congress classification as an online retrieval tool: potentials and limitations. *Information technology and libraries* 1986; 5(3):181-92.
- Chan LM. A subject cataloging code? *Cataloging and classification quarterly* 1989; 10(2):199-202.
- Chan LM. Inter-indexer consistency in subject cataloging. *Information technology and libraries* 1989; 8(4):349-58.
- Chan LM. Subject analysis tools online: the challenge ahead. *Information technology and libraries* 1990; 9(3):258-62.
- Cherry JM. Improving subject access in OPACs: an exploratory study of conversion of users' queries. *Journal of academic librarianship* 1992; 18(2):95-9.
- Chressanthis JD. The reclassification decision: Dewey or Library of Congress? *Cataloging and classification quarterly* 1995; 19(3/4):169-82.
- Clayton M. *Gestión de automatización de bibliotecas*. Madrid: Fundación Germán Sánchez Ruipérez, 1991 (Biblioteca del libro).
- CNEA. *II Jornadas Nacionales y I Latinoamericanas y del Caribe sobre Microisís*: Buenos Aires, 15 al 23 de junio de 1992. Buenos Aires: CNEA, 1992.
- CNEA. *II Jornadas Nacionales y I Latinoamericanas y del Caribe sobre Microisís*: Buenos Aires, 15 al 23 de junio de 1992. Suplemento. Buenos Aires: CNEA,

- 1992.
- CNEA. III Jornadas Nacionales sobre Microisís: Mar del Plata, 24, 25 y 26 de noviembre de 1993. Mar del Plata: Universidad Nacional de Mar del Plata, 1993.
- Cochrane PA. Modern subject access in the online age. *American libraries* 1984; 15(2):80-3; 15(3):145-8, 150; 15(4):250-2, 254-5; 15(5):336-9; 15(6):438-41; 15(7):527-9.
- Connell TH. Techniques to improve subject retrieval in online catalogs: flexible access to elements in the bibliographic records. *Information technology and libraries* 1991; 10(2):87-98.
- Cousins SA. Enhancing subject access to OPACS: controlled vocabulary vs. natural language. *Journal of documentation* 1992; 48(3):291-309.
- Culkin PB. Rethinking OPACS: the design of assertive information systems. *Information technology and libraries* 1989; 8(2):172-7.
- Del Bigio G. Conferencia. En: II Jornadas Nacionales y I Latinoamericanas y del Caribe sobre Microisís: Buenos Aires, 15 al 23 de junio de 1992. Suplemento. Buenos Aires: CNEA, 1992:49-78
- Drabenstott KM, Demeyer AN, Gerckens J, Poe DT. Analysis of a bibliographic database enhanced with a library classification. *Library resources and technical services* 1990; 34(2):179-98.
- Drabenstott KM. The need for machine-readable authority records for topical subdivisions. *Information technology and libraries* 1992; 11(2):91-104.
- Dubois CPR. Free text vs. controlled vocabulary: a reassessment. *Online review* 1987; 11(4):243-53.
- Duval BK, Main L. Automated library systems: a librarian's guide and teaching manual. Westport: Meckler, 1992 (Computers in Libraries; 64).
- Dwyer J. Bibliographic records enhancement: from the drawing board to the catalog screen. *Cataloging and classification* 1991; 13(3/4):29-52.
- Dykstra M. LC subject headings disguised as a thesaurus. *Library journal* 1988; 113:42-6.
- Dykstra M. PRECIS in the online catalog. *Cataloging and classification quarterly* 1989; 10(2): 81-94.
- Ensor P. User characteristics of keyword searching in an OPAC. *College research libraries* 1991; 53(1):72-80.
- Ensor P. User practices in keyword and boolean searching on the online public access catalog. *Information technology and libraries* 1992; 11(3):210-9.
- Federación Internacional de la Documentación. CDU: Clasificación decimal universal. 6 ed. abreviada española. Madrid: AENOR, 1991.
- Fidel R. Searchers' selection of search keys: II. Controlled vocabulary or free-text searching. *Journal of the American Society for Information Science* 1991a; 42(7):501-14.
- Fidel R. Searchers' selection of search keys: III. Searching styles. *Journal of the American Society for Information Science* 1991b; 42(7):515-27.
- Fidel R. Who needs controlled vocabulary? *Special libraries* 1992; 83(1):1-9.
- Foskett AC. The subject approach to information. 5 ed. London: Library Association Publishing, 1996.
- Foskett DJ. Concerning general and special classification. *International classification* 1991; 18(2):87-91.
- Frost CO. Faculty use of subject searching in card and online catalog. *Information technology and libraries* 1987; 6(1):60-3.
- Gerhan DR. LCSH in vivo: subject searching performance and strategy in the OPAC era. *Journal of academic librarianship* 1989; 15(2):83-9.
- Gödert W. Facet classification in online retrieval. *International classification* 1991; 18(2):98-109.
- Gregor D, Mandel C. Cataloging must change. *Library journal* 1991; 116:42-7.
- Guenther RS. The development and implementation of the USMARC format for classification data. *Information technology and libraries* 1992; 11(2):120-31.
- Hancock M. Subject searching behaviour at the library catalogue and at the shelves: implications for online interactive catalogues. *Journal of documentation* 1987; 43(4):303-21.
- Hancock-Beaulieu M. Evaluating the impact of

- an online library catalogue on subject searching behaviour at the catalogue and at the shelves. *Journal of documentation* 1990; 46(4):318-38.
- Harris K. Part-controlled vocabulary for literature studies. *International classification* 1986; 13:133-6.
- Haverkamp DA. La evolución de la automatización en biblioteca: XX Reunión Nacional de Bibliotecarios, Buenos Aires, 10-13 de setiembre de 1984. Buenos Aires: ABGRA, 1984.
- Henige D. Library of Congress subject headings: is euthanasia the answer? *Cataloging and classification quarterly* 1987; 8(1):7-19 y 21-3.
- Hjorland B. The concept of subject in information science. *Journal of documentation* 1992; 48(2):172-200.
- Holley RP, Killheffer RE. Is there an answer to the subject access crisis? *Cataloging and classification quarterly* 1982; 1(2/3):125-33.
- Holley RP. Subject access in the online catalog. *Cataloging and classification quarterly* 1989; 10(2):3-8.
- Holley RP. Report on the IFLA Satellite Meeting Subject Indexing: Principles and Practices in the 90's, August 17-18, 1993, Lisbon, Portugal. *Cataloging and classification quarterly* 1993; 18(2):87-100.
- International Federation of Documentation. FID Directory 1993-1995. The Hague: FID, 1994.
- ISO 2788: Guidelines for the establishment and development of monolingual thesauri. 2 ed. Geneva: ISO, 1986.
- ISO 10162: Information and documentation: Open systems interconnection: search and retrieve application service definition. Geneva: ISO, 1993.
- ISO 10163-1: Information and documentation: Open systems interconnection: search and retrieve application protocol specification. Part 1: Protocol specification. Geneva: ISO, 1993.
- Jacobs PS. Knowledge-intensive natural language generation. *Artificial intelligence* 1987; 33:325-78.
- Jáudenes Casaubón M. Control de autoridades: normalizar la normalización. En: Megan Walls JA, coord. *Tratado básico de biblioteconomía*. 2 ed. Madrid: Complutense, 1996.
- Jouguelet S. Subject access in online catalogs: developments in France. *Cataloging and classification quarterly* 1989; 10(2):213-24.
- Kaske NK. The variability and intensity over time of subject searching in an online public access catalog. *Information technology and libraries* 1988; 7(3):273-87.
- Kaske NK, Sanders NP. On-line subject access: the human side of the problem. *Reference quarterly* 1980; 20(1):52-8.
- Kern-Simirenko C. OPAC user logs: implications for bibliographic instruction. *Library hi tech* 1983; 1(3):27-35.
- King NS. End-user errors: a content analysis of PaperChase transaction logs. *Bulletin of the Medical Library Association* 1993; 81(4):439-41.
- Krugman S. Failures in subject retrieval. *Cataloging and classification quarterly* 1989; 10(2):9-35.
- Knutson G. Subject enhancement: report on an experiment. *College and research libraries* 1991; 52(1):65-9.
- Koh GS. Options in classification available through modern technology. *Cataloging and classification quarterly* 1995; 19(3/4):195-212.
- Lancaster FW. *Vocabulary control for information retrieval*. 2 ed. Arlington, Virginia: Information Resources Press, 1986.
- Lancaster FW. *Indización y resumen: teoría y práctica*. Buenos Aires: EB Publicaciones, 1995.
- Larson RR. Classification clustering, probabilistic information retrieval and the online catalog. *Library quarterly* 1991a; 61(2):133-73.
- Larson RR. The decline of subject searching: long-term trends and patterns of index use in an online catalog. *Journal of the American Society for Information Science* 1991b; 42(4):197-215.
- Liu S, Svenonius E. DORS: DDC online retrieval system. *Library resources and technical services* 1991; 35(4):359-75.
- Lopes MI. Subject authority control in a union catalogue: perspectives for a national cooperative project in Portugal. *Knowledge organization* 1993; 20(3):133-8.

- Mandel CA. Enriching the library catalog record for subject access. *Library resources and technical services* 1985; 29(1):5-15.
- Mandel CA. Cataloging for access. *Reference Librarian* 1991; (34):61-8.
- Markey K. Searching and browsing the Dewey Decimal Classification in an online catalog. *Cataloging and classification quarterly* 1987; 7(3):37-68.
- Markey K. Integrating the machine-readable LCSH into online catalogs. *Information technology and libraries* 1988; 7(3):299-312.
- Markey K, Demeyer A. Dewey Decimal Classification Online Project: integration of a library schedule and index into the subject searching capabilities of an online catalogue. *International cataloguing* 1985; 14(3):31-4.
- Martínez AM, Mangiaterra EN, Ristuccia C, Pichinini M, Pené M. Control de autoridades en catálogos en línea. *Investigación Bibliotecológica* [en prensa].
- McCarthy C. A reference librarian's view of the online subject catalog. *Cataloging and classification quarterly* 1989; 10(2):203-11.
- McIlwaine IC. Guía para el uso de CDU. Madrid: AENOR, 1994.
- McIlwaine IC. UDC Centenary: the present state and future prospects. *Knowledge organization* 1995; 22(2):64-9.
- McIlwaine IC, Williamson EN. A feasibility study on the restructuring of the Universal Decimal Classification into a fully-faceted classification system. *Advances in knowledge organization* 1994; 4:406-13.
- Menelle CV. Recuperación automática de la información por teleprocesamiento. XX Reunión Nacional de Bibliotecarios, Buenos Aires, 10-13 de septiembre de 1984. Buenos Aires: ABGRA, 1984.
- Micco M. The next generation of online public access catalogs: a new look at subject access using hypermedia. *Cataloging and classification quarterly* 1991; 13(4):103-29.
- Micco M, Popp R. Improving library subject access (ILSA): a theory of clustering based in classification. *Library Hi Tech* 1994; 12(1):55-66.
- Millsap L, Ferl TE. Search patterns of remote users: an analysis of OPAC transaction logs. *Information technology and libraries* 1993; 12(3):321-43.
- Mitchell JS. Options in the Dewey Decimal Classification System: the current perspective. *Cataloging and classification quarterly* 1995; 19(3/4):89-104.
- Mitchell JS. The Dewey Decimal Classification at 120: Edition 21 and beyond. *Advances in knowledge organization* 1996; 5:378-85.
- Mulder C. Z39.50: a step toward a searching standard. *MLA News* 1995; (272):8.
- Niehoff R, Mack G. The vocabulary switching system: description of evaluation studies. *International classification* 1985; 12(1):2-6.
- Olsen RJ, Christensen JO, Larsen KA, Willing K. Implementing NOTIS keyword/boolean searching: a case study. *Cataloging and classification quarterly* 1989; 10(2):125-40.
- Peters TA. When smart people fail: an analysis of the transaction log of an online public access catalog. *Journal of academic librarianship* 1989; 15(5):267-73.
- Piasek JM. Enhanced subject access in Ohio public libraries. *Cataloging and classification quarterly* 1993; 16(4):77-91.
- Pietris MKD. LC: new manual, not code, needed. *American libraries* 1987; 18:958.
- Rada R. Connecting and evaluating thesauri: issues and cases. *International classification* 1987; 14(2):63-9.
- Reynolds SJ. In theory there is no solution: the impediments to a subject cataloging code. *Library quarterly* 1989; 59(3):223-38.
- Ríos García Y. Catálogos en línea de acceso público: selección bibliográfica. *Revista española de documentación científica* 1991; 14(2):121-41.
- Roberts N. The pre-history of the information retrieval thesaurus. *Journal of documentation* 1984; 40(4):271-85.
- Rodríguez Ricard T, Trejo Franqui D. Estudio comparativo de tres listas de encabezamientos de materia en español. *Revista española de documentación científica* 1989; 12(4):422-40.
- Rogers MN. Are we on equal terms yet? Subject headings concerning women in LCSH, 1975-1991. *Library resources and*

- technical services 1993; 37(2):181-96.
- Rolland-Thomas P, Mercure G. Subject access in a bilingual online catalogue. *Cataloging and classification quarterly* 1989; 10(2):141-63.8
- Rolland-Thomas P. Thesaural codes: an appraisal of their use in the Library of Congress subject headings. *Cataloging and classification quarterly* 1993; 16(2):71-91.
- Rowley JE. *Abstracting and indexing*. 2 ed. London: Clive Bingley, 1988.
- Satija MP. FID-88 Pre-Conference Seminar on Computer Applications of the UDC. *International classification* 1988; 15(3):149-50.
- Settel B, Cochrane PA. Augmenting subject descriptions for books in online catalogs. *Database* 1982; 5(4):29-37.
- Shubert SB. Critical views of LCSH ten years later: a bibliographic essay. *Cataloging and classification quarterly*, 1992; 15:37-97.
- Silvester JP, Klingbiel PH. An operational system for subject switching between controlled vocabularies. *Information processing and management* 1993; 29(1):47-59.
- Shoham S, Yitzhaki M. Classification systems and the online catalog. *Advances in knowledge organization* 1994; 4:312-9.
- Sloan BG. Remote access: design implications for the online catalog. *Cataloging and classification quarterly* 1991; 13(3/4):133-40.
- Squires SJ. Access to biomedical information: the unified medical language system. *Library trends* 1993; 42(1):127-51.
- Stephens IE. Getting more out of call numbers: displaying holdings, locations and circulation status. *Cataloging and classification quarterly* 1991; 13(3/4):97-102.
- Strachan PD, Oomes FMH. Universal decimal classification update. *Cataloging and classification quarterly* 1995; 19(3/4):119-32.
- Strong GW, Drott MC. A thesaurus for end-user indexing and retrieval. *Information processing and management* 1986; 22(6):487-92.
- Studwell WE. Why not an AACR for subject headings? *Cataloging and classification quarterly* 1985; 6(1):3-9.
- Studwell WE. The 1990s: decade of subject access. *American libraries* 1987; 18:958,960.
- Studwell WE. Of eggs and baskets: getting more access out of LC subject headings in an online environment. *Cataloging and classification quarterly* 1991; 13(3/4):91-6.
- Studwell WE. LC's head in the sand, or, why the subject cataloging manual is not enough. *Technical services quarterly* 1993; 10(3):45-50.
- Studwell WE, Folsom SL. Game shows, elevators, full plates, and other allegories: a look at the present state and future possibilities of LC subject headings. *Serials librarian* 1993; 23(3/4):241-4.
- Su SF. Dialogue with and OPAC: how visionary was Swanson in 1964? *Library quarterly* 1994; 64(2):130-61.
- Sveinsson E. Unanswered questions in the design of controlled vocabularies. *Journal of the American Society for Information Science* 1986; 37(5):331-40.
- Sweeney R. The development of the Dewey decimal classification. *Journal of documentation* 1983; 39(3):192-205.
- Syracuse RO, Poyer RK. Enhancing access to the library's collections: a view from an academic health center library. *Cataloging and classification quarterly* 1991; 13(3/4):53-90.
- Taylor A, Bohdan S. *Wynar: Introduction to cataloging and classification*. 8 ed. Englewood: Libraries Unlimited, 1992.
- Tedd LA. Computer-based library systems: a review of the last twenty-one years. *Journal of documentation* 1987; 43(2):145-65.
- Tillet B. *Authority control in the online environment*. New York: Haworth Press, 1989.
- Tyckoson DA. The twenty-first century limited: designing catalogs for the next century. *Cataloging and classification quarterly* 1991; 13(3/4): 3-28.
- Thomas AR. Bliss classification update. *Cataloging and classification quarterly* 1995; 19(3/4):105-18.
- Trotter R. Electronic Dewey: the CD-ROM version of the Dewey Decimal Classification. *Cataloging and classification quarterly* 1995; 19(3/

- 4):213-34.
- UBA. Sistema de Bibliotecas y de Información. Bibliotecas universitarias: formato común para registro de información documental. Buenos Aires: UBA, 1993.
- UNAM. Catálogo de autoridades de la Universidad Nacional Autónoma de México. México DF: CUIB, UNAM, 1997.
- Vega E. Macros en la impresión de fichas catalográficas. En: III Jornadas Nacionales sobre Microisís: Mar del Plata, 24, 25 y 26 de noviembre de 1993. Mar del Plata: Universidad Nacional de Mar del Plata, 1993:55-62.
- Vickery B, Vickery A. Online search interface design. *Journal of documentation* 1993; 49(2):103-87.
- Vizine-Goetz D, Markey K. Characteristics of subject heading records in the machine-readable Library of Congress subject headings. *Information technology and libraries* 1989; 8:203-9.
- Walker S. OKAPI: evaluating and enhancing an experimental online catalog. *Library trends* 1987; 35(4):631-45.
- Weinberg BH. Library classification and information retrieval thesauri: comparison and contrast. *Cataloging and classification quarterly* 1995; 19(3/4):23-44.
- Williamson NJ. Classification in the computer age. *International classification* 1989a; 16(1):30-4.
- Williamson NJ. The role of classification in online systems. *Cataloging and classification quarterly* 1989b; 10(2):95-104.
- Williamson NJ. The Library of Congress classification: a content analysis of the schedules in preparation for their conversion into machine readable form. Washington DC: Library of Congress, 1994.
- Williamson NJ. Deriving a thesaurus from a restructured UDC. *Advances in knowledge organization* 1996; 5:370-7.
- Woods EW. Requirements for a format for classification data: Final report. sL: IFLA, 1996. <http://ifla.inist.fr>.

Índice de abreviaturas

- AACR: Anglo American cataloging rules
ABGRA: Asociación de Bibliotecarios Graduados de la República Argentina
AENOR: Asociación Española de Normalización
BBC: Bliss bibliographic classification
BE: Bibliotecas especializadas
BIBUN: Bibliotecas universitarias
BIREME: Centro Latinoamericano y del Caribe de Información en Ciencias de la Salud
BLMCP: Birmingham Libraries Cooperative Mechanisation Project
BNE: Biblioteca Nacional de España
BP: Bibliotecas populares
BSO: Broad System of Ordering
BT: Broader term o término genérico
CAICYT: Centro Argentino de Información Científica y Técnica
CC: Colon classification
CDD: Clasificación decimal de Dewey
CDROM: Compac disk read only memory
CDU: Clasificación decimal universal
CENL: Conference of European National Libraries
CEPAL: Comisión Económica para América Latina
COBRA: Computerised Bibliographic Record Actions
COMPASS: Computer Aided Subject System
CRG: Classification Research Group
CSH: Canadian subject headings
CUIB: Centro Universitario de Investigaciones Bibliotecológicas
ETHICS: Eidgenössischen Technischen Hochschule Information Control System
FAO: Organización de las Naciones Unidas para la Agricultura y la Alimentación
FID: Federación Internacional de Información y Documentación
FOCAD: Formato común argentino para documentos
ICFES: Instituto Colombiano para el Fomento de la Educación Superior
IFLA: International Federation of Library Associations
ISO: International Standard Organization
KWIC: Keyword in context
LAN: Local area network o red de área local
LASER: London and South Eastern Library Region
LCAF: Library of Congress Authority File
LCC: Library of Congress classification
LCSAF: Library of Congress Subject Authority File
LCSH: Library of Congress subject headings
LEMB: Lista de encabezamientos de materia para bibliotecas
LPSH: List of Persian subject headings
MARC: Machine readable cataloging
MeSH: Medical subject headings
NASA: National Air and Space Administration
NT: Narrower term o término específico
OC _C: Online Computer Library Center (antes: Ohio College Library Center)
OEA: Organización de Estados Americanos
OIEA: Organización Internacional de Energía Atómica
ONU: Organización de Naciones Unidas
POPSI: Postulate-based Permuted Subject Indexing
PRECIS: Preserved Context Indexing System
RAMEAU: Répertoire d'Autorité Matière Encyclopédique et Alphabétique Unifié
RLiN: Research Libraries Information Network
RT: Related term o término relacionado
SCOLCAP: Scottish Libraries Cooperative Automation Project
SIDCYT: Sistema de Información y Documentación en Ciencia y Técnica
SISBI-UBA: Sistema de Bibliotecas de la Universidad de Buenos Aires
SWALCAP: South Western Academic Libraries Cooperative Automation Project
UF: Used for o usado por
UNAM: Universidad Nacional Autónoma de México
UNESCO: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura
UTLAS: University of Toronto Library Automation Systems
WLN: Washington Library Network

Índice

Presentación	
Prefacio	
1. Catálogos en línea y búsqueda por materia	
2. Los usuarios y la búsqueda por materia	
- Expectativas del usuario	
- Comportamiento del usuario	
- Problemas que enfrenta el usuario	
- Usuario local, usuario remoto	
- Algunas recomendaciones	
3. Descripción de contenido: adición de términos	
4. Descripción de contenido: combinación de términos	
- Lengua natural y lenguajes documentales	
- Acceso alfabético y acceso sistemático	
- Precoordinación y poscoordinación	
5. Exploración y selección de términos	
- Índice de la base de datos	
- Lenguaje documental en línea	
- Lista de autoridades	
- Lista sistemática	
- Signatura topográfica	
- Exploración y selección de términos para la descripción	
6. Lenguajes documentales	
- Listas de epígrafes	
- Sistemas de clasificación	
- Otros lenguajes documentales	
- Conciliación de lenguajes documentales	
- El futuro de los lenguajes documentales	
7. Catálogos en línea en Argentina: un desafío para esta década	
8. Conclusiones	
Bibliografía	
Índice de abreviaturas	

**Este libro se terminó de imprimir en el
Departamento de Medios Audiovisuales de la
Facultad de Humanidades y Ciencias de la Educación
de la Universidad Nacional de La Plata,
en el mes de julio de 1998.**