

Internet, Metadatos y Acceso a la Información en Bibliotecas y Redes en la era Electrónica

COMPILADORES:

Filiberto Felipe Martínez Arellano

Lina Escalona Ríos

ok

...s@biblioteca.com

data@biblioteca

data@comunic

AGENCY DATABASE

cub

infoconsultores, s.c.

La presente obra está bajo una licencia de:

http://creativecommons.org/licenses/by-nc-sa/3.0/deed.es_MX



Atribución-No Comercial-Licenciamiento Recíproco 3.0 Unported

Eres libre de:



copiar, distribuir y comunicar públicamente la obra



hacer obras derivadas

Bajo las condiciones siguientes:



Atribución — Debes reconocer la autoría de la obra en los términos especificados por el propio autor o licenciante.



No comercial — No puedes utilizar esta obra para fines comerciales.

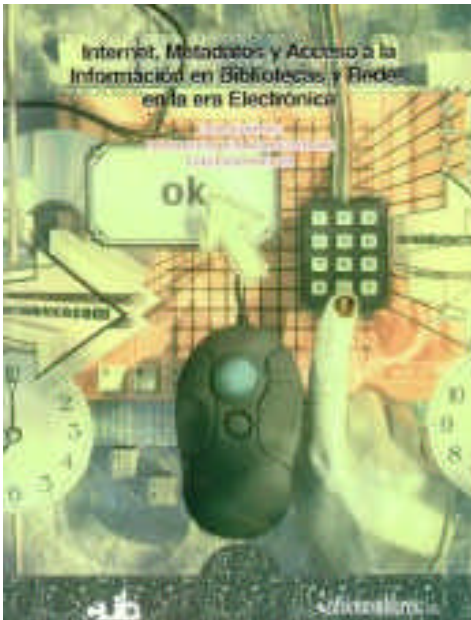


Licenciamiento Recíproco — Si alteras, transformas o creas una obra a partir de esta obra, solo podrás distribuir la obra resultante bajo una licencia igual a ésta.

Esto es un resumen fácilmente legible del:
[texto legal \(de la licencia completa\)](#)

En los casos que sea usada la presente obra, deben respetarse los términos especificados en esta licencia.





Internet, metadatos y acceso a la información en bibliotecas y redes en la era electrónica

Compiladores
Filiberto Felipe Martínez Arellano
Lina Escalona Ríos

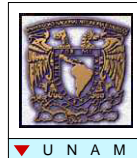
**CENTRO UNIVERSITARIO DE
INVESTIGACIONES BIBLIOTECOLÓGICAS**



▼ CONTENIDO
▼ RESEÑA
▼ SALIR

Internet, metadatos y acceso a la información en bibliotecas y redes en la era electrónica

Primera Edición 2000
DR® UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO



▼ U N A M



▼ C U I B

Créditos de la versión impresa

Coordinación de la edición:
Carlos Ceballos Sosa

Formación Editorial:
Ignacio Rodríguez Sánchez

Revisión especializada:
Francisco González y Ortiz

Diseño de Portada:
Mario Ocampo Chávez

Créditos de la versión electrónica

Diseño y producción:
Carlos Ceballos Sosa

Asesoría en Sistemas:
Alberto Castro Thompson

Edición electrónica 2001
DR® UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO
CENTRO UNIVERSITARIO DE
INVESTIGACIONES BIBLIOTECOLÓGICAS
Producido y hecho en México



Internet, metadatos y acceso a la información en bibliotecas y redes en la era electrónica

Primera Edición 2000
DR® UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO
Ciudad Universitaria, 04510,
México, D.F.
Impreso y hecho en México
ISBN: 968-36-8130-1

RESEÑA

Internet, metadatos y acceso a la información en bibliotecas y redes en la era electrónica, en un tema por demás relevante, en estos momentos en que los bibliotecólogos nos estamos esforzando por adaptarnos a los cambios acelerados que los nuevos formatos de información han traído a nuestra disciplina.

Ciertamente, las nuevas formas de generar, distribuir y consultar información en forma electrónica ha traído consigo grandes retos para la organización de este tipo de información, los que están siendo discutidos a nivel internacional.

Por lo anterior el objetivo de esta obra es difundir las opiniones de destacados bibliotecólogos a nivel internacional que se dieron cita en las instalaciones del CUIB para analizar las características de los sistemas de metadatos usados en la organización de la información electrónica, para discutir la relevancia y pertinencia que tienen los sistemas de clasificación bibliográfica y el acceso temático en la organización y recuperación de recursos electrónicos. Entre los participantes se destaca la presencia de Judith Hopkins, Robin Wendler, Michael Gorman, Gerry Mckirman, Sandy Roe, Ramiro Lafuente, Alvaro Quijano, Margarita Almada, entre otros.

Esperamos que estas ponencias sean del interés de los profesionales de la información que nos dedicamos a la organización y que despierten la inquietud por la investigación sobre estos temas.

Contenido

INTRODUCCIÓN ······	vii
Filiberto Felipe Martínez Arellano Lina Escalona Ríos	
¿Metadatos o catalogación? Un cuestionamiento erróneo? ······	1
Michael Gorman	
USMARC como una estructura de metadatos ······	21
Judith Hopkins	
Diversificación de actividades: habilidades y funciones catalográficas en la era digital ······	36
Robin Wendler	
Puntos de vista: acceso y navegación convencionales y ‘neoconvencionales’ en colecciones digitales ······	49
Gerry McKiernan	
Acceso temático en línea ······	70
Sandy Roe	
Reflexiones en torno a la taxonomía documental para organizar el acceso a la información documental ······	81
Ramiro Lafuente López	
Bibliotecas y tecnologías: una propuesta de integración ······	88
Alvaro Quijano Solís	
Los flujos de información electrónica y la educación del futuro ······	103
Margarita Almada de Ascencio	

Introducción

La época en que vivimos se caracteriza por una serie de cambios en la forma de generar, distribuir y obtener la información. Somos testigos de una transformación de la comunicación impresa hacia una comunicación basada en medios electrónicos y el uso de Internet como alternativa para producir, difundir y obtener información se vuelve un hecho cada día más común. Sin duda alguna, los documentos electrónicos, representan una nueva alternativa para el intercambio de información; sin embargo, al igual que en el mundo de la información impresa, para que los recursos de información electrónica sean conocidos, y por lo tanto utilizados, es necesario organizarlos adecuadamente.

Lo anteriormente expuesto pone de manifiesto uno de los problemas que actualmente enfrenta la Bibliotecología y las Ciencias de la Información, el establecimiento y desarrollo de sistemas que permitan una efectiva organización de los recursos electrónicos para que éstos puedan ser identificados, localizados, obtenidos y utilizados por los miembros de las comunidades a los cuales va dirigida la información que contienen.

Puesto que diversas alternativas, denominadas metadatos, están siendo utilizadas para organizar y recuperar los recursos electrónicos, se vuelve una prioridad para los bibliotecólogos que estamos involucrados en la investigación y práctica de la organización de la información el conocer las características de dichas alternativas.

Tomando en consideración lo anterior, el Centro Universitario de Investigaciones Bibliotecológicas (CUIB) de la Universidad Nacional Autónoma de México (UNAM) seleccionó como tema para su XVI Coloquio Internacional de Investigación el de “Internet, metadatos y acceso a la información en bibliotecas y redes en la era electrónica,” encargándonos la honrosa tarea de organizarlo.

Indudablemente, este tópico es de gran relevancia y trascendencia en nuestros días, en donde los bibliotecólogos estamos adaptándonos a los acelerados cambios que han traído consigo los nuevos formatos de infor-

mación en nuestras actividades profesionales y en nuestra disciplina, los cuales han motivado que nos replantemos nuestro papel como organizadores de la información.

Este tópico ha despertado un gran interés y debate entre la comunidad bibliotecológica internacional puesto que el surgimiento de Internet y los documentos electrónicos, cuyas características difieren grandemente de los impresos, ha puesto en tela de juicio el uso de los sistemas “tradicionales” de catalogación, clasificación e indización en la organización de los recursos electrónicos así como el papel del bibliotecólogo en tal organización.

En este coloquio, investigadores de la Universidad Estatal de California, la Universidad Estatal de Dakota, la Universidad de Harvard, la Universidad de Lund, la Universidad Estatal de Iowa, la Universidad Estatal de Nuevo México, la Universidad Estatal de Nueva York en Búfalo, el Colegio de Bibliotecología de la Facultad de Filosofía y Letras de la UNAM, la Dirección General de Asuntos del Personal Académico de la UNAM, la Dirección General de Bibliotecas de la UNAM, El Colegio de México, la Universidad Autónoma de Chihuahua (Ciudad Juárez), la Universidad de Colima y del Centro Universitario de Investigaciones Bibliotecológicas de la UNAM, nos reunimos para discutir y analizar las tendencias en la organización de los recursos electrónicos, así como las potencialidades que los sistemas de organización bibliográfica “tradicionales”, con los cuales durante muchos años los bibliotecarios hemos organizado colecciones, tienen para organizar los nuevos formatos en que se presenta hoy en día la información.

Considerando que los trabajos presentados por los participantes en el XVI Coloquio de Investigación Bibliotecológica representan un aporte a la investigación bibliotecológica en el campo de la organización de los recursos electrónicos, se decidió compilarlos, traducir los que fueron presentados en inglés y publicar esta obra.

En su documento, Michael Gorman hace un análisis del concepto e importancia de los metadatos, enfatizando las posibilidades del Núcleo de Dublín como una alternativa adecuada para la catalogación de recursos electrónicos, aunque ésta debe ser complementada con el uso de los códigos de catalogación y el formato MARC.

Por su parte, Judith Hopkins analiza la factibilidad de USMARC como un sistema de metadatos para la organización de los recursos electrónicos. Hace mención al proyecto INTERCAT, auspiciado por OCLC, como un ejemplo en donde se ha probado y evaluado la eficacia de utilizar registros bibliográficos de recursos electrónicos en formato USMARC.

En su ponencia “Diversificación de actividades: habilidades y funciones catalográficas en la era digital,” Robin Wendler, acertadamente expone que los bibliotecarios hemos creado sistemas de metadatos durante mucho tiempo y que poseemos los conocimientos y habilidades necesarias para organizar adecuadamente los registros de información electrónica. Asimismo, señala que los bibliotecólogos tenemos ante nosotros el reto que la organización de la información electrónica representa.

Gerry Mckiernan muestra el uso de diferentes sistemas de clasificación, convencionales y neoconvencionales, que se están empleando para la organización y recuperación de la información electrónica. Como ejemplo de los sistemas convencionales menciona a Cyberstacks (sm), una colección de recursos electrónicos organizados con base en el Sistema de Clasificación L.C. y a CyberDewey, en donde se ha utilizado el Sistema de Clasificación de Dewey para la organización de recursos electrónicos. También menciona numerosos ejemplos de casos en donde se han utilizado o experimentado nuevas formas de organización, las cuales denomina “neoconvencionales”.

Sandy Roe aborda en su documento el problema de la falta de aprovechamiento de los vocabularios controlados en la organización y recuperación de la información electrónica. Señala que con la creciente generación de documentos digitales, la búsqueda de información relevante se ha hecho muy lenta, por lo que la búsqueda usando un lenguaje natural, trae consigo serios problemas como la recuperación de grandes cantidades de documentos, los cuales podrían evitarse con el uso de vocabularios controlados.

Alvaro Quijano Solís aborda en su trabajo el impacto que la tecnología ha tenido en las bibliotecas universitarias y por ende en los servicios que se ofrecen a los estudiantes e investigadores. Asimismo, analiza el papel que deberá jugar el bibliotecólogo en la utilización de los recursos tecnológicos para “apoyar las labores bibliotecológicas intelectuales asociadas con los procesos

de valor agregado, que están, a su vez, orientados a lograr la misión y los objetivos de las propias bibliotecas a las que apoyan”.

En su documento, Ramiro Lafuente López señala que ‘las redes de telecomunicación de información están creando nuevos espacios para la comunicación de información, modificando hábitos y estableciendo nuevas formas de documentos digitales que hacen necesario plantearse la conveniencia de establecer principios generales orientados a facilitar la construcción de una taxonomía de los documentos digitales que facilite la construcción de principios generales para crear, organizar, difundir y gestionar acervos digitales.’

Margarita Almada de Ascencio aborda el papel que tendrán que jugar los bibliotecólogos en los programas de educación a distancia y educación virtual, señalando que deberán de participar activamente, de manera interdisciplinaria, en la creación de sistemas de información para apoyar esas modalidades educativas, buscando un adecuado flujo de la información y una adecuada organización de sus contenidos en soportes electrónicos para su interacción a distancia.

Finalmente, es importante mencionar que esta serie de documentos que han sido compilados y traducidos, coadyuvarán a incrementar el conocimiento de las formas que pueden ser utilizadas para organizar en forma adecuada los recursos electrónicos.

Filiberto Felipe Martínez Arellano
Lina Escalona Ríos

¿Metadatos o catalogación?

Un cuestionamiento erróneo

MICHAEL GORMAN

Universidad Estatal de California en Fresno

INTRODUCCIÓN

Las bibliotecas son instituciones que seleccionan, adquieren, proporcionan acceso, ordenan y preservan el conocimiento y la información registrada en todos los formatos existentes, además de auxiliar e instruir sobre el uso de dicho conocimiento e información. Al llevar a cabo esta misión, cada biblioteca posee tres valores inapreciables: un equipo de bibliotecólogos y otros empleados de la biblioteca capacitados y con conocimientos, sus colecciones y la arquitectura de un control bibliográfico que proporciona acceso a éstas. Una biblioteca que carezca de cualquiera de estas tres características no puede ser considerada como tal. Sin control bibliográfico, los bibliotecólogos y cualquier otro tipo de personal no podrían desempeñar su trabajo, tampoco los usuarios emplearían las colecciones de un modo efectivo. Sin los bibliotecólogos y otro tipo de personal, la biblioteca sería solamente un almacén. Sin las colecciones, una biblioteca se reduce al nivel de una oficina de información. Uno debe entender estas tres características esenciales como entrelazadas e interdependientes, cada una tan valiosa y necesaria como la otra. Como dice el viejo acertijo: ¿Cuál es la pata más importante de un taburete de tres patas? Expondré brevemente cada uno de estos factores. Primero...

BIBLIOTECÓLOGOS

En la mayoría de los países, actualmente el término “bibliotecario” se aplica a personas con algún título en bibliotecología (en ocasiones un título superior), aunque a menudo también es empleado para designar a los administradores de alto rango de la biblioteca, independientemente de si poseen o no tal título). En cualquier caso, el trabajo de un bibliotecólogo implica aplicar el conocimiento y la experiencia profesionales a

uno o más componentes de la misión de la biblioteca. Estos componentes son generalmente denominados: desarrollo de colecciones, control bibliográfico, consulta, instrucción bibliotecaria y administración. Cada una de estas grandes tareas pueden aplicarse a la biblioteca como un todo o a colecciones y servicios especializados dentro de ella. La naturaleza del trabajo del bibliotecólogo ha cambiado en las dos últimas décadas. Tal cambio ha obedecido en parte a la tecnología y en parte, a la concentración del profesionalismo y al consecuente aumento en el número e importancia de personal no profesional dentro de las bibliotecas. En suma, contamos con menos bibliotecólogos que antes, pero su trabajo es más profesional. La siguiente característica esencial de una biblioteca son sus...

COLECCIONES

El concepto de colección de la biblioteca también ha cambiado. No hace mucho, la colección de una biblioteca podía definirse como los materiales (casi todos textos impresos en papel) localizados en su edificio o edificios. Hoy debemos asumir una visión más amplia de las colecciones de la biblioteca. La tecnología nos ha traído nuevas formas de registrar el conocimiento y la información, así como permitirnos compartir información acerca de recursos disponibles y, finalmente, compartir tales recursos. La “colección” de una biblioteca moderna consta de cuatro partes:

1. Objetos tangibles (libros, grabaciones, mapas, videos, etc.) propiedad de la biblioteca y resguardados en un lugar físico específico.
2. Recursos (electrónicos) intangibles (CD-ROMS, etc.) existentes en la biblioteca y disponibles para su uso dentro de ésta.
3. Objetos tangibles, (libros, etc.) propiedad de otras bibliotecas pero accesibles a través de sistemas de préstamo interbibliotecario y bases de datos colectivas.
4. Materiales remotos intangibles (electrónicos) que no existen en la biblioteca pero a los cuales la biblioteca tiene acceso.

Cuando tomamos esta visión extensa, podemos ver a la biblioteca como el centro de una serie de amplios círculos concéntricos que en última instancia contienen el conocimiento e información registrada del mundo entero, en tal concepto, la biblioteca propia es a la vez un recurso local y parte de una vasta red de bibliotecas y bases de datos interrelacionadas. Antes de ser transportados por la grandeza de esta metáfora, debemos recordar que ésto es lo que es -una metáfora- y que la grandiosa y universal biblioteca es en realidad, un mosaico de un número casi infinito de programas individuales, servicios, esquemas, etc., todos ellos influenciados por intereses locales. También debemos recordar que los usuarios de la biblioteca prefieren colecciones

locales y bases de datos disponibles localmente porque tienen mayores posibilidades de obtener lo que quieren - acceso rápido a los materiales que necesitan.

Otro aspecto importante de esta visión ampliada de las colecciones de las bibliotecas es la necesidad de hacer una distinción entre el conocimiento registrado y la información. Cuando se usan términos como “la era de la información” y “el profesional de la información” (un sinónimo moderno pero inexacto para el bibliotecólogo), la palabra “información” está casi vacía de significado. Esa falta de significado lleva a la incapacidad de distinguir, en términos de importancia o valor entre, digamos, los poemas de García Lorca y un directorio telefónico. Las palabras tienen significados y tales significados poseen cierto peso, así que es importante definir los términos básicos sobre los cuales descansan todas nuestras especulaciones acerca del futuro.

Es vital que entendamos la distinción entre *información* y *conocimiento registrado*. La información consiste en datos, paquetes individuales de texto e imágenes. La característica más importante de la información es que puede usarse sin un contexto. Uno puede tomar, por ejemplo, el número de personas que viven en América Central, un registro bibliográfico del catálogo de alguna biblioteca, una breve descripción de la geografía de Honduras, como la encontrada en un almanaque o en alguna otra obra de consulta rápida, o una fotografía de la Casa Blanca y emplearlas, sin referencia ningún a otro hecho o información, para añadirlas al almacén de información que uno posee. Por otro lado, el conocimiento registrado es razonado y acumulativo. Avanza de un punto a otro, delineado por el autor o creador del texto o de la secuencia visual. En última instancia, la página de un libro académico, o la fotografía de alguna escena de una película, no tienen significado independiente; su uso depende de lo que le precede y de lo que le sigue. (No me refiero aquí a la escritura creativa o trabajos visuales de arte - aunque éstos también comparten la naturaleza discursiva delineada en el conocimiento registrado y deben leerse o verse como un todo dentro de un contexto.) El punto que estoy tratando de distinguir no es meramente semántico. En realidad, la tecnología electrónica es un bien extraordinario relacionado con la información y no un bien relacionado del todo con el conocimiento registrado. Asimismo, la tecnología electrónica proporciona pocas respuestas a uno de los aspectos más importantes de la bibliotecología, la preservación y transmisión de los registros de la humanidad a las generaciones futuras. Cuando nos referimos a las colecciones, debemos tener en mente que algunos medios de comunicación tienen fortalezas que a otros les faltan y debilidades que necesitan ser subsanadas por otros medios. El futuro global de las bibliotecas depende de entender la distinción entre información y conocimiento registrado ya que cada uno de

los tres principales componentes de los que he hablado, podría cambiarse totalmente si tal distinción y sus consecuencias no son entendidas. Si asumimos, como algunos ya lo han hecho, que el futuro descansa en y sólo en el medio digital para diseminar y preservar los dos aspectos, el conocimiento registrado y la información, los bibliotecólogos llegarían a convertirse en empleados electrónicos, la noción ampliada de las colecciones, de la que he hablado, se reduciría a secuencias de base de datos y el control bibliográfico se convertiría meramente en un asunto de motores de búsqueda, búsquedas por palabras clave y metadatos. Ahora, permítanme hablar acerca del ...

CONTROL BIBLIOGRÁFICO

El control bibliográfico es el elemento vital de la biblioteca. Durante siglos, los bibliotecólogos han creado una arquitectura elegante, además de funcional, basada en la catalogación, clasificación, encabezamientos de materia y en décadas recientes, en sistemas automatizados que han aumentado en gran medida su funcionalidad. La presión hacia la estandarización del control bibliográfico ha sido acelerada por la tecnología. Ahora, por primera vez en la historia de las bibliotecas, estamos en un punto en que el sueño del Control Bibliográfico Universal (*UBC*) puede ser logrado en el mundo real. El mundo anglo parlante es dominado por las *Reglas de Catalogación Angloamericanas, segunda edición (RCA2)* y este código ha tenido mucha influencia en los códigos de catalogación de muchos países cuyo idioma no es el inglés. Todos los países han adoptado la *Descripción Bibliográfica Internacional Normalizada (ISBD)*. Miles de bibliotecas a través del mundo utilizan el *Sistema de Clasificación Decimal de Dewey (CDD)*, la *Clasificación de la Biblioteca del Congreso* o la *Clasificación Decimal Universal (CDU)*. Los países angloparlantes emplean la *Lista de Encabezamiento de Materia de la Biblioteca del Congreso* o alguna otra variante. El sistema MARC para codificar registros bibliográficos electrónicos, está basado en un estándar de la *International Standards Organization (ISO)* y es empleado para intercambiar registros a través del internacionalmente aceptado UNIMARC y casi todos los países, regiones o grupo lingüístico tienen su propia versión de MARC o emplean el formato USMARC para obtener y transmitir registros bibliográficos. MARC es la base para casi todos los catálogos en línea de las bibliotecas y el registro MARC ha hecho posibles las grandes bases de datos actuales, nacionales e internacionales, que son necesarias si queremos lograr un *Control Bibliográfico Universal*. En síntesis, existe un emergente acuerdo internacional sobre el marco y contenido de los registros bibliográficos. Tenemos, o podríamos tener, las herramientas para crear un nuevo orden bibliográfico mundial.

CATALOGACIÓN

La catalogación es una actividad central en el campo del control bibliográfico. Consiste en un número de pasos complejos, el primero de los cuales es describir, de acuerdo con las *ISBD*, la entidad bibliográfica que está siendo catalogada. Dicha entidad podría ser un objeto o un grupo de objetos tangibles (un libro, un microfilm, un conjunto de videocintas, una colección de manuscritos), una serie de datos electrónicos que se encuentra en un objeto tangible que se posee localmente (CD-ROM, cinta electrónica, etc.), o podría ser una fuente remota almacenada en una base de datos distante. Una vez que uno ha descrito la entidad bibliográfica, uno revisa la obra de la cual aquella entidad es una manifestación y asigna puntos de acceso como nombres y título uniforme, encabezamientos de materia y números de clasificación, de acuerdo con las reglas y esquemas utilizados. Finalmente, uno toma todos los datos relacionados con la entidad bibliográfica y con la obra para codificarlos de acuerdo con la versión de MARC usada en la biblioteca, sistema o país al que uno pertenece. Toda esta catalogación se hace tomando en cuenta no sólo las reglas y estándares que ya he mencionado, sino también prácticas antiguas como los registros de los archivos de autoridad (conteniendo en forma estandarizada los nombres y títulos de personas, organismos y obras) y la consulta de manuales sobre la aplicación de encabezamientos de materia y números de clasificación.

Es evidente, con base en la anteriormente mencionado, que la catalogación es una actividad profundamente laboriosa y profesional por naturaleza. Por lo tanto, es un proceso costoso. De hecho, ustedes pueden gastar dinero al principio del proceso para que miles de usuarios de la biblioteca puedan encontrar en el futuro lo que requieren de modo rápido y con un grado aceptable de precisión y pertinencia, o bien, como en el caso de Internet y de la “Web” de nuestros días, el dinero será gastado en el tiempo desperdiciado por miles de usuarios cuando ellos caminen torpemente a través de un caos incontrolado con sólo una ayuda de búsquedas por palabra clave. Simplemente, comparemos la experiencia de localizar conocimientos e información deseados en una biblioteca correctamente ordenada, con el ensordecedor “ruido” producido por la más simple de las búsquedas en la “Web”.

Para ilustrar lo anterior, utilicé la palabra “Honduras” como término temático en el catálogo en línea de mi biblioteca. El resultado fue una lista ordenada alfabéticamente de encabezamientos de materia; cada uno acompañado de una indicación de cuantos materiales sobre cada materia específica existían en el catálogo.

Honduras (4)
Honduras — Administrative and political divisions — Maps (3)
Honduras — Antiquities (7)
[See Also Narrower Term] Cerro Palenque Site (Honduras) (1)
[See Also Narrower Term] Copán Site (Honduras) (10)
Honduras — Armed Forces — Political Activity (1)
Honduras — Armed Forces — Political Activity 20th century (1)
Honduras — Constitutional history (1)
Honduras — Description and travel (5)
Honduras — Discovery and exploration — Maps (1)
Honduras — Economic conditions (3)
Honduras — Economic conditions — 1918 - (3)
Honduras — Economic conditions — 1918 - 1980 (1)
Honduras — Economic conditions — 1918 - Maps (1)
Honduras — Economic policy (1)

A continuación tomé uno de esos encabezamientos, “*Honduras - Antiquities (7)*”, y encontré un breve listado de siete títulos, entre los cuales se encontraban los datos del siguiente:

Pottery of prehistoric Honduras : regional classification and analysis /
Los Angeles : Institute of Archaeology, University of California, Los Angeles,
1993.
312 p. ill.; 28 cm.
Monograph (University of California, Los Angeles. Institute of Archaeology);
35.
Subjects
Indian pottery — Honduras
Indians of Central America—Honduras—Antiquities.
Pottery, Prehistoric— Honduras—Classification.
Honduras—Antiquities.
CSUF STACK F1505.3.P6 P68 1993

[Observemos los encabezamientos de materia que se han incluido, la descripción ISBD, los números de clasificación, cada uno de los cuales requirieron conocimientos de las reglas de catalogación, MARC, etc.]

Después busqué la misma palabra (“*Honduras*”) en AltaVista, uno de los motores de búsqueda más populares. El resultado fueron; 266 970 referencias encontradas, siendo las cuatro referencias que aparecieron en primer lugar (presuntamente “las más relevantes”) las siguientes.

1. 31.07.95.BID Y HONDURAS FIRMAN CONTRATOS POR US\$79 MILLONES

[URL: ww2.iadb.org/prensa/cp15195c.htm]

31.07.95.BID Y HONDURAS FIRMAN CONTRATOS POR US\$79 MILLONES. Samuel Silva: (202) 623-3731 Correo electrónico: samuel@iadb.org CP-151/95 PARA...

2. Dive in Honduras - Main Page

[URL: www.diveinhonduras.com/]

nbsp; Why us? The Resorts. Islands Info. Online Request. Feedback. Mainland Tours. You are Diver number. to dive deeply into our web site Since...

Last modified 6-Feb-98-page size 9K - in English [Translate]

3. A Special Advertising Section on Honduras

[URL: www.washtimes.com/inetnatlads/honduras/coffee.html]

A Special Advertising Section on Honduras. A medium size producer will net about \$800.000 this year. Honduran coffee industry is filling up its cup. The...

Last modified 15-Aug-97-page size 11K - in English [Translate]

4. Central America Business Directory: Guatemala, El Salvador, Honduras, Nicaragua

[URL: www.centramerica.com/]

Business Tourism Exporters Importers Services Banking Directory: Guatemala, El Salvador, Honduras, Nicaragua, Costa Rica

Alta Vista y otras herramientas similares utilizan búsquedas por palabra clave sobre el texto completo de los materiales electrónicos en que ellos buscan. Por definición, estas palabras no son parte de un vocabulario controlado, como aquellos generados por los códigos de catalogación, las listas de encabezamiento de materia y las clasificaciones. El resultado es el “ruido” del que aquí hemos sido testigos. El resultado de una catalogación cuidadosa (y costosa) es que el usuario es conducido muy rápidamente a un material deseado y tener la confianza de que la búsqueda ha obtenido una buena proporción de los materiales relevantes que existen en la biblioteca. Una búsqueda en el gigantesco catálogo colectivo MELVYL, el cual incluye todas las bibliotecas de la Universidad de California, por la materia “*Honduras*”, encontró más registros que aquellos encontrados en el catálogo de mi biblioteca (825). No obstante la cantidad, tales registros fueron de utilidad por estar listados alfabéticamente por autor y porque la búsqueda podía ser limitada fácilmente agregando términos más específicos. Cuando agregué el término “*antiquities*” a la

búsqueda en MELVYL se recuperaron 50 registros y nuevamente, éstos fueron presentados en orden alfabético. Buscando en la “Web” con Alta Vista y agregando la palabra “*antiquities*” a “*Honduras*”, el resultado fue una limitación de los artículos recuperados, se redujeron de más de un cuarto de millón a “únicamente” 86 030, nuevamente en un orden aparentemente aleatorio. El orden fue basado, por supuesto, en un conteo de palabras, una práctica que arrojó como el quinto de los 86 mil artículos “más relevantes” recuperados al siguiente:

5. POPYRUS BOOKS, books on Ancient Art,Antiquities,Archaeology & numismatics [URL: www.papyrusbooks.com/] new, rare and out of print books on the Arts, Archaeology, and Numismatics form Ancient Greece, Rome, Egypt, Near East, Byzantium

Es obvio, desde el punto de vista lógico y desde el de la experiencia, que una catalogación completa de los recursos electrónicos, basada en normas y vocabularios controlados, es infinitamente preferible para recuperar y desplegar registros en lugar de búsquedas basadas en palabras clave dentro de textos completos, vocabularios no controlados y conteo de palabras. Sin embargo, algunos cuestionamientos permanecen:

- * ¿Podríamos afrontar una catalogación completa?
- * ¿Existe alguna tercera forma entre la catalogación completa y la laberíntica búsqueda de palabras clave?
- * ¿Podrían aplicarse los estándares de catalogación existentes a los materiales electrónicos?
- * ¿Qué recursos electrónicos son lo suficientemente útiles para merecer ser catalogados?
- * ¿Cómo podemos preservar los materiales electrónicos una vez que han sido catalogados?

¿QUÉ CATALOGAR? (NO ¿CÓMO CATALOGAR?)

Antes de entrar a la discusión de los metadatos y cómo podrían o no caber dentro del Control Bibliográfico Universal, permítanme considerar la naturaleza de los documentos electrónicos que están disponibles en Internet y en la “Web”. Cualquiera de nosotros que haya dificultosamente transitado por los numerosos resultados de una búsquedas en la “Web” sabe que la gran cantidad de documentos que resultan de esas búsquedas son inútiles, de poco uso, o de un uso que está severa-

mente limitado en el tiempo o en el espacio (i. e., son efímeros o de un fuerte interés local, o ambas cosas). No tengo la manera para determinar el porcentaje de los recursos electrónicos que son de valor duradero y general pero estaría sorprendido si fueran más del uno o dos por ciento. Esto hace surgir la pregunta que involucra una opción con dos alternativas para recuperar los recursos electrónicos -¿catalogación completa para una minúscula minoría y una mejor recuperación, aunque no costosa para la gran mayoría? Existen formas en que la búsqueda por palabras clave puede hacerse más eficiente, aún para el buscador no muy hábil; sin embargo, el punto fundamental es que las palabras clave tienen los siguientes inconvenientes intrínsecos:

- a) Por definición, no son tomadas de un vocabulario controlado.
- b) No tienen contexto. Por ejemplo, una búsqueda sobre “Honduras” encontrará un documento que contenga la frase “este estudio no cubre Honduras”, tan fácilmente como otro que, de hecho, sea acerca de Honduras.
- c) No tienen ninguna relación con otras palabras. Existe un fuerte contraste con las complejas relaciones de una lista de encabezamiento de materia y la posibilidad que estas le proporcionan al usuario para moverse de una materia a otras Relacionadas.
- d) Excepto en las ciencias duras, la terminología está muy lejos de ser precisa y los académicos y escritores emplean palabras con una gran cantidad de matices y falta de exactitud.

El resultado es, por supuesto, que hasta que no tengamos máquinas que verdaderamente puedan pensar y “leer” textos de la misma manera en que lo hace un sofisticado ser humano,¹ no existirá ninguna cantidad de programación que haga posible una búsqueda por palabras clave, la cual sea de utilidad para el usuario como la “tradicional” catalogación bibliotecaria. Por otro lado, si fuésemos capaces de identificar la pequeña minoría de los recursos electrónicos de valor duradero, quizá una búsqueda por palabras clave, ligeramente más desarrollada que la que tenemos ahora, podría ser suficiente para la masa de documentos electrónicos efímeros, de valor local y limitado.

1 Para una discusión interesante sobre qué tan lejos estamos de las “máquinas pensantes,” véase Rawlins, Gregory E.J. *Slaves of the Machine*. MIT Press, 1997.

ESTÁNDARES (ESTRUCTURA Y CONTENIDO)

He mencionado anteriormente los estándares que regulan a la catalogación. Es muy importante para esta exposición que entendamos la diferencia entre los estándares que son *estructura* y aquellos que prescriben el *contenido* del registro bibliográfico. MARC es el mejor ejemplo de un estándar de estructura. Denomina los campos y subcampos de un registro legible por una máquina sin especificar los datos que deban incluirse en esos campos y subcampos. Por ejemplo, el campo 100 es para los asientos principales por nombres personales y el campo 700 para los asientos secundarios de personas (la distinción entre los dos es un vestigio de la catalogación previa a la llegada de la computación, algo que también es incluido, inadecuadamente, dentro de los sistemas de metadatos, pero esa es otra historia...). MARC no le indica al catalogador que nombres deben de ir en el campo 100 y cuáles en el 700, tampoco le indica la forma en que deben ser usados tales nombres. Por lo anterior, uno tiene que recurrir a estándares de contenido como las *RCA2*. Las *ISBDS* son, esencialmente, estándares de estructura; aunque ellas prescriben el contenido en términos generales, uno tiene que ir a su propio código de catalogación nacional para encontrar la forma precisa en que los datos deben de ser ingresados dentro de la estructura *ISBD*. Los esquemas de clasificación y las listas de encabezamiento de materia son, por supuesto, estándares de contenido. Las primeras etapas de metadatos (i. e., de las que estamos siendo testigos en este momento) parecen estar plagadas de problemas debido a una falta de entendimiento por parte de los autores de los metadatos sobre la distinción entre los estándares de estructura y los estándares de contenido, así como de una consecuentemente inhabilidad para decidir si los metadatos están relacionados con el contenido y por lo tanto, que contenido debe aparecer.

¿PUEDEN LOS ESTÁNDARES DE CATALOGACIÓN ACTUALES SER APLICADOS A LOS RECURSOS ELECTRÓNICOS?

Los metadatos, obviamente, están diseñados para responder a necesidades percibidas y desde mi interpretación, parece que esas necesidades son:

- a) Proporcionar una opción para la catalogación de los recursos electrónicos;
- b) proporcionar una tercera alternativa entre los costos de la catalogación “tradicional” y la futilidad de la búsqueda por palabras clave.

Es completamente posible, dado que el concepto original de “metadatos” no surgió entre los bibliotecólogos, que no se haya tomado en cuenta la utilización de la catalogación “tradicional” y aunque ahora los bibliotecólogos están involucrados en

los proyectos de metadatos, la idea de que los recursos electrónicos no pueden ser catalogados empleando los estándares que actualmente existen, podría consolidarse firmemente. Sea como fuere, el hecho es que las entidades bibliográficas electrónicas tienen los mismos atributos que otras entidades bibliográficas. Es perfectamente posible catalogar los materiales electrónicos de tal manera que los registros resultantes pueden ser totalmente integrados en los catálogos de las bibliotecas. Existen unas nuevas *ISBD* para recursos electrónicos,² las cual establecen las bases para la revisión del capítulo 9 de las *RCA2*; los recursos electrónicos poseen títulos y creadores (autores) que pueden emplearse para proporcionar puntos de acceso estandarizados, cubren temas que pueden expresarse en números de clasificación y encabezamientos de materia, además, todos esos datos pueden incorporarse dentro de un registro MARC. En suma, si una de las justificaciones para la invención de los metadatos es que se requieren para facilitar el acceso a los recursos electrónicos por la ausencia de estándares de catalogación, simplemente, tal justificación es errónea.

METADATOS Y MARC

Quizás, la decisión haya sido tomada sin pensar detenidamente en ella. Tal decisión parece indicar que debido a que la “catalogación tradicional” es demasiado costosa, debe haber un compromiso –una tercera forma– que incluya los beneficios de la catalogación pero sin los esfuerzos o gastos que ésta implica. En las palabras de la introducción y las del final del reporte sobre el Proyecto de Metadatos Nórdico³ vemos que:

Muchos especialistas creen que cualquier sistema de metadatos es mejor que no tener ninguno –no necesitamos mantenernos fieles a los rigurosos requerimientos de calidad y formatos complejos de los sistemas de catalogación bibliotecarios–. En su lugar, es posible vivir con lo simple, lo que será fácilmente razonable para los editores, autores y otras personas involucradas en la publicación de los documentos electrónicos. (Las cursivas son mías)

Esta es una de las pocas menciones en este extenso reporte sobre las necesidades percibidas y la naturaleza de los metadatos como una alternativa para la catalogación. Se da por hecho que existe algo entre “los rigurosos requerimientos de calidad” y ninguna calidad, y que también hay algo entre los “formatos complejos” y ningún formato. El cuestionamiento de la catalogación se menciona más adelante

2 *ISBD (ER)* : International Standard Bibliographic Description for Electronic Resources. – München : K.G. Saur, 1997.

3 <http://linnea.helsinki.fi/meta/nmfinal.htm> (Fechado julio de 1998)

en el reporte de una manera más bien extraña. El comité está discutiendo los cambios al “Núcleo de Dublín” que han sido influenciados por el Proyecto de Metadatos Nórdico:

A un nivel más detallado, el Proyecto de Metadatos Nórdico ha tenido una fuerte influencia en cómo los metadatos del Núcleo de Dublín han sido realmente establecidos. Por ejemplo, se decidió que *la forma preferida para los nombres de los autores fuera la normalizada* (“apellido, nombre”). Una razón para esta decisión fue el hecho de que nuestro proyecto apoyaba esa visión fuertemente. Nosotros señalamos en las discusiones que esta forma para el nombre es superior para el ordenamiento de índices y la conversión del Núcleo de Dublín a formatos como MARC, los cuales requiere nombres normalizados. (Las cursivas son mías)

Nótese la premisa acerca de que la única “normalización” que necesitan los nombres para funcionar como mecanismos de recuperación efectiva es la inversión (apellido, nombre), lo cual es normal para la mayoría de los nombres de origen europeo. En efecto, como cualquier catalogador podría decirnos, los nombres no occidentales no son a menudo invertidos y existen muchas variantes en los diferentes tipos de nombres en los que uno necesita las reglas de catalogación para elegir la forma normalizada. Este concepto se repite más tarde en el texto del reporte:

Con el fin de facilitar a los usuarios el elaborar registros con el Núcleo de Dublín, un aspecto esencial son las *ampliamente ignoradas reglas de catalogación*. No existe ningún texto autorizado que nos indique de que parte de la publicación es permitido tomar el nombre del autor, o que defina como poner el nombre en la etiqueta del creador (“apellido, nombre” o “nombre, apellido”). Este tipo de información es actualmente proporcionada a través de guías de usuarios específicas. (Las cursivas son mías)

Nuevamente vemos la idea de que las reglas de catalogación están relacionadas solamente con el nombre del autor y si éste es o no invertido. No se hace ninguna mención de los innumerables problemas de catalogación que surgen en la determinación de los puntos de acceso de autor y título.

EL NÚCLEO DE DUBLÍN

Generalmente, parece ser aceptado que el Núcleo de Dublín es la aplicación más desarrollada de los metadatos y que está a punto de ser ampliamente aceptada. Fue desarrollada por OCLC en sus oficinas centrales en Dublín, Ohio y de este lugar se derivó el nombre. Consiste en 15 elementos descriptivos que son etiquetados. Enumeraré cada uno de estos elementos y mostraré como se relacionan con los estándares y prácticas de catalogación existentes.

Conjunto de Elementos del Núcleo de Dublín

1. *Título*

Etiqueta: “Título”

El nombre dado al material, generalmente por el Creador o Editor.

[El título propiamente dicho como se define por las RCA2. Campo 245 de MARC]

2. *Autor o Creador*

Etiqueta: “Creador”

La persona u organización principal que es responsable de la creación del contenido intelectual del material. Por ejemplo, los autores en el caso de los documentos escritos, los artistas, fotógrafos, o ilustradores en el caso de los materiales visuales.

[El autor, definido casi de la misma forma como se hace en las RCA2, sólo que sin el “contenido artístico.” Campos IXX de MARC]

3. *Materias y Palabras clave*

Etiqueta: “Materias”

El tema del material. Generalmente, las materias son expresadas a través de las palabras clave o frases que describen el tema o contenido del material. Se recomienda el empleo de vocabularios controlados y de esquemas de clasificación formales.

[Materias establecidas por las listas de encabezamientos de materia/tesauros. Campos 6XX de MARC]

4. *Descripción*

Etiqueta: “Descripción”

Una descripción textual del contenido del material, incluyendo resúmenes en el caso de documentos como objetos, o descripciones de contenido en el caso de materiales visuales.

[Notas de contenido o resumen como se definen en las RCA2. Campos 505 o 520 de MARC]

5. *Editor*

Etiqueta: “Editor”

La entidad responsable de que el material esté disponible en su formato actual, tales como una casa editorial, un departamento universitario, o una entidad corporativa. [Editor como se define en las *RCA2*, sub-campo 260 #b de MARC]

6. *Otro colaborador*

Etiqueta: “Colaborador”

Una persona u organización no especificada en el elemento Creador que haya hecho contribuciones intelectuales significativas al material, pero cuya aportación es secundaria para cualquier persona u organización especificada en el elemento Creador (por ejemplo, editor, adaptador e ilustrador).

[Asientos secundarios como se definen en las *RCA2*. Campos 7XX de MARC]

7. *Fecha*

Etiqueta: “Fecha”

Una fecha asociada con la creación o disponibilidad del material. Tal fecha no debe confundirse con la correspondiente al elemento Cobertura, el cual debe de estar asociado con el material solo en los casos que el contenido intelectual se refiera, de alguna manera, a esa fecha.

[Fecha de publicación como se define en las *RCA2*, sub-campo 260 #c]

8. *Tipo de Material*

Etiqueta: “Tipo”

La clase del material, tal como “*homepage*”, novela, poema, documento de trabajo, reporte técnico, ensayo, diccionario.

[Campos de longitud fija de MARC, etiqueta 008, posición 27]

9. *Formato*

Etiqueta: “Formato”

Los datos del formato del material, empleados para identificar la programación (software) y posiblemente el equipo (hardware) requerido para desplegar u operar el material.

[*ISDB(ER)* área 7.5. Campo 516 de MARC]

10. *Identificador de material*

Etiqueta: “Identificador”

Una cadena de signos o número empleado para identificar el material de manera unívoca. Los ejemplos para los materiales existentes en red incluyen los URLs y los URNs (cuando han sido incorporados).

Otros identificadores universalmente unívocos como los números ISBN u otros nombres formales que también son opciones para este elemento.

[Como se define en las *RCA2* y las *ISBD(ER)* 7.5.2 y área 8. Campos 02X del formato MARC]

11. *Fuente*

Etiqueta: “Fuente”

Información acerca de algún material secundario del cual se deriva el material principal. Puesto que generalmente solo son recomendados aquellos elementos que contienen información acerca del material principal, este elemento podrá contener la fecha, creador, formato, identificador u otros metadatos del material secundario considerados importantes para la identificación del material principal.

[Como se define en las RCA2. Campos 76X - 78X del formato MARC]

12. *Idioma*

Etiqueta: “Idioma”

El idioma del contenido intelectual del material.

[Campos de longitud fija de MARC, etiqueta 008, posición 35-37]

13. *Relación*

Etiqueta: “Relación”

Un identificador de un material secundario y su relación con el material principal. Este elemento permite vínculos entre materiales relacionados y las descripciones del material deben de ser indicadas. Los ejemplos incluyen la edición de un trabajo (Versión de), la traducción de un trabajo (Basado en), el capítulo de un libro (Parte de) y una transformación mecánica de una serie de datos en imagen (Formato de).

[Como se define en las RCA2. Campos 76X-78X de MARC]

14. *Cobertura*

Etiqueta: “Cobertura”

Las características espaciales o temporales del contenido intelectual del material. La cobertura espacial se refiere a una región física (por ejemplo, sector celeste); uso de coordenadas (por ejemplo, longitud y latitud), o nombres de lugares que provienen de una lista controlada o escritos en su forma completa. La cobertura temporal se refiere a lo que trata el material, a diferencia de cuando fue creado o puesto a disponibilidad (esto último pertenece al elemento Fecha).

[Campos de longitud fija de MARC, etiqueta 008, posiciones 033 y 043]

15. *Manejo de Derechos*

Etiqueta: “Derechos”

Una mención acerca de la forma de manejar los derechos, un identificador que vincula a una declaración acerca del manejo de los derechos, o un identificador que vincula a algún servicio de obtención de información acerca del manejo de los derechos.

[Campos 506 y 561 del formato MARC]

Es fácil darnos cuenta que lo que tenemos aquí es, en esencia, un subconjunto de MARC. Cada uno de estos elementos tiene su contraparte dentro de MARC y el

contenido de cada uno de ellos está determinado por los códigos de los campos de longitud fija de MARC, los códigos de catalogación de las *ISBD* y/o las listas de encabezamientos de materia/tesauros.

De acuerdo con el documento “web” publicado por el *Survey Institute for Art and Design* en nombre del *ADAM Consortium*,⁴ el Núcleo de Dublín tiene los siguientes atributos positivos:

- * es muy sencillo de aprender
- * tiene elementos repetibles
- * tiene elementos opcionales
- * puede ser ampliado hacia aplicaciones más complejas
- * puede incluirse de modo invisible en las páginas “web”
- * es reconocido dentro del consorcio de la World Wide Web

Todo esto es cierto, pero es escasamente relevante para las preocupaciones básicas acerca de los metadatos, ya que ninguno de estos aspectos se refiere a los puntos centrales del contenido de los registros bibliográficos, o a la limitada naturaleza del subconjunto que esos 15 elementos representan.

El bibliotecólogo australiano Warwick Cathro ha escrito un documento breve e interesante⁵ en el cual examina el debate acerca de la simplicidad del Núcleo de Dublín, o dicho de otra manera, de su estructura y contenido. Al citar a Stuart Weibel de la OCLC⁶ señala,

El punto de vista minimalista refleja un fuerte compromiso hacia la noción de que la principal característica motivacional del ND es su simplicidad. Esta simplicidad es importante para la creación de los metadatos (por ejemplo, por autores no entrenados en las artes de catalogación) como para el uso de los metadatos por herramientas (por ejemplo, generadores de índices que, probablemente, no harán uso de calificadores detallados o esquemas de codificación). El objetivo... solamente puede lograrse si existe un núcleo simple de elementos que sean entendidos para significar la misma cosa en cada caso.

El punto de vista opuesto es ejemplificado en una cita de Roger Clarke,⁷

Las prioridades de los proponentes han sido expresamente orientadas hacia la simplicidad, muy lejos de estructuras sofisticadas. Está implícito en su opción que las dos son incompatibles. El... deseo por la simplicidad ha resultado en un mecanismo que es incapaz de representar la riqueza de los retos del mundo real que se presentan

4 <http://www.adam.ac.uk/adam/metadata.html> (Fechado diciembre de 1997)

5 Cathro, Warwick. The Dublin Core : simplicity or complexity? (National Libraray of Australia. Staff paper) www.nls.gov.au/nla/staffpaper/cathro2.html (Fechado 4 de agosto de 1997)

6 Weibel, Stuart. The 4th. Dublin Core metadata report. *D-Lib magazine*. (Julio 1997)

7 Clarke, Roger. Beyond the Dublin Core, www.anu.edu.au/people/Roger.Clarke/II/DublinCore.html

por ellos mismos... un modelo más rico, más sofisticado, no necesita ser incómodo o inconveniente.

El punto de vista minimalista es que el Núcleo de Dublín deberá contener solamente sus 15 categorías actuales y el contenido no debe normalizarse en la manera que lo hace la gente con habilidades en las “artes de la catalogación” (para usar el término de Weibel). El punto de vista opuesto hace un llamado a más riqueza de contenido e, implícitamente, al ejercicio de aquellas “artes de catalogación” que han mostrado a) una gran y enriquecida extensión y precisión en la recuperación y b) costar más dinero y esfuerzo del que los minimalistas están dispuestos a gastar e invertir. En las palabras de Peter Graham de la Universidad de Rutgers.⁸

El subtexto de la discusión “antiMARC” es un deseo de mantener el Núcleo de Dublín simple. Todo eso está muy bien y es bueno, pero los deseos de los usuarios son complejos... aquellos quienes desean mantener un registro del Núcleo de Dublín “sencillo” ... tienen la obligación de describir cuáles son los límites que el registro del Núcleo de Dublín proporciona. Aquellos (como yo) que desean ver al Núcleo de Dublín posibilitado para expandirse dentro de una capacidad considerable, tienen la obligación de describir cómo puede hacerse esto posible, sin perder el concepto de “núcleo”.

Debemos ser claros en nuestra comprensión acerca lo que aborda esta discusión. Realmente se reduce a una elección entre una forma no costosa y poco efectiva de catalogación en la que los 15 elementos del Núcleo de Dublín se llenan con texto libre, no controlado y sin calificar, por un lado, o una forma de catalogación efectiva y costosa en la cual, por lo menos, algunos elementos del Núcleo de Dublín son llenados con datos controlados y normalizados, decididos sobre las bases de un examen profesional del material. En este último caso, como mínimo, los siguientes elementos del Núcleo de Dublín (ND) necesitarían de intervención humana.

- * Título (ND#1). Decidido de acuerdo con reglas de catalogación (por ejemplo, las RCA2, regla 1.1.B)
- * *Autor o Creador* (ND#2). Decidido de acuerdo con las reglas de catalogación, en términos de la elección del autor/creador y la forma en la cual el nombre es presentado.
- * *Temas y palabras clave* (ND#3). Elegidas de las listas de encabezamientos de materia y tesauros autorizados.

8 En el “4th. Dublín Core workshop”, en Canberra, Australia, marzo de 1997.

- * *Otro colaborador* (ND#6). Decidido de acuerdo con las reglas de catalogación en cuanto a la selección y forma.
- * *Tipo de Material* (ND#8). Una palabra o frase tomada de una lista estandarizada de designaciones.
- * *Fuente* (ND#11). Determinada por el examen del material y presentada en forma estandarizada.
- * *Idioma* (ND#12). Un código tomado de la lista MARC de códigos de idioma.
- * *Relación* (ND#13). Determinada de un examen del material y proporcionada en una forma estandarizada.
- * *Cobertura* (ND#14) Códigos de lugar y tiempo tomados de las listas MARC para códigos de fecha y lugar.

Dicha intervención humana, probablemente no debería tomar grandes cantidades de tiempo y costos como la catalogación completa, pero ciertamente, iría más allá de la simplicidad y los bajos costos deseados por quienes sostienen el punto de vista minimalista.

¿QUÉ DEBEMOS DE HACER?

Este análisis me lleva a creer que tenemos cuatro opciones para el control bibliográfico de los recursos electrónicos: la catalogación completa, registros con un Núcleo de Dublín enriquecido, registros con un Núcleo Dublín mínimo y la opción de búsquedas por palabras clave en texto completo sin estructurar. *Si* estamos preparados para hacer juicios de valor y *si* podemos alcanzar un consenso amplio (ambos “si” muy grandes), me parece que podemos clasificar los recursos electrónicos de acuerdo con su valor, ahora y en el futuro - un proceso no muy diferente a la selección de materiales - y aplicar estos cuatro niveles de control bibliográfico de un modo apropiado. Por ejemplo, aquellos recursos electrónicos relativamente escasos (2%?) y que son vistos como teniendo un valor continuo, podrían ser catalogados completamente de acuerdo con estándares internacionales y nacionales, obteniéndose como resultado registros MARC; el siguiente nivel podría recibir una “catalogación” con el Núcleo de Dublín enriquecido; el subsiguiente nivel, registros mínimos del Núcleo de Dublín; y el último (probablemente la mayoría) podría dejarse a las bondades de los motores de búsqueda. No deseo minimizar la dificultad de alcanzar un acuerdo acerca del valor individual de los materiales electrónicos, pero me parece que el debate sobre la catalogación/Núcleo Dublín y sobre la naturaleza de este último, nunca

podrán resolverse en ausencia de un plan coherente y una aproximación sistemática al control bibliográfico.

PRESERVACIÓN

Es importante darse cuenta que todos los sistemas de control bibliográfico tienen dos funciones primarias. Estas son: a) dar acceso rápido y eficiente a los materiales deseados y b) ayudar a la preservación y transmisión futuras de los registros del conocimiento humano. En el último caso, es obvio que si nosotros no conocemos lo que tenemos, no podríamos preservar esos documentos. Aunque alguna atención ha sido dada a la preservación de los recursos electrónicos, mucha gente ha sido disuadida por la vastedad del problema y la carga extra impuesta por el cambio tecnológico. Creo que nosotros, como bibliotecólogos, tenemos una tarea sin precedentes para asegurar que las generaciones futuras sepan qué es lo que sabemos y como un asunto urgente, debemos de diseñar esquemas para asegurar que esto ocurra.

Las propuestas para desarrollar archivos electrónicos gigantes y sistemas electrónicos de archivo parecen estar relacionadas directamente con cuestiones de tecnología y financiamiento.⁹ ¿Quién va a pagar por tal empresa, especialmente una que va a continuar por siempre? Dado que los registros electrónicos hechos hace sólo diez años, actualmente no pueden usarse por la mayoría, debido al cambio continuo en el hardware y software, ¿cómo podremos asegurarnos que los materiales almacenados en tales archivos puedan ser usados en 10, 50, 100 o 200 años? ¿Vamos a preservar hardware y software obsoleto (una tarea en apariencia imposible durante muchas décadas), o vamos actualizar y “emigrar” todos los documentos y recursos electrónicos cada cierto número de años? Si cada material va a custodiarse solamente en un lugar, ¿qué garantía tenemos de que una copia sobrevivirá indefinidamente? Estas y otras preguntas necesitarán responderse, y responderse pronto, si no vamos a perder para siempre conocimiento registrado e información valiosos.

Casi por accidente, los editores y bibliotecólogos han desarrollado una manera casi infalible de preservar el conocimiento registrado y la información durante los últimos 500 años. Esa respuesta es, por supuesto, imprimir el texto y/o

9 Véase, por ejemplo: Preserving digital information. <http://www.rlg.org/ArchTF/>

imágenes en un papel libre de ácido, hacer copias múltiples y distribuir esas copias de una manera amplia. Es significativo notar la existencia de numerosos libros con más de 400 años de antigüedad que pueden ser perfectamente utilizados. ¿Alguien piensa seriamente que sus equivalentes electrónicos de hoy estarán disponibles y puedan ser usados a fines del siglo 24? Hay quienes aún no consideran ninguna solución que no sea dependiente de la tecnología digital, pero yo les digo lo siguiente a ustedes y a ellos con toda seriedad. La mejor cosa que podríamos hacer, una vez que hayamos seleccionado los materiales electrónicos que vamos a catalogar y/o aquellos a los que debemos hacerles registros con el Núcleo de Dublín enriquecido, es imprimirlos en papel libre de ácido, hacer copias múltiples y distribuir esas copias a bibliotecas de investigación en todo el mundo.

USMARC como una estructura de metadatos

JUDITH HOPKINS

Universidad de Búfalo

CONCEPTOS DE CONTENIDO Y CODIFICACIÓN

Los metadatos, generalmente son definidos como “datos acerca de datos”. Citando una conferencia que dio Robin Wendler al personal de Harvard son: “la información necesaria para identificar, localizar, manejar y acceder un recurso electrónico.”¹

Durante una discusión sobre metadatos en AUTOCAT (la lista de discusión electrónica sobre control de autoridades y catalogación), el día 31 de marzo de 1998, Arlene Taylor,² (profesora de la Escuela de Bibliotecología y Ciencias de la Información de la Universidad de Pittsburgh), introdujo dos conceptos, el de contenido y el de codificación. Muchos escritos sobre metadatos se enfocan al aspecto de contenido (cf. Lynch),³ aunque no en la forma en como éste debería ser tomado, ni como debería ser codificado. Taylor menciona que, con base al uso actual, solamente el contenido que es codificado de acuerdo a algún sistema son metadatos. Citando a Taylor:

La forma en que los metadatos están ahora siendo usados, parece indicar que consideraran ambos aspectos, contenido y codificación. Cuando sólo hay contenido, esto es denominado como un registro bibliográfico o registro representativo. Cuando sólo hay codificación, es identificado como una “estructura”. Así que un registro de contenido

1 <http://sylvia.harvard.edu/~robin/metaslides/sld003.htm>

2 Taylor, Arlene. (1998, Marzo 31) Re: Metadata. AUTOCAD. Disponible por correo electrónico en: LISTSERV@LISTSERV.ACSU.BUFALO.EDU/geetpost-autocat-11038

3 Clifford Lynch da a entender que un recurso de información que se está describiendo no necesita ser digital. En su obra *The Dublin Core Descriptive Metadata Program: Strategic Implications for Libraries and Networked Information Access*, los ejemplos que emplea son los de una pintura, una fotografía de esa pintura y una imagen digitalizada de esa fotografía para las cuales él dice que debe haber tres grupos de metadatos del Núcleo de Dublín, uno para la pintura, otro para la fotografía y uno más para la imagen digitalizada. <http://www.arl.org/newsltr/196/dublin.html>

en RCA2, por sí mismo no podría ser metadatos, aunque un registro de contenido en RCA2 codificado con MARC, sí son metadatos. Existen muchas estructuras de metadatos como: El Núcleo de Dublín, TEI, EAD, GILS, etc. La mayoría de ellas no prescriben el contenido o su forma (aunque TEI sí lo hace). Ellos sugieren qué tipo de información debería de ir en cada campo, pero no proporcionan ninguna guía sobre la forma en que debe de ponerse. Las RCA2 nos proporcionan la guía estandarizada que necesitamos para llenar una estructura MARC. El APPM proporciona a los archivistas una guía para el contenido de un registro EAD (aunque la mayoría de los registros EAD son, hasta ahora, “ayudas de búsqueda” de archivos o manuscritos). Sospecho que a medida que otras comunidades obtengan experiencia, también se verá la necesidad de guías estandarizadas para el contenido.

Por otro lado, Douglas Winship, quién no es bibliotecólogo, sostiene que esta es una definición demasiado limitada. En el siguiente mensaje a AUTOCAT,⁴ él señala:

Como lego, yo pensaría que es aquí donde se rompe el nivel elitista. Si una monografía son datos; y un registro acerca de esa monografía son datos acerca de tales datos, independientemente de la forma del registro, éste es por definición metadatos y los elegantes esquemas de codificación no importan un ápice. Si los metadatos son información acerca de información - perdón por la paráfrasis - entonces, cualquier simple tarjeta de un catálogo de la biblioteca es tan metadatos como el mismo registro en el esquema de codificación más elegante que ustedes pudieran concebir.

¿Qué nivel de codificación se requiere? Una buena tarjeta podría tener todos los datos necesarios acerca del documento en cuestión, aunque no siguiera las RCA o las RCA2, [o] tampoco ser codificado en MARC. ¿Es por ello que no son datos acerca de datos, ni información útil acerca de la información disponible? ¿Deberá de ser ennoblecida por algún esquema computarizado para que en realidad sea digna de valor?

En su respuesta, Taylor⁵ parece estar de acuerdo con Winship en lo general, aunque señala que el uso común tiende a seguir su definición original:

No tengo ninguna objeción personal en llamar metadatos a una tarjeta de catálogo. Simplemente comentaba la forma en que está siendo usada. Por lo que a mí respecta, cualquier información acerca de información son metadatos. Pero en ciertos círculos, ustedes no encontrarán a nadie que admita que los “catalogadores” crean “los metadatos”. Así que en este sentido, sí es elitista, como usted dice.

De esta forma, metadatos son, principalmente, información acerca de recursos electrónicos, en donde esa información es codificada de acuerdo con algún esquema, el que a menudo es electrónico. En forma general, la definición empleada es muy am-

4 Winship, Douglas. (1998, 31 de marzo) Metadata. AUTOCAT. Disponible por correo electrónico en: LISTSERV@LISTSERV.ACSU.BUFFALO.EDU/getpost/autocat/11042

5 Taylor, Arlene. (1998, 31 de marzo) Metadata. AUTOCAT. Disponible por correo electrónico en: LISTSERV@LISTSERV.ACSU.BUFFALO.EDU/getpost/autocat/11051

plia y podría aplicarse a toda la información, esté o no codificada de acuerdo a algún esquema electrónico utilizado para describir cualquier recurso, electrónico o no.

LA ESTRUCTURA Y EL CONTINUUM ESTRUCTURAL

El concepto de “codificación” podría incluir desde el requerir algún esquema electrónico hasta cualquier otra cosa que proporcione una estructura a los datos, facilitando de esa forma su comprensibilidad. Roy Tennant, de hecho define a los metadatos como “información estructurada sobre información” y destaca que la palabra clave es “estructurada.”⁶ Él continúa diciendo que los metadatos, como en la catalogación, (él entonces hace una distinción entre dos conceptos), una descripción de texto libre no es suficiente.

Uno de los temas centrales del 4o. Taller sobre Metadatos del Núcleo de Dublín, llevado a cabo en Canberra, Australia, del 3 al 5 de marzo de 1997, fue la estructura de los elementos o la identificación formal de dicha estructura y los posibles calificadores necesarios para la especificación de cualquier sintaxis dada. En su reporte sobre el Taller de Canberra, Stuart Weibel de OCLC menciona:⁷

Cada uno de los talleres sobre el Núcleo de Dublín celebrados hasta ahora, han tenido una característica particular que refleja algún interés emergente central del grupo. Los resultados del Núcleo de Dublín 4 son mejor conocidos como una manifestación de la tensión entre dos campos pragmáticos: el de los minimalistas y el de los estructuralistas.

Es una sobresimplificación el sugerir que estos dos grupos son distintos y mutuamente excluyentes; de hecho, son dos polos de un continuum y los conferencistas se distribuyeron a través de éste.

El punto de vista minimalista refleja un fuerte compromiso hacia la noción de que la principal característica motivacional del Núcleo de Dublín es su simplicidad. Esta simplicidad es importante tanto para la creación de los metadatos (por ejemplo, por autores no entrenados en las artes de la catalogación), como para el uso de los metadatos por herramientas (por ejemplo, generadores de índices que, probablemente, no harán uso de calificadores detallados o esquemas de codificación). El objetivo de la interoperabilidad semántica entre las comunidades, solamente puede lograrse si hay un núcleo simple de elementos que sea entendido y que signifique la misma cosa en cada caso. Los calificadores adicionales apoyan la especificación, modificación y particularización del significado de un elemento. Dado que esto probablemente será

6 Tennant, Roy. “21st- Century Cataloguing,” *Library Journal*, abril 15, 1998, p. 30.

7 Weibel, Stuart. The 4th Dublin Core Metadata Workshop Report
(<http://www.dlib.org/dlib/june97/metadata/06weibel.html>)

hecho de diferentes maneras por diferentes grupos y en diferentes tiempos, potencialmente conducirá a un flujo semántico en los elementos y a la pérdida consecuente de la interoperabilidad semántica.

Los estructuralistas, como un grupo, aceptan el peligro de este flujo semántico a cambio de una flexibilidad mayor de un medio formal de elementos calificadores o de extensión, de tal forma que puedan ser más útiles para las necesidades de una comunidad particular.

Weibel continua la descripción del continuum de la estructuralidad a la que se ha hecho referencia en las líneas anteriores. Aunque el primer punto de este continuum no son definitivamente los metadatos y el último definitivamente sí lo son, la interpretación queda abierta justo en donde la línea divisoria cae. Las siguientes descripciones de los puntos del continuum están basadas en lo señalado por Weibel, agregando mis comentarios.

Indización de texto completo. - No hay estructura; lo que tenemos es una indización de palabras clave puramente **derivadas** del texto completo de los materiales. La indización puede basarse en el título de una página escaneada o en el texto completo del contenido. Estos no son metadatos.

Representaciones sin campo. - Esto consiste de una colección no diferenciada de términos asignados a un material, sin ningún sistema de etiquetado para identificar tales términos. Aunque ésto aún es contenido puro, es al material por alguien y no derivado como ocurre con las palabras clave.

De acuerdo con mi interpretación, un registro pre-ISBD y pre-MARC de un catálogo de tarjetas, podría caer entre este punto y el que se enumera abajo. Podría considerarse como un ejemplo de un contenido no estrictamente puro, dado que no fue una transcripción exacta del documento fuente pero tuvo alguna estructura implícita. Las reglas convencionales de la construcción de tarjetas del catálogo proporcionaron esa estructura. Dado que ninguno de los datos fue explícitamente etiquetado con su nombre, cada uno de los elementos del grupo de datos pudo determinarse, generalmente, por su colocación en las tarjetas de catálogo.

Representaciones con campos mínimos. - Esto consiste en un número limitado de campos con nombres que identifican los elementos (roles). Esta opción es representada por el campo minimalista en la comunidad del Núcleo de Dublín. Las descripciones ISBD en las tarjetas de catálogo, también podrían caer aquí, pues los elementos son identificados por la puntuación que les precede, además del orden en que son presentados.

Representaciones calificadas. - Los nombres de los campos básicos tienen atributos adicionales que apoyan el refinamiento y la calificación de los nombres de los elementos o de su contenido. Esta alternativa está representada por el campo estructuralista de la comunidad del Núcleo de Dublín.

Representaciones ricamente estructuradas.- Por ejemplo, un registro MARC, un registro TEI.

ESTRUCTURAS DE METADATOS

En su mensaje a AUTOCAT, citado anteriormente, Arlene Taylor se refirió a varias estructuras de metadatos (Núcleo de Dublín, TEI, GILS, MARC). Limitaré esta exposición a MARC, como la opción que tiene mayor relevancia para las bibliotecas, y en particular, para los registros de catálogo, aunque continuaré haciendo referencia al Núcleo de Dublín.

MARC (Machine-Readable Cataloguing)

MARC es una estructura, un contenedor llamado “molde”, en el cual podemos colocar el contenido que hemos creado de acuerdo con estándares externos. En los Estados Unidos, tales estándares son las Reglas de Catalogación Angloamericanas, 2a edición (RCA2),⁸ listas de autoridad de materia tales como los Encabezamientos de Materia de la Biblioteca del Congreso (LCSH),⁹ los Encabezamientos de Materia Médicos (MeSH)¹⁰ o la Lista de Encabezamientos de Materia de Sears,¹¹ esquemas de clasificación como la Clasificación de la Biblioteca del Congreso (LCC),¹² la Clasificación Decimal de Dewey (DCC),¹³ la clasificación de la Biblioteca Nacional de Medicina (NLM),¹⁴ o la clasificación para documentos federales del Superintendent of Documents (SuDocs),¹⁵ etc.

8 2 ed., revisión 1998. Chicago: American Library Association, 1998.

9 20 ed. Washington D.C. : Cataloging Distribution Service. Library of Congress, 1997.

Ahora también está disponible como parte del Cataloguer's Desktop, un producto en CD-ROM del Library of Congress, Cataloging Distribution Service, actualizado trimestralmente.

10 Editada anualmente por la National Library of Medicine; para su venta por el Superintendent of Documents, U.S. Government Printing Office.

11 14 ed. editada por Martha T. Mooney. New York, H. W. Wilson, 1991.

12 Disponible en 47 volúmenes que cubren clases individuales. Muchas de las clases también están disponibles como parte de Classification Plus, un producto en CD-ROM del Library of Congress, Cataloging Distribution Service, actualizado trimestralmente.

13 Dewey decimal classification and relative index. 21 ed., Albany, NY : Forest Press, 1996.

14 5 ed. Bethesda, MD : National Library of Medicine; para su venta por el Superintendent of Documents, U.S. Government Printing Office, 1994.

15 Estados Unidos. Superintendent of Documents. Depository Administration Branch. An explanation of the Superintendent of Documents, classification system. Rev. 1990. Washington D.C. : Depository

MARC no es la única estructura posible, pero es aquella con la cual los bibliotecarios están más familiarizados.

MARC es un término genérico aplicado a un amplio universo de formatos Machine-Readable-Cataloguing [Catalogación Legible por Máquina]. En los Estados Unidos, este término se usa intercambiamente con otro más preciso: USMARC. También existen otras versiones de USMARC (como OCLC MARC y RLIN MARC y aquellas desarrolladas por los diversos vendedores de sistemas integrales de bibliotecas (NOTIS MARC es un ejemplo) que proporcionan designaciones adicionales de contenido: campos, subcampos, valores de los indicadores, etc. Fuera de los Estados Unidos existen otros formatos MARC regionales y nacionales tales como el CANMARC, UKMARC, UNIMARC, aunque actualmente se está llevando a cabo un gran esfuerzo para alinear o “armonizar” estos diferentes formatos.

La tarea de desarrollar un formato legible por máquina para los registros del catálogo se inició en la Biblioteca del Congreso a principios de 1965. La meta era desarrollar un formato que empleara campos de extensión variable en un registro legible por máquina de extensión variable, el cual pudiera emplearse para representar e intercambiar datos bibliográficos entre sistemas automatizados. De esta forma, se diseñó un formato que fuera de COMUNICACIÓN y no solo un formato de procesamiento o almacenamiento. Este se diseñó para los libros, es decir, monografías, pero ha sido ampliado para incluir publicaciones seriadas, materiales cartográficos, partituras y grabaciones, manuscritos, materiales audiovisuales y materiales electrónicos. La versión actual está incluida en la obra *The USMARC format for bibliographic data*,¹⁶ editada por la Biblioteca del Congreso. Existen otros formatos MARC, además del de datos bibliográficos: para datos de autoridad,¹⁷ datos de acervo,¹⁸ datos de clasificación¹⁹ y datos de información para la comunidad.²⁰ Un registro USMARC incluye tres elementos: la estructura del re-

Administration Branch, Library Division, Library Programs Service, U.S. G.P.O., 1990.

16 Ed. 1994. Washington : Cataloguing Distribution Service, Library of Congress, c 1994- . Se han publicado tres actualizaciones, la última en marzo de 1997.

17 USMARC format for authority data. 1993 ed. Cataloging Distribution Service, Library of Congress, 1993- . Se han publicado dos actualizaciones, la última en 1997.

18 USMARC format for holdings data. Ed. 1989. Cataloguing Distribution Service, Library of Congress, 1989 - . Una actualización ha sido publicada en 1995.

19 USMARC for classification data. 1991 ed.. Cataloguing Distribution Service, Library of Congress, 1991 - Se ha publicado una actualización en 1995.

20 USMARC format for community information. 1993 ed. Washington D.C. : Cataloguing Distribution Service, 1993 - . Se ha publicado una actualización en 1996.

gistro, la designación del contenido y el contenido de los datos del registro.²¹ La estructura del registro es una implementación de la norma American National Standard for Bibliographic Information Interchange (ANSIZ39.2), la cual es también un estándar internacional (ISO 2709). Esta estructura tiene tres componentes principales: una guía con una extensión fija, la cual contiene elementos o datos necesarios para procesar el registro; un directorio, el que consiste de una serie de entradas que incluyen la etiqueta, su extensión y la posición de inicio de cada una de ellas dentro del registro; y los campos variables, los cuales conforman la mayor parte del registro.

Aunque la estructura del registro establece los tipos de códigos (por ejemplo, etiquetas de tres caracteres numéricos para identificar cada campo de información, el número de las posiciones de los indicadores que van a ser incluidos en cada campo de extensión variable, etc.), ésta no asigna ninguno de los códigos. La designación de contenido hace eso; define los códigos específicos empleados para identificar las partes específicas de información dentro de un registro, por ejemplo, que 100 es la etiqueta indicando una entrada principal por autor personal, que el subcampo d de ese campo contiene las fechas que pertenecen a la vida de esa persona, etc.

El contenido y la estructura de los elementos de datos en el registro MARC son definidos por estándares ajenos al formato, RCA2, LCSH, Clasificación de la Biblioteca del Congreso. Por ejemplo, las RCA2 nos indica las condiciones bajo las cuales un nombre personal puede ser considerado como una entrada principal (y por lo tanto debe asignarse la etiqueta 100 de MARC) y cómo debe ser formulado aquel nombre, las condiciones bajo las cuales deben agregarse las fechas al nombre personal, que fechas y como éstas deben de ser estructuradas, etc.

¿Qué tan bien funciona USMARC como una estructura de metadatos? Para responder a esta pregunta, necesitamos ver los tipos de metadatos. Roy Tennant,²² un líder de proyectos del Departamento de Investigación y Desarrollo de la Biblioteca Digital de la Universidad de California en Berkeley, describe tres tipos de metadatos aplicables a la información de recursos digitales: descriptivo o intelectual, administrativo y estructural, indicando que USMARC solamente está relacionado con el descriptivo. Esa categoría incluye datos que estamos acostumbrados a ver en los registros bibliográficos (descripción, puntos de acceso, localización de la información, información sobre la manera en que el documen-

21 The USMARC formats: background and principles. Rev. noviembre de 1996.
(<http://lcweb.loc.gov/marc/96principl.html>)

22 Tennant, Roy. "21st-Century Cataloguing", *Library Journal*. abril 15, 1998, p.30.

to descrito está relacionado con otros), los cuales serán empleados para buscar y localizar ese documento.

Los metadatos administrativos describen cómo fue producido un archivo digital y la información sobre su origen (información sobre conversión del documento, autenticidad, etc.), propiedad, derechos de propiedad intelectual, así como términos y condiciones de uso.

Los metadatos estructurales nos indican cómo está estructurado un documento. Tennant emplea el ejemplo del libro impreso que ha sido convertido en un objeto digital para ilustrar esto. En los libros impresos, las páginas se siguen una a otra de modo lineal. No hay, por lo tanto, necesidad de describir la estructura de un libro, con excepción de casos raros. Sin embargo, si el mismo libro se convierte en un objeto digital mediante el escaneo de cada página como una imagen, los metadatos deben proporcionar la información de lo que “encuaderna” esos cientos de archivos de computadora separados para, una vez reunidos, formar un total lógico y proporcionar las formas en que se debe navegar en ese “libro” digital. USMARC no está diseñado para incluir este tipo de información.

No hay necesidad de decir mucho acerca de la conveniencia de USMARC para codificar metadatos descriptivos. Para eso fueron diseñados sus campos de control y de datos. Existen campos de datos diseñados para incluir los nombres de aquellas personas y organismos que crearon el documento o contribuyeron a su creación, diferenciando entre autoría personal y corporativa, así como entre asiento principal y secundario (campos 1XX y 7XX). Proporciona campos para mostrar y dar acceso a varios títulos relacionados con el documento (campos 240, 245, 246, 730, 740). Tienen un campo para la información sobre la edición (250), para la información sobre su publicación (260) y otro para la descripción física (300). La información relacionada es codificada en los campos de serie (4XX y 8XX), en los campos de liga (76X-78X), al igual que en algunos campos de notas (5XX.). Las materias que trata el documento y los géneros que representan se dan en los campos 6XX. El campo de control 041 y el campo para notas 546, proporcionan información acerca de los idiomas del documento. Los campos 05X-09X contienen números de clasificación o información apropiada para la localización física de los documentos, mientras que el campo 856 permite la localización y el acceso de los materiales electrónicos.

USMARC cuenta con opciones para registros de control de archivos, las que pueden expandirse e incorporar elementos de metadatos administrativos. Como actualmente ha sido definido, el código “a” de la posición 08 de la guía para control de archivos indica:

... que el material se describe de acuerdo a las reglas de descripción de archivos, las cuales se enfocan a las relaciones contextuales entre los documentos y a su ori-

gen, más que en los detalles bibliográficos. El grupo específico de las reglas para la descripción puede encontrarse en el campo 040, subcampo #e. Todas las formas de material pueden controlarse a modo de archivo.²³

(El campo 040 corresponde a la Fuente de Catalogación y el subcampo #e a Convenciones de Descripción, excepto si las reglas ALA 1908, 1941 o 1941 han sido empleadas, o si la descripción y puntos de acceso son consistentes con RCA1 o RCA2). No obstante que la definición anterior especifica que todas las formas de material pueden controlarse a modo de archivo, el énfasis del código “a” de la posición 08 de la guía, parece estar en las reglas empleadas para crear el registro.

Los siguientes campos del formato USMARC, probablemente podrían ser empleados para metadatos administrativos, ya sea como actualmente se manejan, o después de alguna redefinición futura:

- 256 Características de los archivos
- 260 Area de publicación, distribución, etc. (impresos) [Incluyo este campo como un ejemplo de metadatos descriptivos; sin embargo, también puede funcionar como un elemento de metadatos administrativos]
- 506 Nota de restricciones de acceso
- 521 Nota de audiencia a quien va dirigido el material
- 524 Nota de preferencia en las citas o descripción de los materiales
- 530 Nota de otros formatos adicionales disponibles
- 534 Nota de la versión original
- 536 Nota de información sobre financiamiento
- 540 Nota de términos que determinan el uso y la reproducción
- 541 Nota de fuente directa de adquisiciones
- 544 Nota de localización de otros materiales de archivo
- 545 Nota de datos biográficos o históricos
- 555 Nota de índice acumulativo/ayudas de búsqueda
- 556 Nota de información sobre la documentación
- 561 Nota de propiedad e historia de la custodia
- 562 Nota de identificación de la copia y versión
- 565 Nota de las características del estuche del archivo
- 567 Nota de metodología
- 581 Nota de publicaciones sobre los materiales descritos
- 583 Nota de acciones sobre el material
- 584 Nota de acumulación y frecuencia de uso
- 786 Asiento de la fuente de datos

23 USMARC format for bibliographic data. Ed. 1994. Washington : Cataloguing Distribution Service, 1994 - . Guía/08.

La mayoría de los campos listados anteriormente son notas y comparten las características generales de todos los campos de notas, es decir, son de texto libre, muchos de ellos son altamente estructurados y contienen subcampos altamente específicos, por ejemplo, los campos 506 y 534. Mientras más estructura exista para un campo, más fácil será de indizar.

La debilidad más grande de USMARC como una estructura de metadatos es su inhabilidad para manejar metadatos estructurales, información sobre cómo compilar el objeto a partir de sus componentes lógicos y cómo facilitar la navegación del objeto. Lo más cercano que tiene USMARC y que podrían usarse para este propósito son varios campos establecidos como puntos de accesos para relaciones. Estos incluyen:

Posición 19 de la Guía. Requerimiento de un registro relacionado. Esta posición de un carácter incluye un código que indica si una nota conteniendo información de identificación básica puede ser generados de los datos en los campos 760-787 y/o en el campo 580 sin tener acceso al registro relacionado.

760-787 Campos de asientos relacionados. Estos campos contienen información que permiten identificar otros registros bibliográficos. Cada uno de estos campos describe una relación diferente entre el documento de origen, descrito en el registro, y el documento relacionado.

525 Nota suplementaria.

580 Nota de complejidad entre asientos relacionados.

Estos campos no son apropiados, tal y como actualmente existen, para los metadatos estructurales, dado que relacionan registros BIBLIOGRAFICOS, no objetos reales que podrían no estar representados por registros bibliográficos separados, por ejemplo, imágenes de páginas individuales de un libro o un artículo de revista.

En resumen, ¿cuáles son las VENTAJAS de usar MARC como una estructura de metadatos?

- * MARC es un formato bien constituido (más de 30 años) con procedimientos aceptables para actualizarse.
- * Una gran inversión en software y experiencia han sido hechas en MARC.
- * Los registros MARC tienen una gran coherencia organizacional.
- * MARC proporciona especificaciones detalladas para los elementos (aunque no todas las que podrían desearse), lo cual permite una manipulación precisa de los datos.
- * El uso de MARC permite que los registros de metadatos sean incorporados dentro de los catálogos de biblioteca.

¿Cuáles son las DESVENTAJAS de MARC como una estructura de metadatos?

- * La construcción y mantenimiento de los registros MARC son costosos

- * El sistema MARC es complejo y requiere creadores de registros perfectamente entrenados y un software complejo.
- * El diseño original de MARC estuvo basado en las tarjetas de catálogo, enfocado al libro impreso y orientado hacia materiales impresos no continuos.
- * MARC no es apropiado para la inclusión de metadatos estructurales y administrativos.

La Biblioteca del Congreso ha previsto formas de conversión o “mapeo” entre USMARC y otras estructuras de metadatos.²⁴ La alternativa “Núcleo de Dublín / MARC / GILS” podría emplearse para convertir metadatos con sintaxis del Núcleo de Dublín a MARC. Una cita sobre esta forma de conversión indica.²⁵

Un “mapeo” entre los elementos del Núcleo de Dublín y los campos de USMARC es necesario para que las conversiones entre diferentes sintaxis puedan realizarse con exactitud. Una vez que el estilo de los metadatos del Núcleo de Dublín sea ampliamente establecido, podría interactuar con los registros MARC en varias formas, como las que se indican a continuación:

Enriquecimiento de registros con una descripción sencilla de recursos. Una agencia de catalogación podría desear tomar los metadatos proporcionados en el estilo del Núcleo de Dublín (presumiblemente en HTML o SGML) y convertir esos datos en campos MARC, resultando un registro estructural. Entonces, ese registro podría ampliarse, tanto como fuese necesario, agregando información adicional, generalmente incluida en el catálogo particular.

Aunque la cita anterior menciona que un registro MARC, basado en los elementos del Núcleo de Dublín, “podría” ser ampliado para proporcionar información adicional, en realidad, tales registros seguramente necesitarían ser ampliados (véase el segundo párrafo de la cita). Nótese que el propósito mencionado para el “mapeo” entre los elementos del Núcleo de Dublín y los campos USMARC es asegurar que “las conversiones entre diferentes sintaxis pueden realizarse con exactitud”. La sintaxis y no la semántica. Uno podría tener una conversión perfecta de un registro del Núcleo Dublín a un registro USMARC, de tal forma que los datos incluidos en el campo del creador en el Núcleo de Dublín vayan al campo apropiado 1XX de USMARC, etc., pero el resultado no será un registro RCA2/MARC, sin un manejo adicional de los datos para asegurar que éstos cubran los estándares de RCA2, LCSH, etc.

Debe de hacerse énfasis en que el Núcleo de Dublín no es catalogación; no incluye reglas de descripción, guías o prácticas comunes. Tomando como ejemplo los elementos incluidos en el área del creador o contribuyente de un registro en el

²⁴ <http://lcweb.loc.gov/marc/>

²⁵ <http://lcweb.loc.gov/marc/dccross.html>

Núcleo de Dublín, no existe ninguna regla para especificar la forma en la cual los nombres personales o de organismos deban de ser asentados. (¿Debe asentarse un nombre personal en forma directo, con el nombre de pila precediendo al apellido, o en un orden indirecto, incluyendo el apellido, coma, nombre de pila? ¿Cuál deberá de ser el elemento de entrada para un apellido compuesto o un apellido con prefijos? ¿Qué tan completo debe ser el nombre de pila?, etc.) Todas estas son cuestiones cubiertas en RCA2 y en otras reglas de catalogación. El resultado podría ser que dos objetos, a los cuales se les hayan asignado elementos con el Núcleo Dublín, podrían tener campos para el creador donde este presente el nombre del mismo individuo de diferentes maneras; las reglas de catalogación están concebidas para prevenir este tipo de inconsistencia.

No obstante lo anterior, en el 4o. Taller del Núcleo Dublín (llevado a cabo en Canberra, en marzo de 1997) se definió un modelo para la calificación de los datos. Los calificadores refinan los elementos o datos. Existen tres tipos de calificadores: TYPE (que modifica el nombre del elemento), SCHEME y LANG (los cuales modifican el contenido del elemento). De esta forma, ahora es posible para un elemento del Núcleo de Dublín tener un calificador, especificando que tal elemento fue creado de acuerdo a un grupo particular de reglas, que proviene de una lista de autoridad de términos específica, o bien, que se encuentra en determinado idioma. Ejemplos:

```
ELEMENT= contributor TYPE =name SCHEME= AACR2  
CONTENT=Knutson, Roger M., 1993-  
ELEMENT=subject SCHEME=lsh LANG=en-us CONTENT=Cookery, Mexican
```

Los registros MARC basados en registros del Núcleo de Dublín, conteniendo elementos que han sido calificados para mostrar que han sido estructurados de acuerdo a AACR2, o tomados de LCSH o alguna otra lista estándar de autoridades de materia para bibliotecas, necesitarían muy poca o ninguna ampliación. La principal ampliación que podría ser necesaria en el registro sería la adición de campos para proporcionar información no incluida en algún otro elemento del Núcleo de Dublín.

CATALOGACIÓN DE RECURSOS ELECTRÓNICOS: EL PROYECTO INTERCAT

Quizás, el proyecto más conocido para crear registros catalográficos para recursos electrónicos, empleando USMARC, es el Proyecto Intercat, el cual es aus-

piciado por OCLC. Establecido en 1995 con fondos provenientes del Departamento de Educación de los Estados Unidos, el proyecto solicitó la ayuda de voluntarios de bibliotecas (no tenían que ser miembros de la OCLC) con la finalidad de crear registros de catálogo para recursos electrónicos de su elección.

El objetivo de este proyecto fue probar y evaluar la eficacia de utilizar registros bibliográficos en formato USMARC, incluyendo localización electrónica e información sobre su acceso (campo 856 de USMARC), para proporcionar la descripción, la localización e información sobre el acceso de objetos de información electrónica disponibles en forma remota.

Total de registros en InterCat hasta el 31 de marzo de 1998: 29552

Total de registros agregados en marzo de 1998: 2846

Máximo de registros agregados en un sólo mes: 2846 (marzo de 1998)

Máximo anterior en un sólo mes: 2101 (enero de 1998)

Total de registros agregados en el primer trimestre de 1998: 6261²⁶

Total de registros en InterCat hasta el 15 de julio de 1998: 40192

Máximo de registros agregados en un sólo mes: 3,912 (mayo de 1998)

Conjuntamente con este proyecto, la reconocida catalogadora estadounidense de materiales audiovisuales, Nancy B. Olson, escribió el manual *Cataloguing Internet Resources: a manual and practical guide*, cuya segunda edición está disponible, tanto en formato impreso como en documento web.²⁷ El manual es compatible con las normas ISBD(ER) (*International Standard Bibliographic Description for Electronic Resources*), 2a. ed., borrador 1995. Todos los ejemplos del manual incluyen la correspondiente codificación en USMARC. El capítulo 8 consiste de ejemplos de registros USMARC existentes en OCLC WorldCat para once tipos de publicaciones electrónicas diferentes, mientras que un apéndice contiene una descripción y los lineamientos para el uso del campo 856 de USMARC: Acceso y localización electrónica.

En la introducción de esta edición del manual, Olson señala:

Las premisa básicas de los proyectos OCLC relacionados con el acceso a recursos de Internet son:

1. Existe una gran cantidad de información valiosa disponible a través de Internet.
2. Estos recursos necesitan estar organizadas para tener acceso a ellos.

26 Jul, Erik. Posting 104 (1 de abril de 1998) en intrcat@oclc.org. Disponible por correo electrónico en: [LISTSERVE@OCLC.ORG/getpost intercat 104](mailto:LISTSERVE@OCLC.ORG/getpost%20intercat%20104); actualizado en julio 16 de 1998. Disponible por correo electrónico en: [LISTSERVE@OCLC.ORG/getpost intercat 169](mailto:LISTSERVE@OCLC.ORG/getpost%20intercat%20169)

27 2 ed. Dublín, Ohio: OCLC, 1997. Disponible en el sitio web de OCLC en : <http://www.purl.org/oclc/cataloguing-internet>.

3. La utilización de las técnicas y procedimientos empleados en bibliotecas y la creación de registros para su recuperación a través de catálogos en línea actuales es el método más eficiente para poder tener acceso a estos recursos.²⁸

Un campo especial de MARC, la etiqueta 856,²⁹ Acceso y localización electrónica, fue desarrollado especialmente para este tipo de información.

En otra parte del libro, Olson menciona: “El capítulo 9 de las RCA2 incluye especificaciones especiales para catalogar archivos de computadora disponibles a través ‘acceso remoto’. El acceso remoto es definido en las RCA2 como: ‘El uso de archivos de computadora a través de dispositivos de entrada/salida conectados electrónicamente a una computadora’ Lo anterior muestra diferencias con el acceso directo: ‘El uso de archivos de computadora vía dispositivos (e. g., discos, cassettes, cartuchos) diseñados para ser insertados en una computadora o en su equipo auxiliar por parte del usuario’³⁰

Olson continua mencionando, “Las especificaciones especiales del capítulo 9 de las RCA2, para archivos disponibles por acceso remoto, son aplicables a los recursos de Internet. Estas incluyen el uso del Área 3 para la descripción del archivo, la ausencia de la descripción física y una nota que indica el ‘modo de acceso’ del archivo.”³¹

Aunque el libro de Olson es el único manual para la catalogación de recursos de Internet, ella no es la única persona que ha pensado y escrito acerca de este tema. Ingrid Hsieh-Yee de la Escuela de Bibliotecología y Ciencias de la Información de la Universidad Católica de América, ha propuesto lo que ella denomina “un estándar modificado para la catalogación de nivel mínimo.”³² Ella difiere de Olson sobre un número de puntos a los que no me referiré en este trabajo.

Ella propone que los registros creados en OCLC deberán de ser colocados en Internet para competir con otros motores de búsqueda.

28 Olson. p. 2.

29 <http://www.loc.gov/marc/856guide.html>

30 Olson, p.5.

31 Olson, p.5

32 xxxii <http://www.oclc.org/oclc/man/colloq/hsieh.htm>.

CONCLUSIÓN

¿Cuál es el futuro de MARC y de todos los MARCs? MARC está ligado al futuro de las bibliotecas y al de los catálogos de acceso público en línea (OPACs) y existe una creencia muy extendida en muchas comunidades de que los OPACs ya no podrá ser el centro del mundo de la información. Actualmente, las búsquedas hechas en los OPACs son sólo una pequeña proporción de las búsquedas efectuadas con motores de búsqueda. MARC y AACR2 no serán adoptadas por el resto del mundo. Pero, ¿existe alguna razón para abandonar MARC? Yo no lo creo. Aparte del hecho de que la mayoría de las bibliotecas tienen una gran cantidad de inversión en MARC, éste posee una trayectoria comprobada para proporcionar registros descriptivos altamente estructurados y manejables para materiales que alguien haya seleccionado como apropiados para la colección de alguna biblioteca. Durante algunos años más, las bibliotecas continuarán usando MARC (conjuntamente con los diversos estándares de contenido tales como las RCA2) para proporcionar registros descriptivos de alta calidad para los materiales seleccionados. MARC continuará su desarrollo en respuesta a necesidades nuevas (por ejemplo, la necesidad de representar datos estructurales y administrativos), o bien, será complementado con métodos alternativos para representar tales datos; en cualquier caso, MARC seguirá siendo usado por las bibliotecas.

Diversificación de actividades: habilidades y funciones catalográficas en la era digital

ROBIN WENDLER

Oficina de Sistemas de Información,
Biblioteca de la Universidad de Harvard
rwendler@harvard.edu

Metadatos es la palabra de moda en labios de todos en nuestra sociedad fuertemente influenciada por la información y como bibliotecarios debemos de estar emocionados. Después de todo, siempre hemos creado metadatos para controlar y proporcionar acceso a nuestras colecciones. Mi propia definición de metadatos dentro del contexto bibliotecario es una muy amplia: la información necesita ser identificada, localizada, controlada y a través del acceso de los materiales, la biblioteca desea hacerla disponible a sus usuarios. Desde esta perspectiva, la catalogación son metadatos, aunque no todos los metadatos son catalogación. Nuestros principios de la catalogación han sido desarrollados y refinados por décadas para apoyar, de la mejor manera, la misión de la biblioteca. Sin embargo, dos desarrollos tecnológicos tienen grandes implicaciones para la función de la catalogación en nuestras instituciones: 1) la explosión de la publicación electrónica, 2) la conversión de metadatos no previamente disponibles en línea hacia una forma legible por máquina.

La misión institucional de las bibliotecas de investigación no incluye el control de Internet. En su lugar, las bibliotecas de investigación enfrentan la necesidad de desarrollar una infraestructura sólida para el control y acceso de los recursos digitales creados, comprados, o seleccionados por nuestras instituciones, algunos de los cuales son recursos de Internet. Debemos:

- * facilitar el acceso a las colecciones digitales
- * integrar las colecciones digitales y las colecciones tradicionales
- * revalorar los estándares y prácticas de catalogación, considerando nuevas formas de publicación
- * crear un ambiente de información coherente que reúna la catalogación heterogénea y los metadatos generados por una gran diversidad de bibliotecas, archivos y museos.

PUBLICACIONES ELECTRÓNICAS

La explosión de las publicaciones digitales tiene el potencial para revolucionar la bibliotecología, no sin antes requerir una reexaminación, rediseño y ampliación de los procesos y estándares existentes. El rápido incremento en el número y evolución de las formas de publicación digital no se apega a la taxonomía que hemos aceptado de los tipos de publicaciones, las cuales requieren actividades de administración de colecciones, no aplicables a las publicaciones en otros medios, así como sistemas de acceso innovador para maximizar su valor y uso.

CONTROL INTELECTUAL DE LOS RECURSOS ELECTRÓNICOS

Algunas características de los recursos electrónicos hacen que éstos sean difíciles de describir:

Variabilidad en la presentación. Las configuraciones de los navegadores locales o el estilo de las páginas pueden determinar cómo aparece un documento. En algunos casos, la apariencia, y aún el orden del contenido, son generados de manera aleatoria. Aún no sabemos cómo crear una catalogación con sentido y que pueda ser compartida acerca de objetos como éstos.

Formas no familiares de publicación. La comunidad bibliotecaria nunca ha entendido muy bien qué hacer con las publicaciones de hojas sueltas, y en algún sentido, ésta es la analogía más cercana de muchas publicaciones electrónicas. El nuevo contenido está impreso en intervalos los cuales podrían aumentar o reemplazar a algunos o todos los contenidos previos, frecuentemente, sin una numeración o cualquier otra indicación de que la versión o su estatus ha cambiado.

Relaciones mal definidas de los materiales con otras formas. Un subconjunto de las publicaciones electrónicas se deriva, o está, de alguna forma, relacionado con publicaciones en otros medios. No siempre es posible determinar el grado de traslape en el contenido de estas variantes o el significado de las diferencias funcionales entre ellas y no existe ningún consenso dentro de la comunidad acerca de como deberán expresarse estas relaciones en la catalogación.

Mutabilidad. Excepto en los casos donde los materiales son archivados localmente en el lugar que se efectúa la descripción, no hay garantía de que éstos no serán modificados radicalmente sin alguna advertencia, volviendo su catalogación inexacta y posiblemente inútil.

Movilidad. En tanto que no se implementen ampliamente algunas medidas como los Nombres Uniformes de los Recursos (*Uniform Resource Names, URNs*) y los sitios se reorganicen, las ligas seguirán modificándose y se perderá el acceso.

Ausencia de un receptor físico. La falta de un usuario convencional elimina el punto clave sobre el cual la mayoría de las bibliotecas han construido sus procedimientos de mantenimiento del catálogo.

Estas son algunas razones del por que las nuevas formas de publicación alteran nuestros modelos establecidos para un control intelectual. Ellas continúan evolucionando en formas que no podemos anticipar. Las formas básicas de las publicaciones electrónicas aún no son estables, no las entendemos, y por lo tanto, no sabemos cómo describirlas y controlarlas. Los conceptos establecidos en las reglas de catalogación, tales como la fuente de información principal, están volviéndose algo sin sentido en un mundo en donde la apariencia visual, y aún el contenido, pueden variar de un usuario a otro. Las categorías de formatos que hemos empleado previamente, tales como monografías, series, textos y audiovisuales, ya no describen adecuadamente estas publicaciones. Se requiere de un trabajo intelectual serio para adaptar los modelos abstractos y las reglas, prácticas, códigos y procedimientos de la bibliotecología a las características de las publicaciones electrónicas e, indudablemente, esto es trabajo de los catalogadores.

LA ADMINISTRACIÓN DE LOS RECURSOS ELECTRÓNICOS

Además de los problemas de descripción que traen consigo las publicaciones electrónicas, éstas requieren de una administración más activa que cualquier otro medio. Algunas o todas las funciones siguientes podrían ser necesarias para manejar las colecciones electrónicas y cada una de ellas requiere de la creación y mantenimiento de nuevos tipos de metadatos:

- * administración de contratos (registro de acuerdos de compras en consorcio, condiciones de licencia, etc.)
- * restricción de acceso y administración de derechos, lo que requiere tanto de instrucciones textuales para los usuarios como de datos estructurados para el procesamiento automatizado (tales como el número de usuarios simultáneos, lista de rangos IP autorizados, números de identificación de acceso y contraseñas, guiones (*scripts*) de acceso local.
- * apoyo técnico, lo cual requiere información acerca de contactos en caso de falla, clientes propietarios, compatibilidad de navegadores , etc.

Para el contenido electrónico que la propia biblioteca crea, tales como herramientas de indización y resúmenes, o los productos de la conversión digital, se requieren todavía de más metadatos:

- * archivación y migración (programas de actualización; especificaciones técnicas tales como fecha de escaneo, perfil de color, resolución, fuente de luz, formato del archivo, profundidad de bit, dimensiones de píxeles, etc.; historia de los cambios)
- * generación de las interfaces de presentación y navegación las que requieren de metadatos estructurales que documenten la organización original del objeto digital para apoyar funciones como el cambio de página, etc.

Este nivel de administración de los recursos es un nuevo trabajo que necesita de un lugar dentro de la organización de la biblioteca. Algunos aspectos de éste recaerán en el personal de adquisiciones, servicios al público y sistemas, aunque algunos otros pueden pertenecer perfectamente al departamento de catalogación.

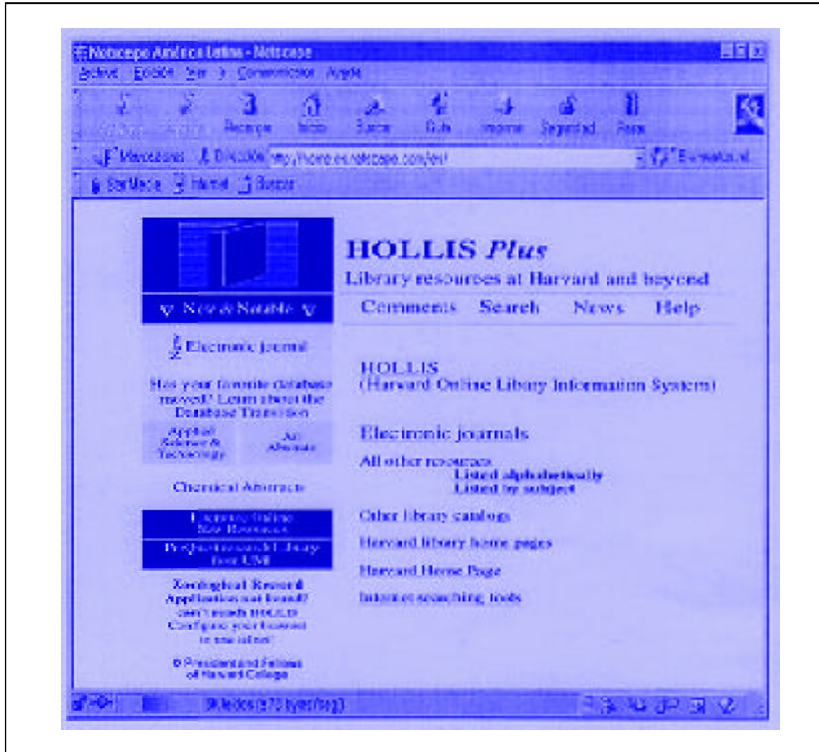
ACCESO A LOS RECURSOS ELECTRÓNICOS

Para proporcionar acceso a los recursos electrónicos, muchas instituciones de investigación han desarrollado, de una manera muy sorprendente, modelos similares: los catalogadores crean registros más o menos tradicionales en el catálogo en línea, mientras los bibliotecarios a cargo de la selección o de los servicios al público crean, en forma independiente, una estructura de acceso basada en el “Web”, en donde a través de un menú destacan los recursos electrónicos más populares, proporcionando información completa acerca de los materiales seleccionados (y a menudo costosos), en la forma que podría hacerse en un registro del catálogo diseñado para tal fin. HOLLIS Plus de Harvard es una de tales interfases y la Página de Recursos Electrónicos de la Universidad Estatal de Michigan es otro (*Ejemplos 1 y 2*).¹

Generalmente, tales sitios “web” incluyen una división de primer nivel (por ejemplo, revistas electrónicas y materiales de consulta). Cada segmento es accesible a través de listas alfabéticas, listas de materia y/o búsquedas por palabras clave (*Ejemplo 3*). Cada uno de los materiales electrónicos tiene una página de acceso que cubre varias funciones:

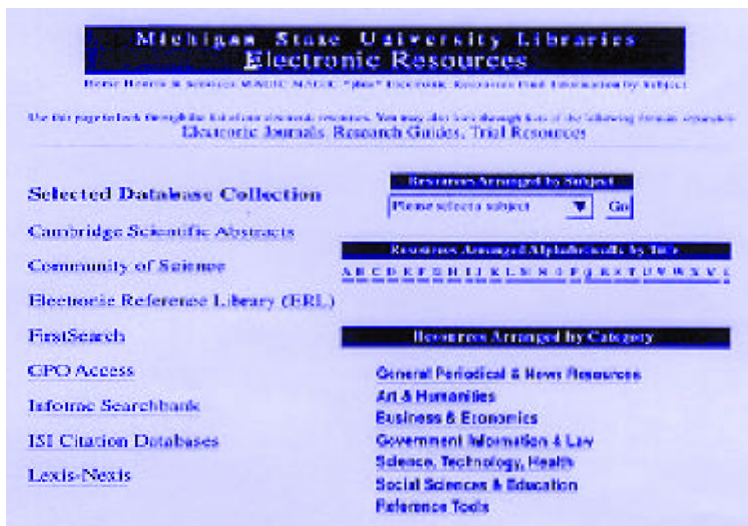
1 HOLLIS Plus, <http://hplus.harvard.edu/>; Michigan State University Libraries, http://www.lib.msu.edu/e_res/

Example 1. Harvard's HOLLIS Plus Web Site



- * proporcionar una descripción textual amplia (el vocabulario que puede ser utilizado en las búsquedas para identificar materiales relevantes *dentro de este sistema de menús*)
- * describir su cobertura, restricciones de acceso y requerimientos técnicos
- * proporcionar términos temáticos (frecuentemente no controlados) junto con ligas a otra información acerca del material como ayudas de búsqueda.
- * proporcionar un vínculo permanente al “conjunto de instrucciones” (*script*) que permite a los usuarios autorizados un acceso transparente a un recurso restringido y posicionarse en el nivel adecuado dentro de un sitio multinivel.

Example 2. Michigan State University Libraries Electronic Resources



Comúnmente, esta estructura de acceso, basada en menús, está completamente divorciada de los datos del catálogo. Existen dos problemas en esto. Las descripciones creadas por personal con diferente tipo de instrucción y perspectivas diferirán significativamente, y en consecuencia, los usuarios podrían enfrentarse con un laberinto de datos relacionados sin ninguna correspondencia clara. Por ejemplo, la forma en que el servicio Lexis-Nexis es descrito en el catálogo en línea de Harvard (*Ejemplo 4*) y la del sistema de menús de Harvard, tienen encabezamientos de materia completamente diferentes y hasta diferentes títulos.

Segundo, la información que existe en dos lugares se vuelve obsoleta en esos dos lugares. Los cambios deben efectuarse solamente una vez en un sistema, por personal capacitado que entienda todas las formas en que los datos serán usados, y por lo tanto, entienda el impacto de tales cambios.

Parece ser que existe un consenso en que ambas opciones son valiosas; por un lado, separar y marcar los recursos electrónicos en un sitio "web" para un acceso fácil, y por el otro, la integración de las fuentes electrónicas con los recursos impresos y otros materiales que aborden los mismos tópicos. Muchas bibliotecas están dirigiendo su atención hacia

un modelo que haga un mejor uso de los registros del catálogo como las bases para generar en forma automática menús “web” clasificados. La Universidad de Columbia ha usado esta técnica para proporcionar acceso a su colección de libros en línea, en donde una jerarquía temática de tres niveles es generada de los registros del catálogo con divisiones temáticas derivadas de los números de clasificación (Ejemplo 5).²

Example 3. HOLLIS Plus Web Access page for Lexis-Nexis UNIVERse

Lexis-Nexis Academic Universe

Abstract:

Lexis-Nexis Academic Universe is a gateway to digitally updated (mostly full-text) documents. The information sources included in UNIVERse were compiled specifically for the news, business and legal research needs of academic institutions. The UNIVERse database includes unmodified scanning of all format HD holders.

UNIVERse offers access to more than 7000 English language newspapers, newsmagazines, magazines, trade journals, wire services and business transcripts. Additionally, UNIVERse offers international resources in French, Italian, and German.

The legal component of the database includes comprehensive current legal information. This consists of federal and state case law, citation secondary services such as law reviews, and some legal encyclopedias.

Availability notes: This service may not be available on Sundays from Jan to May because of scheduled system maintenance.

Warning: Internet Microsoft Internet Explorer (MSIE) version 4.0: This resource will not work properly if you are using MSIE version 4 as your web browser. HOLLIS Plus developers are investigating ways to work around this bug in MSIE 4 software, but for now your only option is to use another browser (e.g., Netscape or an earlier version of MSIE). Sorry for the inconvenience.

To find out what version of MSIE you are using, from the MSIE menu bar select Help, then select About Internet Explorer. The version number will appear near the bottom of the document that displays.

Search Guide:
[Lexis-Nexis Academic Universe Search Guide](#)

Subjects:
[INTERNET AND INFORMATION SERVICES AND MEDIA](#), [Economic News and Governmental Resources](#), [LAW](#), [LAW LIBRARY](#), [Periodicals](#), [LAW LIBRARY](#)

Restrictions:
Web-based HD required.

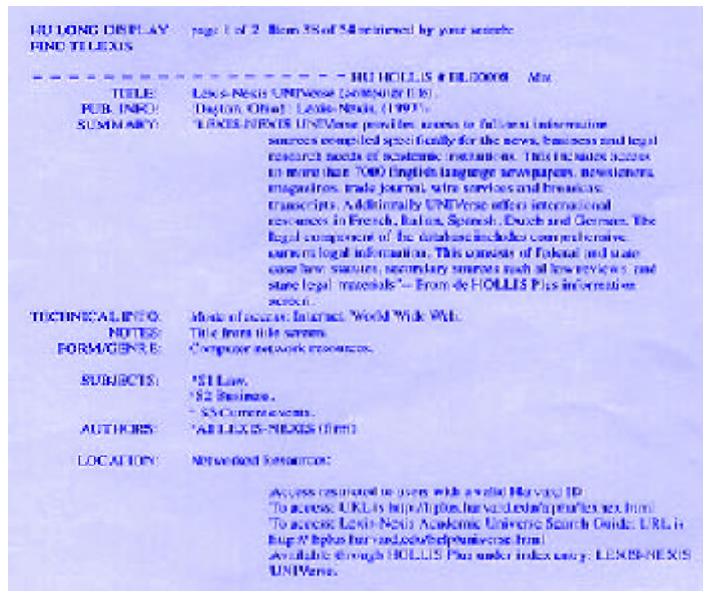
[Comment in Lexis-Nexis Academic Universe](#)

[HOLLIS Plus Home](#) | [News](#) | [Search](#) | [Comment](#) | [Help](#)

2 Columbia University Online Books Evaluation Project,
<http://www.columbia.edu/cu/libraries/digital/texts/about.html>

En el sistema de Columbia, las páginas “web” son generadas automáticamente pero no improvisadas. En otras universidades se emplea una búsqueda dinámica para generar las páginas, aunque a menudo hace falta un formateo sofisticado como el que hace Columbia.

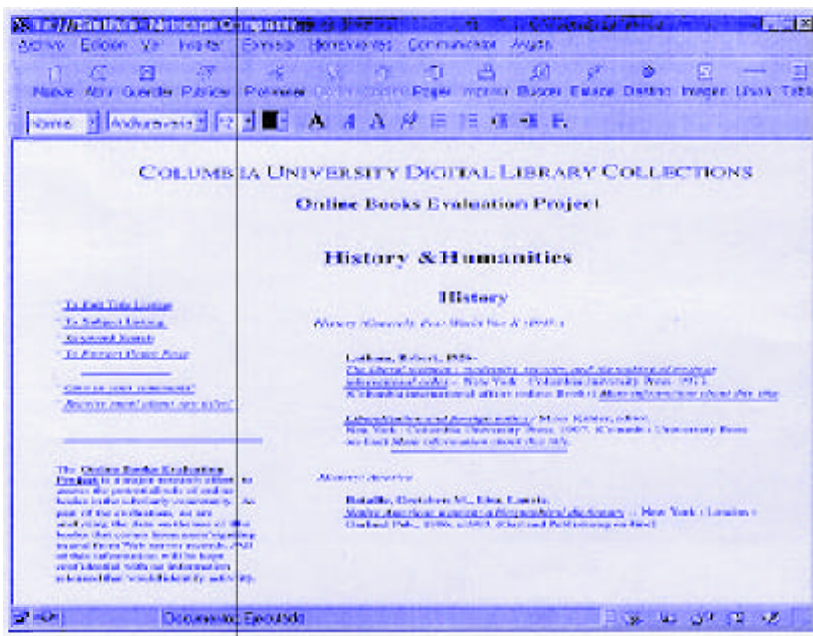
Example 4. Online catalog Record for Lexis-Nexis UNIVerse



Clifford Lynch ha señalado que la sofisticación de los sistemas comerciales “web” tales como *Amazon.com*, inevitablemente incrementarán las expectativas de los usuarios de bibliotecas en relación con metadatos que incluyan mayor valor agregado tales como ligas a reseñas y recomendaciones sucintas. Si estos tipos de funcionalidad fuesen generalmente incluidas en los sistemas de bibliotecas, y no existe ninguna razón técnica para

que no sea logrado, las bibliografías clasificadas y anotadas en línea sobre cualquier tema podrían generarse automáticamente, incluyendo todos los tipos de material. La promesa de este modelo tiene más impacto por cada dólar invertido en la catalogación: más productos que satisfagan las diferentes necesidades de acceso, con la ventaja de una consistencia interna y de un mantenimiento único. Esta es la visión de un sistema de información más rico y más integrado que cualquier catálogo en línea de la primera generación.

Example 5. Columbia University's Online Books Evaluation Project, History & Humanities Web Page



Sin embargo, la implementación de esa visión tomará trabajo. Aunque el cambio no es técnicamente difícil, organizativa y culturalmente si lo podría ser. Un análisis de costo beneficio debe de ser realizado todas las veces que la biblioteca considere proporcionar nuevos servicios. Antes de que se establezcan más ligas a información

bibliográfica cualitativa en nuestros OPACs, necesitamos determinar nuestras expectativas para los estándares y el mantenimiento de los datos, así como establecer los mecanismos apropiados para asegurar que ese mantenimiento tenga lugar. Los conceptos de control intelectual no son ampliamente enseñados a los que no son catalogadores, así que el departamento de catalogación deberá de ajustar los nuevos tipos de metadatos que serán incluidos en el catálogo, o bien, los no catalogadores deberán de ser entrenados para adoptar algunos valores de la catalogación.

Existen roles importantes para las habilidades de los catalogadores en el control intelectual, la organización y los modelos de acceso para los recursos electrónicos. Carol Mandel ha lamentado lo que ella llama la “actitud procoraza” en los departamentos de catalogación: si la información no se ajusta a un registro MARC, ese no es nuestro trabajo. Ella argumenta, convincentemente, que si los metadatos estructurales y administrativos son creados “en algún otro lugar”, el contenido digital creado localmente también será manejado “en algún otro lugar”, y si las presentaciones “web” están diseñadas “en algún otro lugar”, repentinamente, el trabajo de la biblioteca está siendo hecho “en algún otro lugar.”³ Los catalogadores tienen una oportunidad para jugar un papel crucial en el desarrollo del manejo de la información en la próxima era, siempre y cuando quieran hacerse cargo de esto.

COLOCANDO MÁS METADATOS EN LÍNEA

En adición a los nuevos tipos de metadatos anteriormente mencionados, los tipos de metadatos que actualmente existen (de acuerdo a mi definición amplia), y que diferentes departamentos de la biblioteca han continuamente creado, están ahora volviéndose accesibles en línea en forma inmediata. Como consecuencia, las relaciones y discrepancias entre ellos, así como la duplicación de esfuerzos involucrados en su creación es más obvia. Por lo tanto, algunas comunidades necesitarán desarrollar estándares donde ninguno ha existido, mientras que el reto para la mayoría de la comunidad bibliotecaria (y particularmente para los catalogadores), es construir un ambiente de información coherente fuera de los sistemas disparatados de *metadatos* en desarrollo.

Con el advenimiento de los sistemas integrales de biblioteca, el personal de adquisiciones, circulación y catalogación aprendió a crear un cuerpo coherente de

3 Mandel,, Carol. “Manifestations of Cataloguing in the Era of Metadatos,” En *Managing Metadatos for the Digital Library, Crosswaks or Chaos?*, Mayo 4-5, 1998, <http://www.columbia.edu/libraries/inside/projects/metadata/presentation/alcstlita/>

información basado en descripciones de materiales compartidos. Sin embargo, la bibliografía, los archivos y las colecciones visuales fueron dejadas de lado por esos esfuerzos. El trabajo en esas áreas está relativamente aislado; por ejemplo, la catalogación por copia y el préstamo interbibliotecario son considerablemente irrelevantes y por lo tanto, el proceso de construcción de estándares comunes no ha sido desarrollado como en otros departamentos, en donde estas funciones se consideran importantes. Los estándares interinstitucionales han sido vistos como una práctica local restrictiva que añade una carga de trabajo sin obtener ningún beneficio. Sin embargo, recientemente se han emprendido iniciativas en ambas áreas, las que eventualmente podrían conducirnos a una integración mayor con otros servicios y metadatos de la biblioteca.

COLECCIONES DE ARCHIVOS

La descripción a nivel de colecciones existente en MARC, nunca ha sido suficiente para describir colecciones de archivos en detalle. Las herramientas de búsqueda creadas por los archivistas, solamente han sido accesibles al investigador en forma personal o a través de extensas comunicaciones por correo. El desarrollo del *SGML Document Type Definition for Encoded Archival Description (EAD)*,⁴ ha permitido el acceso remoto a herramientas de búsqueda y efectuar rastreos interinstitucionales en éstas, lo cual incrementa la posibilidad de utilizar las herramientas de búsqueda junto con los registros de los catálogos para obtener un cuadro más completo de los materiales de la biblioteca. En Harvard, Berkeley y otras instituciones, existen catálogos colectivos en donde se utilizan las herramientas de búsqueda EAD.⁵

El EAD DTD define un conjunto de datos y sintaxis comunes pero también permite una gran amplitud para registrar prácticas locales. A medida de que los archivistas obtengan experiencia en la conversión de las herramientas de búsqueda, creadas para papel, a un ambiente de búsquedas en bases de datos, algunos de ellos habrán empezado a reconocer el valor de los estándares de contenido. Como un ejemplo, las fechas en las herramientas de búsqueda impresa pueden ser entendidas en una gran variedad de formatos, pero cuando esas herramientas son accesibles a través de búsquedas en bases de datos, el formato de las fechas es crucial para la recuperación. Sin formatos

4 Encoded Archival Description (EAD) DTDS <http://lcweb.loc.gov/ead/>

5 The California Digital Library Online Archive of California, <http://sunsite2.berkeley.edu/eduloc/>; The Harvard/Radcliffe Digital Finding Aids Project, <http://findingaids.harvard.edu/>

de fechas que puedan ser analizados automáticamente, sería imposible buscar por “1930” y recuperar en una herramienta de búsqueda materiales fechados entre 1870 y 1942. El desarrollo de estándares de contenido podría ser un cambio radical para una comunidad que privilegia las practicas locales y pondera altamente su autonomía. Sin embargo, debido a las crecientes expectativas en torno al uso de las herramientas de búsqueda en línea, podemos vislumbrar un amplio panorama para las habilidades y estándares del catalogador dentro de la comunidad archivística.

MATERIALES VISUALES

La comunidad que maneja materiales visuales también ha sido energizada por el potencial de una diseminación amplia de la información a través de la “Web.” Diversas organizaciones y proyectos tales como el *Consortium for The Computer Interchange of Museum Information (CIMI)*,⁶ el *Museum Educational Site Licencing Project (MESL)*,⁷ el *Record Export for Art and Cultural Heritage (REACH)*,⁸ el *Museum Digital Licensing Colective (MDLC)*⁹ y el *Art Museum Image Consortium (AMICO)*,¹⁰ avalan el creciente interés en sistemas de información cooperativos dentro de la comunidad relacionada con la herencia cultural; asimismo, esfuerzos similares están llevándose a cabo en las colecciones de historia natural.

En Harvard, las bibliotecas con acervos de transparencias, los museos y los archivos gráficos, han unido esfuerzos para desarrollar un “catálogo colectivo” de los materiales visuales. Los repositorios individuales retendrán sus sistemas para la organización de su colección local pero exportarán un grupo común de datos al catálogo colectivo, basado en parte en las categorías básicas de la Asociación de Recursos Visuales [*Visual Resources Association*].¹¹ Los participantes reconocen la necesidad de trabajar hacia una semántica común y a la implementación de vocabularios controlados con el propósito de crear una base de

6 Consortium for the Computer Interchange of Museum Information, <http://www.cimi.org>

7 Museum Educational Site Licencing Project,
<http://www.gii.getty.edu/mesl/>

8 Record Export for Art and Cultural Heritage,
<http://www.rlg.org/reach.html>

9 Museum Digital Licensing Collective, Inc., <http://www.museumlicensing.org/>

10 Art Museum Image Consortium, <http://www.amico.net>

11 Visual Resources Association Core Categories, Version 2.0,
<http://www.oberlin.edu/~art/vra/wcl.html>

datos colectiva efectiva. Aunque existe gran entusiasmo en esta comunidad para la cooperación, los tipos de repositorios involucrados y consecuentemente, sus necesidades de información, son extremadamente diversas. El proceso de desarrollar estándares comunes y ajustar las culturas institucionales en apoyo a tales estándares será una enorme tarea. Aquí nuevamente vemos que los principios de la catalogación ganan terreno y que la teoría de la catalogación, además de las habilidades del catalogador, crecen en importancia.

EN CONCLUSIÓN

El reto venidero será crear un ambiente de información coherente alejado de esos sistemas de acceso disparatados. En Harvard esperamos tener tantos sistemas locales como sean necesarios para llenar las funciones de organización de la colección, aunque para el acceso público tendremos una serie de catálogos interrelacionados, cada uno diseñado de acuerdo a las características y requerimientos funcionales de los tipos particulares de materiales, por ejemplo, imágenes y objetos de museo, archivos y colecciones de biblioteca. Además, una plataforma de acceso común podrá permitirle al usuario buscar a través de cualquier combinación de esos catálogos, especificando qué tipos de material desea recuperar y obtener un grupo de resultados significativos, manejables, navegables e integrados. Para lograr esto, aún dentro de una sola institución, se requerirá de un gran trabajo y buena fe por parte de mucha, mucha gente. El lograr esto a una escala mayor, será aún más difícil.

Es claro que las bibliotecas enfrentan un cambio gradual pero profundo en la organización y la presentación de sus colecciones. Estas tareas no pueden ser realizadas exclusivamente por los catalogadores, pero bajo ninguna circunstancia, éstas deben llevarse a cabo sin los catalogadores y sin las habilidades y entrenamiento particular que ellos poseen. A medida que estándares de metadatos muy simples como los del Núcleo Dublín sean desarrollados, los catalogadores son quienes serán capaces de entender cómo pueden implementarse en las bibliotecas, bajo qué circunstancias y dentro de qué líneas de acción, sin comprometer la misión de la biblioteca. Las comunidades que han hecho mofa de los misterios de las RCA2 y MARC, están empezando a ver el valor de los estándares de contenido y podrían beneficiarse de la experiencia y perspectivas de los catalogadores a medida que desarrollen sus propios estándares. En la medida que el trabajo descriptivo en esas comunidades reproduzca más estándares de ese tipo, habrá una mayor demanda y se dará más valor a la habilidad para establecer un orden en el caos, lo cual es el objetivo principal de la teoría y práctica de la catalogación.

Puntos de vista: acceso y navegación convencionales y “neoconvencionales” en colecciones digitales

GERRY MCKIERNAN

Universidad Estatal de Iowa
gerrymck@iastate.edu

*La mejor manera de predecir
el futuro es inventarlo*
Alan Kay

CONVENCIONALES

En el otoño de 1995, *Cyberstacks(sm)*, una colección de fuentes de referencia *World Wide Web (WWW)* sobre ciencia y tecnología en Internet, fue formalmente establecido en la Universidad Estatal de Iowa.¹ Además de ser una colección selecta de fuentes significativas de referencia en texto completo, *Cyberstacks(sm)* también ha servido como un prototipo para investigar el uso de uno de los principales esquemas de clasificación bibliográfica como marco organizativo y ambiental en el cual se apliquen o añadan otras convenciones de la organización bibliotecaria, previamente utilizadas, para mejorar la identificación y el uso de los recursos de la “Red.”² Adicionalmente, *Cyberstacks(sm)* ha sido

- 1 CyberStacks(sm). Página WWW en Internet, URL: <http://www.public.iastate.edu/~CYBERSTACKS> (versión actualizada al 4 de julio 1998).
- 2 Gerry McKiernan, “The New/Old World Wide Web Order: The Application of ‘Neo-Conventional’ Functionality to Facilitate Access Use of a WWW Database of Science and Technology Internet Resources,” *Journal of Internet Cataloging 1*, no. 1 (1997): 47-55.

utilizado como la base para examinar y analizar las características y funcionalidad de esfuerzos similares.³

SISTEMAS DE CLASIFICACIÓN BIBLIOGRÁFICA

El sistema de clasificación de la Biblioteca del Congreso es un sistema apropiadamente desarrollado que ha sido usado por generaciones en las bibliotecas para organizar una variedad de materiales impresos y no impresos. Dentro de su esquema de organización del conocimiento,⁴ este sistema de clasificación no solamente denota una cobertura temática y de contenido, sino también el formato de la información y relaciones conceptuales. Considerándolo un adecuado marco organizativo, en la etapa inicial de la creación de Cyberstacks(sm) creímos que éstas y otras características similares, además de las generales de este sistema de clasificación, podrían proporcionar el contexto y estructura adecuadas para facilitar el acceso a la “Red.” A partir de su establecimiento, más de 1,500 recursos significativos de Internet sobre ciencia y tecnología, además de otros de varias disciplinas de las ciencias sociales y las humanidades, han sido seleccionados para la colección de Cyberstacks(sm) y listados en su Índice de Títulos como recursos perfilados o candidatos. Varias docenas de individuos, instituciones y organizaciones de más de seis continentes, actualmente han establecido ligas a Cyberstacks(sm), o a alguna de sus categorías principales. Aunque hemos creído que la utilización de un sistema de clasificación reconocido proporciona un marco de valor agregado, suficiente para la navegación de los recursos “web”, el escepticismo de algunos en relación con la pertinencia de esta opción, nos llevó a realizar una revisión sistemática de otros proyectos que habían adoptado sistemas de clasificación bibliográfica para organizar los recursos de Internet. A partir de ese estudio, identificamos varios esfuerzos mundiales que han utilizado esquemas de clasificación bibliográfica locales, nacionales o internacionales para la organización de la “Web.” Entre estos se encuentran aquellos que han usado esquemas internacionales tales como la Clasificación Decimal de Dewey (CDD) (e. g., CyberDewey) y la Clasificación Decimal Universal (CDU)

3 Gerry McKiernan, “Beyond Bookmarks: A Review of Frameworks, Features, and Functionalities of Schemes for Organizing the Web,” *Internet Reference Services Quarterly* 3, no. 1 (1998): 69-82.

4 Library of Congress. Office for Subject Cataloging Policy, *LC classification outline*, 6th ed. (Washington, DC: Library of Congress, Office for Subject Cataloging Policy, 1990).

(e. g., WWW Subject Tree of WAIS Database [Nordic WAIS/World Wide Web Project]), o sistemas de clasificación nacionales y locales tales como la Clasificación de la Biblioteca Danesa de Agricultura y Veterinaria (Internet Efter Emne), el sistema Nederlandse Basisclassificatie (DutchESS) y la clasificación Sveriges Allmänna Biblioteksörening (SAB) [Systematisk Internetkatalog (Möndals Stadsbibliotek)].

VOCABULARIO CONTROLADO

Adicionalmente a la organización de los recursos “web” por medio del uso de sistemas de clasificación tradicionales, algunos sitios han ampliado el acceso a sus colecciones particulares por medio de recursos indizados con descriptores o encabezamientos de materia normalizados. Una de tales colecciones, *ADAM (Art, Design, Achitecture & Media Information Gateway)* — un servicio desarrollado para ayudar en la identificación de las fuentes de información de “calidad asegurada” en Internet en las áreas de bellas artes, diseño, arquitectura, artes aplicadas, multimedia y estudios de museo y conservación — ha asignado los términos incluidos en el Tesoro de Arte y Arquitectura a los recursos seleccionados, además del número de la Clasificación Decimal de Dewey. Otro sitio que emplea ambos elementos, un sistema de clasificación y un vocabulario temático para enriquecer el acceso a su colección particular es *Scout Report Signpost*, una colección organizada por un esquema abreviado de la clasificación de la Biblioteca del Congreso, la cual puede ser recorrida usando una lista alfabética de los Encabezamientos de Materia de la Biblioteca del Congreso (*Library of Congress Subject Headings, LCSH*).⁵ Uno de los sitios “web” más sofisticados que emplean un vocabulario controlado para navegar en su colección es OMNI (*Organising Medical Networked Information*), el portal británico para los recursos de Internet sobre medicina, biomedicina, aspectos de salud, administración sanitaria y tópicos relacionados. Además de indizar su colección empleando los encabezamientos de materia médicos estandarizados del MeSH® (*Medical Subject Headings*), este sitio también proporciona la oportunidad a sus usuarios de revisar y buscar por términos y conceptos *relacionados* con un término o frase de búsqueda empleando el Sistema Unificado de Lenguaje Médico (*Unified Medi-*

5 Aimee D. Glassel and Amy Tracy Wells, “Scout Report Signpost: Design and Development for Access to Cataloged Internet Resources,” *Journal of Internet Cataloging* 1, no. 3 (1998): 15-45.

cal Language System, UMLS®), desarrollado por la Biblioteca Nacional de Medicina de los E.U (*National Library of Medicine, NLM*).⁶

Reporte DESIRE

Como parte de un estudio mayor, desarrollado bajo los auspicios del proyecto *DESIRE (Development of a European Service for Information on Research and Education)*, una revisión detallada y completa sobre la aplicación de la clasificación bibliográfica y los vocabularios controlados para organizar y mejorar el acceso a los recursos “web,” ha sido preparada por Traugott Koch del *NetLab* de la Universidad Lund en Suecia y Michael Day de *UKOLN*, Gran Bretaña, con la colaboración de otros colegas.⁷ *DESIRE* es un proyecto a gran escala auspiciado por el Cuarto Programa Estructural de la Unión Europea sobre Telemática para el Sector de Investigación. Estos y otros esfuerzos han sido incluidos en un centro distribuidor de información denominado *Beyond Bookmarks (Figura. 1)*.⁸

Figura 1. Beyond Bookmarks: Schemes for Organizing the Web
<http://www.public.iastate.edu/~CYBERSTACKS/CTW.htm>



6 National Library of Medicine, *Fact Sheet: Unified Medical Language*. Internet Página WWW en Internet, URL: <<http://www.nlm.nih.gov/pubs/factsheets/uMLS.html>> (versión actualizada al 5 de julio de 1998).

7 Traugott Koch and Michael Day. *The Role of Classification Schemes in Internet Resource Description and Discovery, Work Package 3 of Telematics for Research project DESIRE (RE 1004)*. Página WWW en Internet, URL: <<http://www.ukoln.ac.uk/metadata/desire/classification/>> (versión actualizada al 5 de julio de 1998).

8 *Beyond Bookmarks: Schemes for Organizing the Web*. Página WWW en Internet, URL: <<http://www.public.iastate.edu/~CYBERSTACKS/CTW.htm>> (versión actualizada al 5 de julio 1998).

NEOCONVENCIONALES

Aunque creado para facilitar el acceso y uso de los recursos “web” importantes, a través de la aplicación de un esquema de clasificación bibliográfica internacional y otros métodos de organización bibliográfica convencionales, *Cyberstacks(sm)* también ha servido como un medio para explorar otras opciones y alternativas para la organización, administración y navegación de los recursos de la “Red.”⁹ Estas investigaciones han incluido una revisión de la “categorización automatizada” de los recursos “web” y un encuesta extensa de las tecnologías sobre visualización de información. Más recientemente, hemos establecido un centro de distribución de información dedicado a los proyectos “web” y “no web” de despliegue auditivo,¹⁰ llevado a cabo una revisión preliminar de los dispositivos interactivos hápticos¹¹ y comenzado un examen de los proyectos de ambiente virtual.^{12,13}

CATEGORIZACIÓN AUTOMATIZADA

Desde su establecimiento, el asunto del crecimiento y ampliación del modelo *Ciberstacks(sm)* para la organización de la “Web” ha sido de interés general y señalado por varios usuarios. Aunque la adopción de los esquemas de clasificación bibliográfica por muchas bibliotecas y agencias de servicios de información es una evidencia muy fuerte de que los métodos tradicionales de organización bibliográfica pueden ofrecer un marco de trabajo para la identificación y uso efectivo de los

- 9 Gerry McKiernan, “Hand-Made in Iowa: Organizing the Web Along the Lincoln Highway,” *D-Lib Magazine* (Feb. 1997). Página WWW en Internet, URL: <<http://www.dlib.org/dlib/february97/02mckiernan.html>> (versión actualizada al 5 de julio de 1998).
- 10 The Next WAVE(sm): Auditory Browsing in Web and non-Web Databases. Página WWW en Internet, URL: <<http://www.public.iastate.edu/~CYBERSTACKS/Wave.htm>> (versión actualizada al 5 de julio de 1998).
- 11 *The Magic Touch(sm): Haptic Interaction in Web and non-Web Databases*. Página WWW en Internet, URL: <<http://www.public.iastate.edu/~CYBERSTACKS/Touch.htm>> (versión actualizada al 5 de julio de 1998).
- 12 *Sensory Information Navigation in Virtual Environments*. Página WWW en Internet, URL: <<http://www.public.iastate.edu/~CYBERSTACKS/SINVE.htm>> (versión actualizada al 5 de julio de 1998).
- 13 Gerry McKiernan, “As the World (Wide Web) Turns: Resources at Iowa State,” *D-Lib Magazine* (July/Aug. 1998). Página WWW en Internet, URL: <<http://www.dlib.org/dlib/july98/07clips.html#GERRY>> (versión actualizada al 5 de julio de 1998).

recursos “web”, tales medidas, generalmente requieren de un esfuerzo constante para su creación y mantenimiento. Las limitaciones inherentes a esta opción, subsecuentemente nos condujeron a una investigación de los esfuerzos que proponen la categorización automatizada de los recursos “web”¹⁴ y a la creación de un centro de intercambio de información, el cual incluye los perfiles de proyectos, servicios, investigación y productos importantes que proporcionan ese desarrollo en las estaciones de trabajo, sistemas y niveles de redes.¹⁵ A nivel de sistema, existen numerosos proyectos que emplean agentes de software inteligente para navegar la “Web” en búsqueda de productos, servicios o recursos que concuerden con el interés de un usuario (e. g., *Firefly*), o que facilitan el acceso a los recursos de la “Red” por medio del establecimiento de espacios conceptuales generales o específicos, enfocados a la búsqueda (e. g., *WebGlimpse*). A nivel de redes, existen proyectos notables tales como el *WAIS/World Wide Web*, el cual es llevado a cabo por grupos de trabajo de la Universidad Lund de Suecia y de la Biblioteca Tecnológica Nacional de Dinamarca; o bien, *Scorpion*, un proyecto de investigación bajo el liderazgo de Keith Shafer de la Oficina de Investigación y Proyectos Especiales de OCLC. Estos dos últimos proyectos no solamente se han esforzado por crear un acceso más amplio a los recursos de la Red a través de la categorización u organización automatizada de los recursos de Internet, sino que también ha extendido tal categorización hacia la clasificación automática. Ahora que el proyecto *WAIS/World Wide Web* ha sido concluido, éste puede identificar y clasificar exitosamente, las bases de datos *WAIS (Wide Area Information Server)* dentro de la CDU y construir un árbol temático *WAIS*, basado en los dos niveles principales de este esquema de clasificación.¹⁶ *Scorpion* ha buscado crear un servicio similar por medio del análisis del contenido de los recursos electrónicos y la asignación de números de clasificación tentativos, tomados del sistema CDD.¹⁷ La tesis de *Scorpion* es que la Clasificación Decimal de Dewey

14 Gerry McKiernan, “Automated Categorisation of Web Resources: A Profile of Selected Projects, Research, Products, and Services,” *The New Review of Information Networking* 2: 15-40.

15 *Project Aristotle(sm): Automated Categorization of Web Resources*. Página WWW en Internet, URL: <<http://www.public.iastate.edu/~CYBERSTACKS/Aristotle.htm>> (versión actualizada al 5 de julio de 1998).

16 Anders Ardö *et al.*, “Improving Resource Discovery and Retrieval on the Internet: The Nordic WAIS/World Wide Web Project-Summary Report,” *NORDIN-FO-Nytt*, 1994. no. 4: 13-28.

17 Keith Shafer, *The Scorpion HomePage*. Página WWW en Internet, URL: <<http://orc.rsch.oclc.org:6109>> (versión actualizada al 5 de julio de 1998).

puede emplearse para llevar a cabo una asignación temática automática para los documentos electrónicos, es decir, el esquema puede ser utilizado para clasificar un documento y sugerir los encabezamientos de materia asociados, y además, comunicar el contenido.

En adición a la identificación de proyectos significativos que buscan la organización de los recursos de la “Red,” empleando sistemas tradicionales de clasificación bibliográfica dentro de un ambiente automatizado, esta revisión también identificó varios proyectos que empleaban agentes de software inteligente,¹⁸ o inteligencia artificial (IA),¹⁹ para asegurar o mejorar la organización de los recursos de la “Red”. Entre los más fascinantes está la aplicación en la organización de recursos “web” del método de redes neuronales desarrollado por Kohonen, denominado Mapa de Autoorganización (*Self-Organizing Map, SOM*); otro proyecto notable es el denominado *ET-Space*,²⁰ una colección de sitios “web” de entretenimiento desarrollado por Hsinchun Chen de la Universidad de Arizona con la participación del Proyecto de Biblioteca Digital [Digital Library Initiative (DLI)] de la Universidad de Illinois en Urbana-Champaign. La técnica SOM de Kohonen es un algoritmo de aprendizaje no supervisado para analizar y visualizar datos estadísticos de alta dimensión, desarrollado por Teuvo Kohonen y su equipo en el Centro de Investigación de Redes Neuronales de la Universidad Tecnológica de Helsinki, Finlandia. El SOM organiza documentos automáticamente en una matriz de dos dimensiones para que los documentos relacionados aparezcan físicamente cerca uno del otro. Kohonen y su equipo, también han empleado el método MAO para organizar los recursos de Internet y comunicaciones electrónicas sobresalientes acerca de este asunto se encuentran en su proyecto *WebMAO*.²¹ Estos y otros esfuerzos han sido incluidos en un centro de intercambio de información denominado *Aristotle Project (Figura 2)*²²

18 Nick R. Jennings and Michael J. Wooldridge, eds. *Agent Technology: Foundations, Applications, and Markets* (Berlin: Springer, 1998).

19 George F. Luger, and William A. Stubblefield, *Artificial Intelligence: Structure and Strategies for Complex Problem Solving*, 3d de. (Reading, MA: Addison-Wesley, 1998).

20 Hsinchun Chen, Chris Schuffels and Richard Orwig, “Internet Categorization and Search: A Self-Organizing Approach,” *Journal of Visual Communication and Image Representation* 7, no. 1 (March 1996): 88-102.

21 Timo Honkela et al., “Self-Organizing Maps of Document Collections,” *alma* 2. Página WWW en Internet, URL: <<http://www.diemme.it/~luigi/websom.html>> (versión actualizada al 5 de julio de 1998).

22 *Project Aristotle(sm)*, *op. cit.*

Figure 2. Aristotle Project (sm): Automated Categorización of the Web<http://www.public.iastate.edu/~CYBERSTACKS/Aristotle.htm>



EL SURGIMIENTO DEL CEREBRO EXACTO

En adición a proporcionar un medio para organizar los recursos “web” en forma automática, las aplicaciones de SOM, al igual que otras identificadas en las investigaciones sobre categorización automática, también ofrecieron formatos y métodos alternativos para desplegar y acceder a los recursos dentro de una colección determinada. Tales presentaciones alternativas, estimularon el interés en la visualización de la información y su potencialidad para enriquecer la identificación y uso de los recursos de la “Red”,²³ subsecuentemente, enviamos solicitudes a varias listas de discusión y grupos de noticias, requiriendo información acerca de los esfuerzos actuales en este campo. De las respuestas a esas solicitudes y de una subsecuente revisión exhaustiva de la literatura, identificamos varias decenas de grandes proyectos y productos que empleaban esta opción y establecimos un nuevo centro de intercambio de información dedicado a esta tecnología y a sus aplicaciones, directas o potenciales, para mejorar la navegación en la “Web.”²⁴ Entre los esfuerzos más notables, identificados a través de esta investigación, estuvieron *SiteMap*, un proyecto de Xia Lin, ahora de la Universidad Drexel, y *SPIREtm*, una aplicación novedosa de visualización desarrollada por Jamos A. Wise, James J. Thomas y el equipo del Laboratorio Nacional del Pacífico Noroeste en

23 Gerry McKiernan, Information Visualisation: The World Wide Web Gets Really Graphical. *Intelligence: The Magazine of the Information Age* (Dec. 1996). 116-118.

24 *The Big Picture(sm): Visual Browsing in Web and non-Web Databases*. Página WWW en Internet, URL: <<http://www.public.iastate.edu/~CYBERSTACKS/BigPic.htm>> (versión actualizada al 5 de julio de 1998).

Richland, Washington. *SiteMap* es una aplicación de Java que visualiza un sitio “web” o una colección de ligas.

A través de *WebCrawler*, *SiteMap* primeramente recorre cada liga del sitio “web”, recolecta datos estadísticos e indiza todas las palabras y páginas del sitio. Basado en los datos estadísticos y en la indización, *SiteMap* convierte cada página del sitio en un vector y usa estos vectores para instruir a una red neuronal. Como resultado, la red neuronal entrenada presenta el sitio en un mapa organizado: las áreas temáticas son identificadas y etiquetadas; sus tamaños y localizaciones son determinados por las relaciones entre los temas y por sus frecuencias de ocurrencia y co-ocurrencia. Las ligas o vínculos son agrupados y ubicados dentro de sus áreas temáticas respectivas, representadas por puntos de colores... . Para auxiliar a los usuarios en su interacción con el mapa, *SiteMap* proporciona varias herramientas interactivas. Por ejemplo, las áreas pueden ser etiquetadas en más/menos detalle a través del ajuste de una barra de avance/retroceso; los vínculos pueden seleccionarse, pulsándolos o arrastrándolos con el “ratón”; los contenidos de cualquier vínculo seleccionado pueden mostrarse en una ventana por separado, etc.²⁵

*SPIRE*_{tm}, o *Spacial Paradigm for Information Retrieval and Exploration*, “acepta grandes volúmenes de texto en casi cualquier formato, determina las relaciones dentro del texto y las presenta en un formato visual que es natural para la mente humana.”²⁶

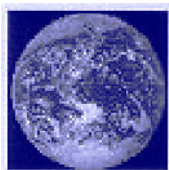
*SPIRE*_{tm} despliega en forma gráfica imágenes basadas en similitudes de palabras y temas en el texto. No se requiere un conocimiento previo de la información o de la selección de los temas o tópicos. *SPIRE*_{tm} crea sus visualizaciones por medio del procesamiento de esas similitudes dentro de temas y tópicos clave y organiza los datos dentro de representaciones visuales que le permiten al usuario explorar y descubrir las relaciones entre el texto de los documentos. Dos tecnologías dentro de *SPIRE*_{tm}, *Galaxies* y *Themescape* proporcionan metáforas visuales naturales que requieren de muy poco entrenamiento para su uso. *Galaxies* cuenta las similitudes y patrones de las palabras en los documentos y después despliega los documentos dentro de la pantalla de la computadora como en un universo de “docuestrellas”. Los documentos que están estrechamente relacionados se apiñan en un grupo estrecho, mientras los documentos no relacionados se encuentran separados por grandes espacios. En *Themescape*, los temas dentro de los espacios del documento aparecen en la pantalla de la computadora como un mapa en relieve de un terreno natural. Las montañas en *Themescape* indican donde los temas son dominantes, los valles indican los temas irrelevantes. Sus formas — una depresión amplia o un pináculo alto — reflejan cómo la información temática es distribuida y relacionada a través de los documentos

25 Xia Lin, *Visual SiteMap*. Página WWW en Internet, URL: <<http://lislin.gws.uky.edu/Sitemap/Sitemap.html>> (versión actualizada al 5 de julio de 1998).

26 Pacific NorthWest National Laboratory, *Showcase: VisualText Analysis: SPIRE*. Página WWW en Internet, URL: <http://multimedia.pnl.gov:2080/showcase/?it_content/spirenode> (versión actualizada al 5 de julio de 1998).

transversales relacionados. Los temas cercanos al contenido serán visualizados más estrechamente, basados en las numerosas relaciones dentro de los espacios del texto.²⁷ Estos y otros esfuerzos han sido incluidos en un centro de intercambio de información denominado *The Big Picture (sm)* (Figura. 3).²⁸

Figura 3. *The Big Picture(sm)* :
Visual Browsing of Web and non-Web Databases
<http://www.public.iastate.edu/~CYBERSTACKS/Big Pic.htm>



NAVEGACIÓN AUDITIVA

Mientras muchos piensan que la visualización puede ofrecer una mejor navegación de las colecciones “web”, otros creen que muchas visualizaciones pueden ser difíciles de usar, debido en parte a la complejidad de algunas técnicas de visualización o las inherentes limitaciones de la mayoría de los monitores de las estaciones de trabajo actuales. Para aumentar el despliegue visual de los documentos “web”, algunos investigadores han empezado a estudiar la aplicación de las tecnologías de despliegue auditivo para enriquecer la interfase gráfica del usuario (*GUI*) y la navegación de las páginas “web”. La mayoría de esas opciones han sido desarrolladas para auxiliar al discapacitado en el uso de la Red y para otros tienen la potencialidad de ofrecer una mejor navegación en la “Web”. Entre los productos más sofisticados para ayudar a los invidentes en el uso de documentos “web” está *pwWebSpeak*™, “un navegador de Internet diseñado para usuarios que desean acceder Internet de un modo no visual, o combinado modos visuales y auditivos.”²⁹

27 *Ibid.*

28 *The Big Picture(sm)*, *op.cit.*

29 Productivity Works, Inc., *pwWebSpeak overview*. Página WWW en Internet, URL:
<<http://www.prodworks.com/pwvovw.htm>> (versión actualizada al 5 de julio de 1998).

“*pwWebSpeak*™ está específicamente diseñado para interactuar directamente con la información de las páginas “web” y traducir el contenido de la información a una forma hablada y a un despliegue visual simplificado. El usuario podría navegar a través de la estructura de un documento basado en su contenido, párrafos y oraciones, más que tener que lidiar con teclas de avance y retroceso de texto y la interpretación de un despliegue de pantalla estructurado. La inteligencia construida dentro de *pwWebSpeak* entiende los constructos de HTML y automáticamente evita aquellos otros que no tienen relación con el contenido de la información de un documento. Ambos aspectos, el oral y la interpretación de los principales caracteres de las páginas “web”, están diseñados para que todas las clases de usuarios puedan emplear el software de una manera efectiva.”³⁰

Una tecnología relacionada puede ser encontrada en el proyecto *AHA* (*Audio HTML Access*), desarrollado por Frankie James en la Universidad de Stanford. “El sistema *AHA* está basado sobre el principio de que los archivos *HTML* explícitamente incluyen el contenido textual y estructural de un documento, y que estos dos tipos de contenido son esenciales para la comprensión del documento.”³¹ Uno de los usos más novedosos del enriquecimiento auditivo fueron incluidos en los estudios llevados a cabo por Mark E. Rorvig, quién ahora trabaja en la Universidad del Norte de Texas en Denton. En una investigación pionera de hace aproximadamente diez años, él exploró la representación musical de los esquemas de la Clasificación Decimal de Dewey. En sus experimentos, Rorvig convirtió en “auditivas”... secuencias repetitivas de las clases del sistema de Clasificación Decimal de Dewey (CDD), tomadas de los esquemas y de los números de clasificación del índice....

[L]as clasificaciones del índice para ‘Hombres’, ‘Mujeres’, ‘Juegos’, y ‘Civil’ fueron convertidas a valores tonales y reproducidos por la máquina. Adicionalmente, los esquemas para los números de las clasificaciones 355.1 y 220, ‘Vida Militar’ y la ‘Biblia’, fueron también fueron procesados y reproducidos.”³²

Estos y otros esfuerzos han sido incluidos en un centro de intercambio de información denominado *The Next WAVE(sm)* (*Figura 4*).³³

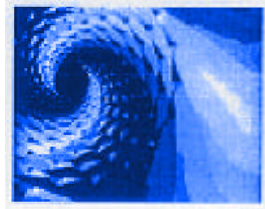
30 *Ibid.*

31 Frankie James, “AHA: Audio HTML Access. *Computer Networks and ISDN Systems* 29 nos. 8-13 (sept. 1997): 1395-1404.

32 Mark E. Rorvig and Paul. T. Orkiszewski, “An Aural Interface for Exploration of Large Bibliographic Databases: The Sound of Subjects,” En *Proceedings of the Tenth National Online Meeting, May 9-11, 1989, New York, New York*, eds. Carol Nixon and Lauree Padgett (Medford, NJ: Learned Information, 1989).

33 *The Next WAVE(sm) op. cit.*

Figura 4. The Next WAVE(sm) :
Auditory Browsing of Web and non-Web Databases
<http://www.public.iastate.edu/~CYBERSTACKS/Wave.htm>



DISPOSITIVOS INTERACTIVOS HÁPTICOS

Puede esperarse que la habilidad para interactuar con las colecciones “web” sea mejorada con la comercialización y el uso de los dispositivos interactivos hápticos, actualmente en desarrollo, o recientemente introducidos en el mercado de los videojuegos. Los dispositivos hápticos son tecnologías que permiten al usuario, literalmente sentir y tocar los objetos dentro de un ambiente basado en la computadora. Una de las tecnologías hápticas más promisorias, que puede esperarse facilite aún más la navegación de las colecciones “web” en un futuro cercano, es el *FEEL-itMouse*, un producto de la Corporación *Immersion*, recientemente anunciado.

Con el uso del *FEEL-itMouse*, cualquier objeto desplegado en su pantalla puede percibirse como sensaciones táctiles reales. Arrastre un icono con el *FEEL-itMouse* y sentirá el *estirón* como si verdaderamente fuera elástico. Al explorar texturas, líquidos, superficies, vibraciones, el *FEEL-itMouse* le permite sentir casi cualquier fenómeno físico, trayendo un realismo natural a las interacciones con el software que usted nunca soñó que pudieran ser posibles.

El *FEEL-itMouse* no es solamente un señalador, es una “tecnología de información” sofisticada que permite una comunicación física bidireccional entre el usuario y la máquina. El usuario empuja el ratón para controlar el cursor y el ratón empuja al usuario para simular encuentros físicos.³⁴

34 Immersion Corporation, FEELit MouseOverview, Página WWW en Internet, URL:
<<http://www.force-feedback.com/feelit/feelit.html>> (versión actualizada al 5 de julio de 1998).

Una fuente notable acerca de los dispositivos hápticos es elFAQ mantenido en el sitio “web,” *sci.virtual-worlds*.³⁵

Esta y otras tecnologías similares están actualmente siendo revisadas para ser incluidas en un centro de intercambio de información, en fase de planeación, denominado *The Magic Touch(sm)* (Figura 5).³⁶

**Figura 5. *The Magic Touch(sm):
Haptic Interaction in Web and non-Web Databases***
<http://www.public.iastate.edu/~CYBERSTACKS/Touch.htm>



PROCESAMIENTO DE LENGUAJE NATURAL

El Procesamiento de Lenguaje Natural (PLN) es “un grupo de técnicas computacionales para analizar y representar en forma natural los textos que aparecen en uno o más niveles del análisis lingüístico, con el propósito de lograr un procesamiento del lenguaje, semejante al humano, para un rango de tareas o aplicaciones particulares.”³⁷ De acuerdo con Liddy,³⁸ estos niveles del análisis lingüístico comprenden: “léxico (análisis a nivel de las palabras, incluyendo el significado lexicográfico y el análisis de las partes del habla); sintáctico (análisis de palabras en una oración con el propósito de descubrir la estructura gramatical de dicha oración); semántico (la determinación de los posibles significados de una oración, incluyendo la desambigüedad en el contexto); discurso (interpretación

35 *sci.virtual-words*, *Haptics Frequently Asked Questions* (FAQ), Página WWW en Internet, URL: <http://www.hitl.washington.edu/projects/knowledge_base/virtual-worlds/haptics-faq.html> (versión actualizada al 5 de julio de 1998).

36 *The Magic Touch(sm)*, *op.cit.*

37 Elizabeth D Liddy, “Enhanced Text Retrieval Using Natural Language Processing,” *Bulletin of the American Society for Information Science* 24, no. 4 (Apr./May 1998): 14.

38 *Ibid.*, 14-15.

de la estructura y el significado transmitidos por los textos mayores a una oración); y pragmático (el entendimiento del uso propositivo del lenguaje situacional, particularmente aquellos aspectos del lenguaje que requieren conocimiento universal)". Aunque muchos sistemas de computadoras han hecho uso de algún nivel del procesamiento lingüístico para mejorar la recuperación de la información para más de una generación, ninguno ha incorporado tantos niveles de procesamiento del lenguaje como el modelo *DR-LINK*, delineado por Liddy hace diez años.³⁹

DR-LINK

DR-LINK es "un sistema de procesamiento de textos inteligente que analiza texto completo no sólo por sus palabras componentes, sino también por el significado inherente en la estructura del texto. Esto le permite al usuario localizar ideas difíciles de encontrar tales como tendencias, predicciones o comparaciones.... recupera información sobre conceptos relacionados, aún si las palabras no corresponden exactamente a la búsqueda original. Usa el contexto para evitar el sentido ambiguo de las palabras con el propósito de minimizar los resultados falsos. El sistema construye bases de texto automáticamente en cualquier área temática."⁴⁰

CHESS

Basándose en la funcionalidad del avanzado procesamiento lingüístico de *DR-LINK*, Liddy y sus colegas de *TextWise*tm, una compañía de investigación y desarrollo localizada en el Centro de Ingeniería de Software y Aplicaciones de Cómputo de la Universidad de Syracuse, han desarrollado *CHESS*, *Chronological Information Extraction System*, el cual "lleva la recuperación de la información más allá de las fronteras de la recuperación basada en documentos."⁴¹

Por medio de la extracción de bloques básicos del conocimiento incluido en el texto, este sistema automáticamente crea una base de conocimiento que agrega información acerca de cualquier entidad nombrada — gente, lugares, eventos, organizaciones, compañías o ideas— y organiza ese conocimiento en una línea de tiempo que cubre el periodo completo de la base de conocimiento. La información es presentada al usuario como un sumario cronológico con vínculos a los documentos originales y como una

39 Elizabeth DuRoss Liddy, "The Discourse-level Structure of Natural Language Texts: An Exploratory Study of Empirical Abstracts (Ph. D. diss., Syracuse University, 1988).

40 TextWise. *Project: DR-LINK*, Página WWW en Internet, URL: <<http://www.textwise.com/drlink.html>> (versión actualizada al 5 de julio de 1998).

41 TextWise. *Project: Chess*, Página WWW en Internet, URL: <<http://www.textwise.com/drlink.html>> (versión actualizada al 5 de julio de 1998).

serie de vínculos a temas relacionados con aspectos principales del tema. *CHES* es un sistema automático de construcción de bases de conocimiento de áreas interrelacionadas y un sistema de consulta.⁴²

La avanzada tecnología del Procesamiento de Lenguaje Natural, no sólo proporciona una mejor identificación del contenido de la colección de textos, sino también ofrece una oportunidad significativa para proporcionar un acceso y estructura de valor agregado a todos y cada uno de los objetos digitales, recursos electrónicos notables y trabajos monográficos de texto completo.

KNOW-IT

Un nuevo sistema integrado de organización y extracción del conocimiento está siendo desarrollado por Liddy y *TextWis*tm, el cual llevará a *DR-LINK* y a *CHES* a un nivel de funcionalidad más alto. Como sus creadores lo previeron, el sistema *KNOW-IT* (*KNOW*ledge base *IN*formation *T*ools) permitirá, tanto el uso del texto en bruto como de las bases de conocimiento estructurado como elementos de entrada para cualquier área temática y la creación de una base combinada de conocimiento a partir de ellos, por medio de la automatización del proceso de estructuración del conocimiento y su accesibilidad.

KNOW-IT resolverá los problemas de los grandes tiempos invertidos en la organización y en el uso de los elementos de entrada manuales. Debido a que la extracción del conocimiento y la creación de ontologías, o estructuras jerárquicas, son automatizadas, las bases de conocimiento serán construidas rápidamente, haciendo al sistema de utilidad para resolver problemas en tiempo real.⁴³

KNOW-IT será un sistema modular que incorpore herramientas para la extracción automatizada del conocimiento, la organización y estructuración, el mantenimiento y el acceso... Aunque los tres módulos pueden emplearse por separado con otras aplicaciones, estarán integrados en *KNOW-IT* como una herramienta sencilla de utilizar por usuarios no especializados de PC... Al integrar la extracción de conocimiento en forma automatizada inteligente con la capacidad de construir automáticamente jerarquías estructuradas de conceptos u 'ontologías' del conocimiento extraído, creará una 'ontología de referencia', basada en fuentes existentes de conocimiento para un campo... la ontología de referencia puede ser editada y cambiada empleando una herramienta complementaria que representa gráficamente los patrones de conocimiento dentro del espacio de éste, siendo posible

42 *Ibid.*

43 TextWise. *Project: KNOW-IT (KNOWledge base Information Tools)*, Página WWW en Internet, URL: <<http://www.textwise.com/drlink.html>> (versión actualizada al 5 de julio de 1998).

la reunión de tomadores de decisión... para agregar, corregir, o enriquecer la jerarquía de un modo colaborativo... *KNOW-IT* operará sobre un texto sin depurar, tanto a nivel de conocimiento de carácter general, como de conocimiento de un área.⁴⁴

Ontologías

Como ha sido definida por Sowa, (citado en Vickery, 1997),⁴⁵ una “ontología es un catálogo de todas las cosas que integran un [posible] mundo, cómo están estructuradas y cómo funcionan... [Es] un catálogo de conceptos y tipos de relación.”⁴⁶ En el campo de la ingeniería del conocimiento, una ontología puede definirse como “una especificación explícita de una conceptualización..., [e]sto es, una ontología es una descripción (como una especificación formal de un programa) de los conceptos y relaciones que pueden existir para un agente o una comunidad de agentes.”⁴⁷ Entre los proyectos más notables de ontología está el *CYC®*, una gran máquina de inferencia y bases de conocimiento multicontextual desarrollada por *Cycorp, Inc.*⁴⁸

El objetivo del proyecto *CYC®* es romper con el “software considerado cuello de botella quebradizo” de una vez y para siempre al construir las bases de un conocimiento básico de “sentido común” - un substrato semántico de términos, reglas y relaciones — que permitirá una variedad de productos y servicios de conocimiento intensivo. *CYC®* es propuesto para proporcionar una capa “profunda” de entendimiento que pueda ser usado por otros programas para hacerlos más flexibles.⁴⁹

En el campo de la documentación, Tom Gruber ha desarrollado una ontología de *datos bibliográficos* que “definen los términos empleados para describir referencias bibliográficas.”⁵⁰ Esta teoría define:

44 *Ibid.*

45 B. C. Vickery, “Ontologies.” *Journal of Information Science* 23, no. 4 (1997): 278.

46 *Ibid.*

47 Tom Gruber, *What is an ontology*, Página WWW en Internet, URL:

<<http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>> (versión actualizada al 5 de julio de 1998).

48 Vickery, *op. cit.*, p. 279.

49 Cycorp, Inc. *The CYC Technology*, Página WWW en Internet, URL:

<<http://www.cyc.com/tech.html>> (versión actualizada al 5 de julio de 1998).

50 Thomas R. Gruber, *Theory Bibliographic-Data*, Página WWW en Internet, URL:

<<http://ksl.stanford.edu/knowledge-sharing/ontologies/html/bibliographic-data/index.html>> (versión actualizada al 5 de julio de 1998).

la clase básica para los objetos de referencia y los tipos (clases) para los objetos de los datos que aparecen en las referencias, tales como autores y títulos. Las bases de datos específicas emplearán una *esquemización* que asocian las referencias con varias combinaciones de los objetos de datos, generalmente como campos en un registro. Esta ontología está propuesta para proporcionar los tipos básicos desde los cuales un esquema específico de base de datos podría ser definido.⁵¹

Recientemente, Weinstein ha desarrollado una ontología para los datos de registros *MARC*,⁵² conjuntamente con el proyecto de la Biblioteca Digital de la Universidad de Michigan.⁵³ En este proyecto, una ontología formal de relaciones bibliográficas fue creada y subsecuentemente, fue generada una base de conocimiento de los metadatos de aproximadamente 500 registros *USMARC* sobre trabajos de y acerca de Beethoven, permitiéndole a los usuarios revisar “familias de trabajos al navegar a través de las relaciones entre los trabajos,” empleando una interfase en Java.⁵⁴

Para los individuos interesados en la exploración del tema de ontologías de manera más profunda, Franconi⁵⁵ ha preparado un excelente centro de intercambio de información sobre “proyectos, gentes, conferencias y recursos específicos sobre ontología y campos relacionados.”

PROYECTOS DE DIGITALIZACIÓN

Desde los primeros días de la era de la computación, ha existido el interés en la creación de versiones digitales de libros y otros trabajos impresos. Uno de los primeros esfuerzos para digitalizar monografías y otros tipos de literatura y hacer colecciones electrónicas ampliamente disponibles es el proyecto *Gutenberg*. Iniciado en 1971 por Michael Hart, el propósito del proyecto *Gutenberg* es “hacer accesibles

51 *Ibid.*

52 Peter C. Weinstein, “Ontology-Based Metadata: Transforming the MARC Legacy,” En *Proceedings of the Third ACM Conference on Digital Libraries, June 23-26, 1998, Pittsburgh, PA* (New York: Association for Computing Machinery, 1998).

53 University of Michigan. Digital Library Initiative. *The NSF/DARPA/NASA University of Michigan Digital Library Project*, Página WWW en Internet, URL: <<http://www.si.umich.edu/UMDL/>> (versión actualizada al 5 de julio de 1998).

54 Peter Weinstein, *Ontology-based Metadata*, Página WWW en Internet, URL: <<http://www.personal.umich.edu/~peterw/Ontology/Beethoven/demo.html>> (versión actualizada al 5 de julio de 1998).

55 Enrico Franconi, *Ontology!*, Página WWW en Internet, URL: <<http://mnemosyne.itc.it:1024/ontology.html>> (versión actualizada al 5 de julio de 1998).

la información, libros y otros recursos al público en general, en forma que una vasta mayoría de las computadoras, programas y la gente, pueda leer, usar, citar y buscar de una forma fácil.”⁵⁶ Recientemente, *UMI* ha anunciado su *Digital Vault Initiative*, un esfuerzo para digitalizar su archivo de 5.5 billones de páginas contiendo documentos que datan desde los primeros textos impresos en inglés hace más de 500 años.⁵⁷ Estos no son sino sólo dos de los numerosos proyectos que van en aumento, que buscan convertir trabajos impresos a formatos electrónicos.^{58 59 60}

Un desarrollo igualmente significativo es el crecimiento en el número de trabajos de texto completo que están siendo presentados electrónicamente por los editores que participan en el programa de Catalogación Electrónica en la Publicación (*ECIP*) de la Biblioteca del Congreso. Una etapa adicional que ha sido planeada dentro de este programa es la utilización del sistema *ConText Option*, por la Corporación *Oracle*, para generar “un resumen global de un trabajo dado, al igual que sumarios de los temas secundarios incluidos dentro del trabajo,”⁶¹ lo que indudablemente tendrá un efecto profundo en cómo otras bibliotecas indizarán, catalogarán y administrarán los recursos digitales en el no muy lejano futuro.

Como producto del creciente interés en convertir o publicar, ya sea documentos o colecciones enteras en formatos digitales, y del incremento de la investigación sobre la representación formal de los sistemas de conocimiento como ontologías, uno puede esperar nuevas oportunidades para la aplicación de la sofisticada tecnologías del Procesamiento de Lenguaje Natural. El PLN no sólo facilitará un conocimiento superior del contenido de los recursos digitales, sino también proporcionará, simultáneamente, un nuevo esquema para representar estos objetos digitales, mejorando su acceso y navegación.

56 Michael Hart, *What is Project Gutenberg?*, Página WWW en Internet, URL:

<<http://www.promo.net/pg/history.html>> (versión actualizada al 5 de julio de 1998).

57 *UMI, Digital Vault Initiative : 500 Years of Information*, Página WWW en Internet, URL:

<<http://www.umi.com/hp/Features/DVault/>> (versión actualizada al 5 de julio de 1998).

58 International Federation of Library Associations and Institutions, *Digital Libraries: Resources and Projects*, Página WWW en Internet, URL: <<http://www.nlc-bnc.ca/ifla/II/diglib.htm>> (versión actualizada al 5 de julio de 1998).

59 Library of Congress. *American Memory: Historical Collections for the National Digital Library*, Página WWW en Internet, URL: <<http://lcweb2.loc.gov/>> (versión actualizada al 5 de julio de 1998).

60 Toronto Reference Library. *Toronto Reference Library Digitization Project*, Página WWW en Internet, URL: <<http://www.mtrl.toronto.on.ca/centres/digit/index.html>> (versión actualizada al 5 de julio de 1998).

61 Lynn El-Hoshy, “Report to ALA ALCTS CCS Subject Cataloging Committee (SAC) on Library Subject Cataloging, June 28, 1998,” Fotocopia.

SNOWCAP(sm)

Una opción para mejorar el acceso y uso de los recursos digitales es aplicar las tecnologías PLN para caracterizar el contenido y naturaleza de los recursos digitales. Empleando los sistemas disponibles actualmente, o en desarrollo, del *Text-Wisem*, al igual que una ontología bibliográfica apropiada, uno puede prever la caracterización de un objeto digital individual como una serie de clases estructuradas con miembros múltiples dentro de una clase. A un nivel básico, las tecnologías PLN podrían identificar, nombres, títulos, lugares y materias. Un objeto digital dado, por ejemplo, un libro digital, podría contener muchos nombres individuales, lugares, títulos y materias. Todos los distintos nombres podrían ser colocados en la clase “Nombre”, los distintos títulos en la clase “Título”, los lugares distintos en la clase “Lugar”, y las distintas materias en la clase “Materia”.

Cada miembro podría ser incorporado dentro de la clase de acuerdo a su rango con una indicación del porcentaje relativo de ocurrencia dentro del corpus del trabajo digital. Uno podría imaginar a los agentes de software inteligente, jugando un papel primordial en el manejo y organización de los miembros dentro de las clases, al igual que en el de las clases mismas.^{62 63 64} Todas las clases y sus miembros respectivos podrían entonces estar “envueltos” como una unidad única y ser colocadas en un “*Repositorio*” separado. Al objeto digital fuente le podría ser asignado automáticamente un identificador único, el cual podría estar asociado con el objeto. Este identificador único, también podría ser agregado como un apéndice a la unidad envuelta que constituye la esencia de ese objeto. El objeto digital podría, de este modo, tener una liga a un agrupamiento de clases representando el objeto y el agrupamiento podría tener una liga correspondiente. Mientras más objetos digitales fuesen procesados por los sistemas PLN, su esencia podría ser colocada en el repositorio y con el transcurso del tiempo, una colección de esencias podría ser creada. Como una mención descriptiva, uno podría considerar la esencia de cualquier objeto digital como un Paquete de Resumen Navegable Ontológico-Envuelto, Conceptual y de Contenido (Summary Navigable Ontologically-Wrapped Conceptual Aboutness Package, o Snow-Cap(sm)).

62 *LibraryAgents: Library Applications of Intelligent Software Agents*, Página WWW en Internet, URL: <<http://www.public.iastate.edu/~CYBERSTACKS/Agents.htm>> (versión actualizada al 5 de julio de 1998).

63 Gerry McKiernan, “ABCD: Agent-Based Collection Development,” *Technicalities*. En prensa.

64 Gerry McKiernan, “SerialAgents(sm): The use of Intelligent Software Agents to Identify, Organize and Manage E-Serials.” En preparación.

Con relaciones directas a los objetos fuente, tal repositorio podría ser visto tanto como la colección y el catálogo de la colección que representa. Los objetos digitales por si mismos, podrían existir en cualquier lugar — en un servidor local, en un servidor de una editorial, o en cualquier lugar en el ciberespacio. Dentro del repositorio, uno puede aplicar un procesamiento similar de PLN, u otro procesamiento semántico sofisticado para construir la ontología de este catálogo. Tal ontología podría constituir “La Ontología” de los objetos digitales “catalogados” y representar todas las relaciones entre los objetos dentro del repositorio.

Una modificación de la funcionalidad *centroide*, tal como se encontró dentro de la arquitectura del sistema *ROADS* (versión 2), podría probar su utilidad para implementar el modelo *SnowCaps(sm)*, como fue definido por quienes desarrollaron este software,

[Un] *centroide* puede ser pensado como un sumario del índice de una base de datos - incluyendo la información que normalmente es usada para encontrar que registros dentro de la base de datos corresponden a una solicitud dada, pero excluyendo los apuntadores para los mismos registros presentes.⁶⁵

En lugar de apuntar a los registros individuales, un registro de información de las bases de datos seleccionadas es mantenido en un “servidor índice”, el cual refiere a los usuarios a las bases de datos que contienen fuentes relevantes relacionadas con la solicitud.⁶⁶ Dentro de nuestro sistema previsto, las esencias de los objetos digitales podrían ser almacenadas en un servidor similar. En tal configuración, las búsquedas podrían realizarse dentro de este servidor indizador y los objetos digitales completos, o secciones relevantes, podrían ser recuperadas con base en la selección hecha por el usuario.

Búsqueda

Uno podría imaginar una interfase de búsqueda PLN que podría traducir un enunciado en lenguaje natural a una solicitud analizada sintácticamente, la que podría ser ejecutada contra el repositorio *SnowCap(sm)*, o sus ontologías relacionadas, para identificar aquellos objetos que mejor concuerden con el significado explícito, además del implícito de una solicitud. Uno podría imaginar la navegación y el despliegue de la ontología para este repositorio, empleando la

65 *ROADS Cross Searching*, Página WWW en Internet, URL:

<<http://www.ilrt.bris.ac.uk/roads/cross/>> (versión actualizada al 5 de julio de 1998).

66 *Ibid.*

tecnología de visualización de información. Indudablemente, la visualización podría ser

Una de las herramientas primarias para manejar y usar una ontología. La visualización permite a los usuarios ver los datos desde perspectivas múltiples para que ellos puedan identificar las relaciones de tales datos que les ayudarán a encontrar lo que están buscando. Una versión visual de la ontología podría permitir al usuario seguir visualmente un concepto hasta sus vecinos más próximos, o analizar el espacio total de conceptos interesantes, relacionados o no.⁶⁷

Uno también podría considerar el complementar la visualización con un despliegue auditivo adjunto para proporcionar pistas auditivas apropiadas para el usuario. En tal caso, ambas opciones podrían ser complementadas con un dispositivo apropiado de interfase háptico, tal como el *FEEL-it Mouse*, para permitirle al usuario navegar físicamente en el espacio de información de la ontología.

LA COLECCIÓN ES EL CATÁLOGO/ EL CATÁLOGO ES LA COLECCIÓN

Con la digitalización generalizada de las colecciones de los documentos, ejemplificada en el proyecto *Digital Vault Initiative* y en otros desarrollos, además de la adopción de las tecnologías interactivas y de los despliegues avanzados tales como la visualización de información, la revisión auditiva y los dispositivos interactivos hápticos, veremos un cambio fundamental en la forma en que la información es hecha accesible y utilizada dentro de las áreas experimental y comercial. Asimismo, el refinamiento de tecnologías como el Procesamiento Natural del Lenguaje y otros desarrollos, así como la implementación de las ontologías basadas en computadora en muchas más bases de conocimiento, también cambiará substancialmente las formas en que los objetos digitales son accedidos y organizados. Sin embargo, estos avances no serán confinados a los laboratorios o ambientes corporativos. Dentro de los próximos años, seremos testigos de un cambio igualmente fundamental en la manera en que las colecciones impresas y electrónicas son catalogadas y manejadas por las bibliotecas. Tales cambios sustanciales tendrán un efecto significativo en la profesión bibliotecaria y redefinirán el papel de los bibliotecarios en el siglo XXI.

67 Daniel E. O'Leary, "Using AI in Knowledge Management: Knowledge Bases and Ontologies," *IEEE Intelligent Systems* 13, no. 3 (May/June 1998): 34-39.

Acceso temático en línea

SANDY ROE

Biblioteca Karl E. Mund,
Universidad Estatal de Dakota
roes@columbia.dsu.edu

INTRODUCCIÓN

He planeado compartir con ustedes un poco de mi experiencia con los recursos electrónicos y los metadatos. Hablaré acerca de algunos resultados de las investigaciones en la profesión sobre los problemas comunes del acceso temático y sus soluciones. Por último, les daré una introducción a unos cuantos proyectos que han aplicado tales soluciones para el acceso temático de los recursos electrónicos. El día de ayer, Judith nos explicó el concepto de metadatos, en términos del contenedor y el contenido. Yo no hablaré del contenedor (i. e. MARC, TEI, Núcleo de Dublín), sino del contenido (clasificación, temas, palabras clave) y sus estructuras.

Antes de llegar a la bibliotecología trabajé en el campo de la edición. Ahí elaboré índices de autor y de materia y asigné términos de un vocabulario controlado utilizando mi primer tesauro. Ayudé a poner en línea revistas que previamente solo estaban disponibles en forma impresa y realicé una gran cantidad de búsquedas temáticas en la “Web.” Aprendí a preparar metadatos - al principio, simplemente agregando etiquetas para la descripción y las palabras clave en el encabezado de documentos HTML y más tarde, analizando gramaticalmente textos etiquetados con SGML para el proyecto de la Biblioteca Digital (*Digital Library Initiative*) de la Universidad de Illinois en Urbana, Champaign. En este proyecto se tomó el texto completo, codificado con SGML, de varios editores de revistas sobre ingeniería para crear una base de prueba de artículos. En ambos casos, uno de los propósitos de los metadatos fue facilitar la recuperación en línea.

El asunto de la organización y acceso a los recursos electrónicos estuvo presente en cada una de estas situaciones. Cada una de ellas contenía aspectos que no satisfacían mis expectativas. Cuando trabajé en el lado de la creación de información, necesitaba más conocimientos sobre cómo identificar las combinaciones más útiles de

los términos controlados y no controlados incluidos en los metadatos para que los documentos que yo codificaba pudiesen ser encontrados, aún cuando poco o nada acerca del documento fuese conocido por quién efectuaba las búsquedas. Cuando trabajé en el lado de las búsquedas, necesité ayuda para la formulación de los términos de búsqueda para mis consultas. ¿Qué términos debería yo escoger para garantizar que estaba buscando la información más relevante para mi tema? ¿Cómo puede el acceso, particularmente el acceso temático, ser mejorado?

Sospeché que los bibliotecarios conocían algo acerca de esto. Tuve una curiosidad muy intensa para descubrir que era aquel algo y también la extrema buena fortuna de tomar tres semestres de mi formación con la Profesora Pauline A. Cochrane durante mi estancia en la escuela de bibliotecología, lo que me ayudó en mi búsqueda de respuestas. Lo que sigue es un breve resumen de algunas de las respuestas que encontré.

RESUMEN DE INVESTIGACIONES ANTERIORES

Casi 100 años de esfuerzos se concentraron sobre la catalogación descriptiva para ayudar a los usuarios a diferenciar trabajos similares o ediciones diferentes del mismo trabajo. El trabajo de autoridades de nombres permitió diferenciar los trabajos de diferentes autores con el mismo nombre, o colocar juntos los trabajos de un autor que escribe bajo diferentes nombres. Ambos aspectos, la catalogación descriptiva y el trabajo de autoridad de nombres, ayudan a los usuarios a encontrar un material cuando conocen el autor o título. Esto es denominado como una búsqueda por materiales conocidos.

Este tipo de búsquedas, durante mucho tiempo han sido aceptadas como el tipo más frecuente que llevan a cabo los usuarios. Las investigaciones sobre el uso del catálogo de tarjetas apoyan esta opinión, indicando que las búsquedas por materiales conocidos eran hechas más frecuentemente que las búsquedas por materia. En consecuencia, el análisis temático no fue enfatizado, aún cuando los objetivos planteados por Cutter para el catálogo de la biblioteca incluían estos tres aspectos (la capacidad para encontrar un material cuyo autor, título o materia fuese conocido).¹

Hasta 1983, se creía que la búsqueda por materiales conocidos era la que prevalecía. Entonces, los resultados de un estudio llevado a cabo por el Consejo de Recursos Bibliotecarios (*Council on Library Resources*) reportaron un volumen mayor de

1 Cutter, Charles Ammi. *Rules for a dictionary catalog*, 4th. Ed., rewritten. Washintong, DC: GPO 1904, 12.

búsquedas por materia en comparación con las búsquedas por materiales conocidos. El reporte de este estudio también mencionó que los usuarios estaban teniendo problemas al efectuar sus búsquedas de materia en los catálogos en línea. Los usuarios deseaban mayores capacidades y flexibilidad en las búsquedas por materia.² Ambos descubrimientos —el predominio de las búsquedas por materia y los problemas que los usuarios estaban teniendo con este tipo de búsquedas— eran completamente inesperados. Consecuentemente, los estudios experimentales sobre los catálogos en línea empezaron a incluir pruebas sobre las diversas características de las búsquedas por materia, los estudios anteriores recibieron una atención renovada y el personal de la biblioteca solicitó mejoras a los catálogos en línea.³ Una década de investigación sobre el uso de catálogo en línea, particularmente en el área del acceso temático, dio inicio.

Se han identificado algunos problemas comunes relacionados con las búsquedas por materia (primero en los OPACs, continuando con los catálogos de las bibliotecas colocados en la “Web” y ciertamente, ahora en las búsquedas generales en la “Web”). Estos son:

- * La no recuperación de materiales⁴
- * Demasiados materiales recuperados⁵
- * Recuperación de solamente materiales no relacionados
- * Recuperación de algunas pero no de todas o las mejores fuentes⁶
- * Cuando el usuario cambia de una base de datos a otra o de un sitio “web” a otro, el vocabulario utilizado para expresar sus búsquedas es diferente, dando lugar al efecto de la “Torre de Babel.”

Las causas y consecuencias de estos problemas comunes incluyen:

- * No se recupera ningún material como resultado de las equivocaciones en el vocabulario utilizado para efectuar las búsquedas por materia. Los estudios sobre la conducta de búsqueda de información demuestran que la gente llega al catálogo con información incompleta para cualquiera de los puntos de

2 Lancaster, F. Wilfred y Amy J. Warner. *Information Retrieval Today*. Arlington, VA: Information Resources Press, 1993.

3 Besant, Larry. “Users of public online catalogs want sophisticated subject access.” *American Libraries* 14 (Marzo, 1982), 160.

4 Un tercio de las búsquedas por tema que los usuarios efectuaron en los catálogos en línea no obtuvieron ningún resultado (Lynch, C.A. Large database and multiple database problems in online catalogs. En *OPACs and beyond*. Dublín, OH: OCLC, 1989)

5 Las búsquedas que produjeron un gran número de registros recuperados desalentaron a los usuarios para revisar los resultados (Lynch, 1989).

6 Hildreth, Charles R. *Intelligent Interfaces and Retrieval Methods for Subject Searching in Bibliographic Retrieval Systems*. Dublin, OH: OCLC, 1989.

acceso.⁷ Los usuarios desean la posibilidad de poder ver una lista de palabras relacionadas con sus palabras de búsqueda.⁸

- * Cuando los materiales recuperados son demasiados, los usuarios se desalientan tanto que no examinan, para nada, los materiales recuperados.
- * Cuando sólo se recuperan artículos no relacionados.

El sistema podría no tener un control de homónimos (*rings* - ¿geométricos o gimnásticos?, *cold* - ¿la temperatura o el resfriado común?, *bear* - ¿el verbo o el sustantivo?).

Frecuentemente, los usuarios a menudo utilizan términos más amplios o más específicos que los correspondientes a los temas que en realidad desean.⁹ Podrían no encontrar ninguna información sobre “*poodles*” al utilizar como término de búsqueda “*dogs*” y asumir que el catálogo no contiene ninguna información sobre el tópico “*podles*”.

- * A menudo, el usuario no encontrará todas o las mejores fuentes porque tiene dificultad para relacionar la información incluida en un material encontrado con los sistemas de clasificación o con los vocabularios controlados, los cuales podrían permitirles ver términos y materiales relacionados.¹⁰ Ellos necesitan ayuda para encontrar más registros como el encontrado. En ocasiones, nuestros catálogos en línea contienen encabezamientos de materia de diferentes vocabularios controlados (LSCH, MeSH, Sears, etc.), aunque sólo incluyen la estructura (términos generales, específicos o relacionados) de uno de ellos, generalmente el LCSH. Si el material encontrado más relevante contiene solamente encabezamientos de materia de uno de los vocabularios, al usuario no se le proporciona ninguna dirección para establecer relaciones con términos de su preferencia, ya sean términos generales, específicos o relacionados. Peor aún, existen sistemas en donde no se incluyen del todo, ninguna estructura de encabezamientos de materia o de referencias cruzadas. Esto evita, efectivamente,

7 Borgamn, Christine L. “Why Are Online Catalogs Still Hard to Use”? *JASIS* 47 (7) 493-503, 1996.

8 Markey, Karen. *Subject Searching in Library Catalogs: Before and after the introduction of Online Catalogs*, OCLC Library, Information, and Computer Science Series 4. Dublin, OH: OCLC, 1984.

9 Hildreth, 1989.

10 Hildreth, 1989; Cochrane Pauline A. Bibliographic Data Bases-Their Effect on User Interface Design in Interactive Retrieval Systems. En *Redesign of Catalogs and Index for Improved Online Subject Access: Selected Papers of Pauline A. Cochrane*, Phoenix, AZ: Oryx Press, 1985.

que el usuario de los sistemas en línea encuentre las relaciones entre los encajados de materia seleccionados.

- * No existen buenos mecanismos que permitan una correspondencia o mapeo entre un vocabulario controlado y otro, por ejemplo de LCSH a Sears, de MeSH a LCC, etc. Quizás usted esté completamente satisfecho en su institución con un tipo de clasificación y se encuentra en una biblioteca o en un catálogo en línea con otra diferente. La notación de la clasificación que usted tiene se convierte en un callejón sin salida.

Las soluciones para los problemas comunes de las búsquedas por materia son:

- * Los sistemas deberán conducir al usuario durante el proceso de búsqueda a los términos preferidos a través de referencias de “véase” y a los términos relacionados con referencias de “véase también”.^{11, 12} Las experiencias de los usuarios que buscan en los OPACs demuestran que ellos se benefician con las referencias cruzadas que los conducen a materiales pertinentes. No debe esperarse que el usuario abandone su búsqueda para ir a un tesoro y después regresar al proceso de búsqueda. Los usuarios mencionaron que llevaban a cabo la tarea de buscar por términos generales o específicos.¹³ Las ayudas en línea para la búsqueda, la revisión de términos y la selección del vocabulario controlado son un imperativo.¹⁴
- * Necesitan ser desarrollados mecanismos de interacción para que la clasificación y los tesauros puedan proporcionar un mapa de rutas de búsqueda óptimas, de términos generales a términos específicos en varias bases de datos.^{15, 16}
- * La búsqueda por medio de la revisión de términos adyacentes al término ingresado por el usuario es la opción más útil y preferida cuando, la búsqueda no es específica (relacionada, por ejemplo, con la disciplina o tópico, el tipo de publicación, el nivel de tratamiento, la perspectiva, etc.), los resultados deseados no son conocidos por adelantado con precisión, los términos co-

11 Hildreth, Charles R. *Online public access catalogs: The user interface*. Dublin OH: OCLC, 1982.

12 Cochrane, Pauline A. Correspondencia por correo electrónico, Julio 13, 1998.

13 Karen, Markey, 1984.

14 Karen, Markey, 1984.

15 Cochrane, Pauline A. Correspondencia por correo electrónico, Julio 13, 1998.

16 Cochrane, Pauline A. & Johnson, Eric H. “Visual Dewey: DDC in a Hypertextual Browser for the Library User.” En *Knowledge Organization and Change*. Proceedings of the Fourth International ISKO Conference, 15-18 July 1996, Frankfurt/Main, Indeks, Verlag.

rectos para representar la consulta (la cual podría ser muy vaga) no son conocidos desde el principio.

Una o más de estas circunstancias podrían estar presentes en la mayoría de las actividades de las búsquedas por materia.¹⁷

En resumen, los usuarios necesitan el auxilio de un vocabulario en línea dentro de los OPACs y otros sistemas de recuperación. Para lograr esto, por lo menos tres cosas son necesarias.

- * Las estructuras necesarias para el control del vocabulario deben existir, permitiendo el control de los homónimos, sinónimos y las variantes morfológicas u ortográficas de las palabras y estableciendo los vínculos entre los términos, siendo éstos de términos generales a términos específicos, o entre términos relacionados. Cualquier sistema de indización automática como las listas de co-ocurrencia deberá de estar vinculado con los mecanismos de vocabulario controlado.
- * Las estructuras de vocabularios controlados deben de estar disponibles para el usuario desde el principio y a través de su proceso de búsqueda.
- * Más sistemas de correspondencia o de intercambio de vocabulario (lo que Clifford Lynch llama “una visión semántica común” deben estar disponibles para auxiliarnos en la búsqueda dentro del creciente número de bases de datos o repositorios de información diseminadas en la “red”, las que son diferentes y mantenidas de manera individual, como si esas bases de datos fueran un solo recurso de información consistente y unificado.¹⁸

Tres proyectos en donde se han aplicado estas soluciones para resolver los problemas actuales de las búsquedas por materia son mencionados a continuación.

EL INSTITUTO GETTY

Los vocabularios estructurados del Instituto Getty son el Tesoro de Arte y Arquitectura (*Art & Architecture Thesaurus, AAT*), la Lista Colectiva de Nombres de Artistas (*Union List of Artist Names, ULAN*) y el Tesoro Getty de Nombres Geográficos (*Getty Thesaurus of Geographical Names, TGN*). Mostraré algunos ejemplos empleando el Tesoro de Arte y Arquitectura y el proyecto de la base de datos a.k.a.²⁰

17 Hildreth, 1989.

18 Lynch, Clifford. “The Dublin Core Descriptive Metadatos Program: Strategic Implications for Libraries and Networked Information Access.” *ARL Newsletter*, Febrero 1998.

20 www.ahip.getty.edu/aka/

La base de datos a.k.a del Instituto Getty es una base de datos virtual integrada por diferentes bases de datos, las cuales contienen vocabularios diferentes. El objetivo de este proyecto es probar la validez de enriquecer los índices de recuperación temática de la base a través de la utilización de los recursos para el control del vocabulario del Instituto (AAT, ULAN, TGN).

Al buscar en una base de datos Getty dentro de a.k.a., empleando el Tesoro de Arte y Arquitectura (AAT), el módulo automático funciona de la siguiente forma. Si se encuentra una correspondencia exacta para su término o términos de búsqueda en AAT, ese término, junto con sus sinónimos y variantes, será empleado para buscar en la base de datos que usted ha seleccionado. También se incluyen aspectos multilingües. Actualmente, la cantidad de términos es de 120,000.²⁰

Con el propósito de demostrar lo anterior, efectué una búsqueda *sobre* “*railway bridges*”. Seleccioné una base de datos y utilicé la opción de vocabulario ampliado del Tesoro de Arte y Arquitectura (módulo automático). Primeramente, el proceso de búsqueda tomó mi término, “*railway bridges*,” y lo buscó en el Tesoro de Arte y Arquitectura, obteniendo los sinónimos y variantes. Estos términos aparecieron arriba de los resultados de mi búsqueda. Cinco variantes y mi término original fueron utilizados para buscar en el Índice Avery de Publicaciones Periódicas de Arquitectura (*Avery Index to Architectural Periodicals*). Las variantes empleadas fueron:

- * *railroad bridge* (forma en singular)
- * *bridges, railway* (forma invertida)
- * *bridges, railroad* (forma invertida en plural)
- * *railroad viaducts* (otra forma)
- * *viaducts, railroad* (otra forma invertida)

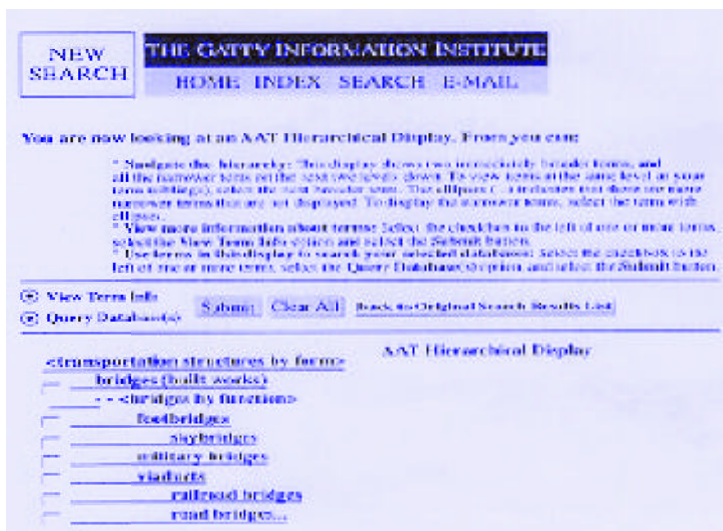
Los resultados de mi búsqueda fueron colocados a continuación y ordenados por orden de relevancia. El título de la primera cita listada fue, “*Wisconsin River railroad bridges: a historic resource*.” A ésta le fue dado un peso mayor que a la segunda cita porque además de abordar el tema, el título contiene una correspondencia exacta de la frase de búsqueda que yo utilicé.

Para ampliar o restringir mi búsqueda, pude regresar a la pantalla y cambiar la opción de vocabulario ampliado que había seleccionado al módulo de revisión del Tesoro de Arte y Arquitectura. Ahora, el mismo término de búsqueda me condujo al tesoro y me mostró un desplegado jerárquico facetado del ATT— los términos generales y específicos para el término que yo había seleccionado (*Figura 1*). En este caso, AAT no incluía términos más específicos para “*railroad bridges*”. Sin embargo, pude

20 Austin, David L. *ALA LITA/ALCTS Programa “Metadatos and Authority Control.”* Washington, DC, Junio 28, 1998.

ver que mis términos estaban incluidos en una categoría más general, “viaducts”, la que a su vez caía en otras más generales, “bridges by function” y “bridges (built works).”

Figura 1.



SISTEMA DE LENGUAJE MÉDICO UNIFICADO (UMLS)

Iniciado en 1986, el Sistema de Lenguaje Médico Unificado (*UMLS, Unified Medical Language System*)²¹ se desarrolló para superar los problemas de recuperación ocasionados por las diferencias en la terminología y la dispersión de información relevantes en las bases de datos. UMLS consiste de cuatro piezas o fuentes de conocimiento—el Metatesauro UMLS, el lexicón SPECIALIST, la Red Semántica UMLS y el Mapa de Fuentes de Información UMLS.

El Metatesauro es el componente central del vocabulario y actualmente incluye aproximadamente 40 vocabularios biomédicos y clasificaciones, estableciendo las relaciones

21 www.nlm.nih.gov/pubs/factsheets/umls.html

entre las numerosas y diferentes formas de nombrar los mismos conceptos. Está organizado por conceptos o significados y su propósito es relacionar las diferentes alternativas y facetas del mismo concepto e identificar las relaciones útiles entre conceptos diferentes. Preserva los significados, las conexiones jerárquicas y otras relaciones entre los términos incluidos en los vocabularios fuente, y establece nuevas relaciones entre conceptos y términos de otros vocabularios fuentes diferentes, creando un sistema de conversión o mapeo entre vocabularios.

La Biblioteca Nacional de Medicina (*National Library of Medicine, NLM*) y otras instituciones están aplicando el UMLS en una gran variedad de maneras. Una de las aplicaciones propias de la NLM incluye el sitio *Internet Grateful Med*®, donde el Metatesauro es considerado el más significativo de todos sus recursos. La NLM compara el Metatesauro con un “Piedra Roseta electrónica.” En 1996 (su décimo aniversario), este Metatesauro incluía más de 589,000 nombres de aproximadamente 253,000 conceptos en biomedicina.²² La versión 1998 del Metatesauro incluye traducciones al español, francés, alemán, portugués y ruso del MeSH (La Lista de Encabezamientos de Materia de la NLM).

IODyne

Eric Johnson de la UTUC ha desarrollado un navegador de tesauros hipertextual llamado IODyne.^{23, 24} Esta herramienta proporciona un despliegue interactivo de sugerencias de términos, empleando tesauros temáticos, listas de co-ocurrencia, palabras alternadas dentro del contexto y despliegues de palabras clave. IODyne permite un mecanismo de arrastre y vaciado desde cualquier campo del documento recuperado y regresar al tesauro, la palabra clave en el contexto, el espacio del concepto (lista de co-ocurrencia), el registro del A&I (Abstract e Índice), o la ventana con la clasificación. El desplegado del tesauro es hipertextual, permitiéndolo abrir la parte de la jerarquía que usted desee ver y manteniendo otras porciones plegadas o cerradas.

En la *Figura 2*, la ventana de la parte superior izquierda muestra un despliegue del Tesauro INSPEC de palabras clave en el contexto, producto de una búsqueda sobre el tópico “*databases*.” Esta sugiere al usuario el término “*deductive databases*,” quien lo teclea en la siguiente ventana de la derecha titulada “*Concept Space*” y abre el Tesauro INSPEC (ventana inferior derecha) marcando “*deductive databases*” y la jerarquía que lo circunda. La ventana “*Concept Space*” es una lista de co-ocurrencia, construida de los registros del A&I

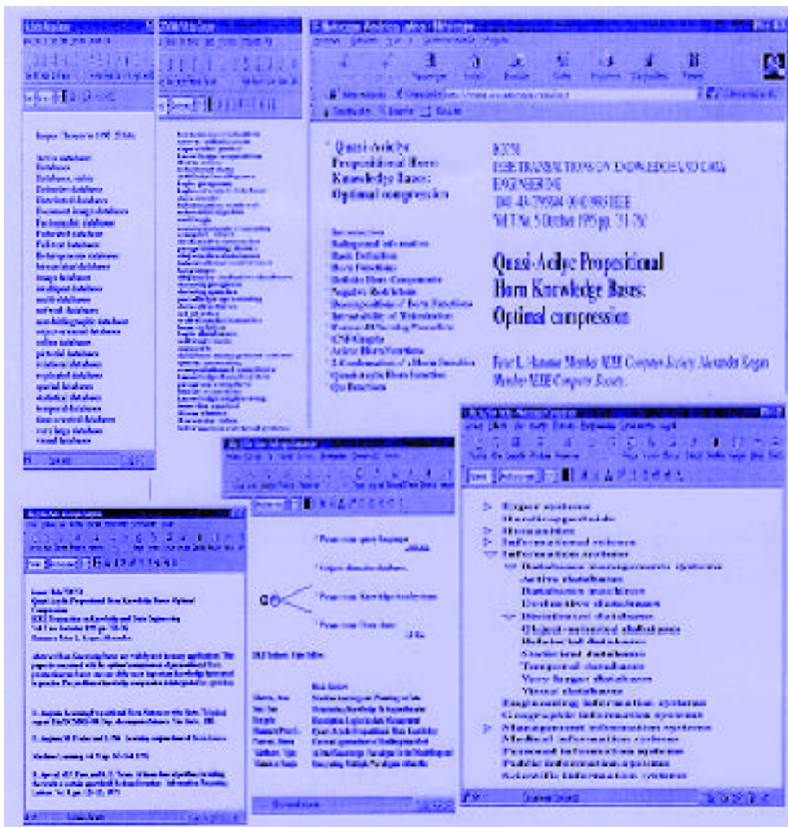
22 www.nlm.nih.gov/pubs/factsheets/igm.html

23 Cochrane and Johnson, 1996.

24 Johnson, Eric H. Extending an Interactive Thesaurus by Dragging. *ACM SIGLINK Newsletter*, September, 1995.

INSPEC, la cual nos muestra los términos que ocurren con mayor frecuencia dentro de “deductive databases.” La ventana central de la parte inferior de la pantalla contiene el documento buscado. En esta también se encuentran desplegadas todas las búsquedas previas del usuario (“query lenguajes”, “deductive databases”) incluyendo una búsqueda Booleana sobre “knowledge-based systems” y “Horn clause.” En la parte inferior se despliega el conjunto de artículos recuperados. Al seleccionar el artículo “Quasi-Acyclic Propositional Horn Knowledge” de Peter L. Hammer, se abrirá el registro A&I (parte inferior izquierda) y (o) el texto completo del artículo (parte superior derecha).

Figura 2.



Mientras estuve en la UICU, tuve la oportunidad de ver la demostración de IODyne y usarlo yo misma con el Tesauro INSPEC cargado. También ha sido exitosamente probado con el Vocabulario del Índice Legislativo (*Legislative Index Vocabulary, LIV*) del Servicio de Investigación de la Biblioteca del Congreso (*Library of Congress/Congressional Research Service*) y con la clasificación CDD.

CONCLUSIONES

Al nivel de materiales, continúe la asignación de vocabularios controlados – sistemas de clasificación, encabezamientos de materia —. Aumente sus registros con términos del material para apoyar la búsqueda por palabras clave y la creación de mecanismos de ayuda como las listas de co-ocurrencia o las de palabras clave en el contexto.

Los fundamentos de investigación que nos proporcionan la dirección existen. Las tecnologías que existen hacen la aplicación de esa investigación más fácil que en cualquier otra época. Los principios continúan siendo los mismos, ya sea que la colección esté en un OPAC, o en un catálogo de biblioteca en la “Web”, o bien en un conjunto de documentos electrónicos como los del Proyecto de la Biblioteca Digital (*Digital Library Initiative*), o en una colección virtual que se encuentra en diferentes localizaciones geográficas en la “*World Wide Web*”, accesada a través de un motor de búsqueda como el del Instituto Getty, o *Grateful Med®* de la NLM. Entonces, piense más allá de lo común para proporcionar acceso a su colección local (independientemente si ésta es una base de datos, un OPAC, o una colección digital) e involúcrese en las maneras de cómo aplicar esta información o cómo relacionarla con otras que están en la posición de aplicarlas.

Reflexiones en torno a la taxonomía documental para organizar el acceso a la información documental

RAMIRO LAFUENTE LÓPEZ

Centro Universitario de Investigaciones

Bibliotecológicas de la UNAM

lafuente@servidor.unam.mx

Los documentos digitales en línea no son un conjunto aislado de textos automatizados independientes del contexto social en que se presentan. Son producto de una compleja interrelación entre la tecnología, el desarrollo económico social y nuevas propuestas para cualificar, organizar y dividir el trabajo en la sociedad de nuestro tiempo. En la estructura de este tipo de documentos podemos encontrar elementos de significación indicativos del contexto social en el cual se generan, su estrecha relación con diversas formas de reorganización de las actividades productivas, de investigación y docencia, hacen necesario el plantearse la construcción de una taxonomía que permita comprender la naturaleza de estos documentos con el fin de diseñar sistemas de clasificación y organización para facilitar el acceso y difusión de los mismos.

La naturaleza de los documentos digitales no sólo encierra un conjunto de símbolos, mensajes y tecnologías, sino que forman parte de una vasta gama de fenómenos sociales y económicos que influyen y determinan tanto su forma como su funcionalidad, orientadas a lograr un óptimo uso y difusión de información para dar cumplimiento a finalidades empresariales, de investigación y de docencia. Para entender y explicar la funcionalidad de los documentos digitales, es fundamental crear esquemas de referencia que nos faciliten desarrollar conceptos que nos sirvan para aclarar las razones por las cuales se generan y utilizan en determinados ámbitos y círculos sociales.

Deliberadamente utilizamos el término documento digital para aludir a un fenómeno que encierra diversas características relacionadas con prácticas sociales acerca del uso de la tecnología de la información y las redes de telecomunicación, como medio para registrar, estructurar y difundir información. Es importante depurar la terminología, dado que si utilizamos la idea de lo electrónico, no es factible reconocer las propiedades únicas de los documentos digitales, que de-

rivan justamente del uso de las tecnologías digitales, que determinan la naturaleza de la forma para el registro, transmisión y proceso de información.

Para estudiar y comprender la naturaleza de los documentos digitales, sería de utilidad una taxonomía que estableciera marcos de referencia para entender y explicar los tópicos involucrados en estos asuntos. Desafortunadamente, las tecnologías digitales se están desarrollando más rápidamente que nuestra habilidad para crear relaciones taxonómicas relevantes para definir los fenómenos relacionados con la creación y aprovechamiento de registros digitales para estructurar documentos. El componer relaciones taxonómicas implica no solamente el análisis lógico y semántico para producirlas, también es necesario generar un consenso en torno a ellas, de tal suerte que reflejen determinados significados para una comunidad epistémica específica.

Una primera forma de aproximarnos al estudio de los documentos digitales, con la finalidad de derivar principios taxonómicos que nos permitan clasificarlos, sería el repasar las características de su interfaz y funcionamiento, dado que estas representan la única forma que se tiene de acceder a ellos. La interfaz para acceder, consultar y recuperar información de un documento digital, obedece al propósito de dar cumplimiento a un objetivo primario: *guiar al público acerca de los contenidos temáticos* del documento. Como la interfaz puede adoptar distintas formas y cumplir diversas funciones, es necesario crear principios que permitan organizarlas conforme a criterios orientados a evaluar su coherencia y significado en relación a considerar si la forma de funcionamiento de la interfaz del documento realmente constituya una manera adecuada, no sólo para recorrer los contenidos temáticos del documento digital, sino para transmitir una idea general de su contenido y desempeño.

Para diseñar y construir interfaces que faciliten el establecer recorridos acerca de los contenidos temáticos de un documento digital, es indispensable fundamentarla en principios, métodos y fines, que sirvan como punto de partida para mostrar ordenadamente, un contenido temático por medio de la pantalla del monitor de una computadora. El diseño de la interfaz del documento digital debe responder a los fines que se pretenda cumplir con la misma, y tener determinadas características lógicas y formales, es decir, es ineludible que responda a un método que facilite tanto su construcción como su utilización por parte del público en general.

Para lograr un cuerpo unificado de conocimiento acerca de los documentos digitales, es necesario analizar tanto las partes de que se compone como las formas de uso del mismo por parte de distintas comunidades y en distintos momentos, para desarrollar métodos de análisis que nos conduzcan a comprender su naturaleza. El estudio de las formas sobre cómo determinadas comunidades se organizan para producir, difundir, y usar documentos, tiene un carácter relevante porque de

ella deriva la estructura lógica del documento digital, así por ejemplo, un documento construido a partir de la tecnología del hipertexto, puede adoptar diversas formas según las finalidades a cumplir. Si se desea es que al usarlo exista plena libertad de recorrerlo, su estructura de ligas y relaciones tiene que ser lo suficientemente flexible para permitirlo, pero si se trata de cumplir fines docentes, entonces la estructura tendrá que responder a la necesidad de guiar para cumplir objetivos de aprendizaje en detrimento de la prerrogativa de examinarlo como se desee.

Los fenómenos inherentes a la producción, difusión y uso de documentos, abarcan una amplia gama de asuntos vinculados a factores económicos, educativos, académicos, políticos, antropológicos, etc. Fenómenos que para efectos de originar una taxonomía los podemos caracterizar para de sus características derivar tipologías de documentos digitales en las cuales sustentar métodos orientados a la formación de colecciones de estos. Las categorías establecidas para cada tipo de documento digital proporcionará los elementos necesarios para construir los métodos adecuados para seleccionar, describir, clasificar y administrar una colección de documentos pertenecientes a esa categoría.

Una forma de resolver los problemas relativos al diseño, construcción y administración de colecciones de documentos digitales radica en pensar que las tecnologías de la información definen otras reglas del juego para la producción y manejo de documentos, al cambiar la manera de organizarlos, transmitirlos y difundirlos. Con ello abre nuevas posibilidades, y por tanto nuevos problemas, y aunque no obligatoriamente hace más fácil el manejo y organización de documentos, idealmente el uso de las tecnologías de la información y las redes de telecomunicación generan otras opciones para la disposición y sistematización de la información.

El registrar información por medio de los dispositivos de una máquina digital, no significa simplemente una nueva forma de almacenar y procesar los tradicionales registros de textos impresos, puesto que los documentos digitales son esencialmente distintos, pues las redes de telecomunicación ofrecen la posibilidad de usarlos de maneras diferente a los impresos. El documento digital tiene un más holgado y flexible manejo de sus contenidos. En razón de esta flexibilidad, antes de decidir como organizar colecciones de documentos digitales, deben especificarse los propósitos a los que pueden servir, y los modos para cumplirlos. Los principios del diseño de colecciones de documentos digitales es deseable se conciben en función de la probabilidad de correspondencia de los contenidos de un documento con los de otro a través de la aplicación de esquemas de clasificación que los relacionen.

La colección y organización de documentos digitales únicamente adquiere sentido en el contexto de necesidades y conductas de los miembros de una comuni-

dad. El proceso intelectual de interpretación de la forma y/o los contenidos temáticos de los documentos, para representarlos por medio de códigos o palabras con la finalidad de facilitar su organización y/o localización, obtiene importancia cuando está enfocado al cumplimiento de una intencionalidad que se expresa a través de algún tipo de producto o servicio. Al manifestar las finalidades de un acervo a través de algún tipo de servicio, se dota a éste de significados relevantes para los miembros de una comunidad, al hacer patente el medio que el público tiene para beneficiarse de la organización de los documentos que se logra por medio de su análisis y clasificación.

Como los servicios, es necesario expresarlos por medio de sistemas de control administrativo, de instrumentos de acceso a la colección, de representaciones documentales y de instrumentos para disponer de la misma. Si el proceso organizador de los servicios exclusivamente se sitúa como instrumento de búsqueda de la eficiencia y se margina el hacer explícitas las intenciones, significados y finalidades de la representación de documentos, indudablemente estaremos frente a la presencia de una maquinaria administrativa que puede ser impresionantemente eficaz para resolver problemas estándar para el almacenamiento y recuperación de información, pero que poco tendrá para ofrecer al individuo en busca de información para generar conocimientos aplicables a la solución de problemas específicos.

La creación de colecciones de documentos digitales involucra el preguntarse acerca de la naturaleza del diálogo que se presenta o debe presentarse con la comunidad para la cual se analizan y clasifican los documentos, asunto que evidentemente está relacionado con aspectos inherentes a la lógica de la construcción de sistemas de clasificación y a la determinación de significados y relevancia que se le otorguen a los contenidos de los documentos. Pero sobre todo, implica establecer las cualidades que deben reunir los servicios por medio de los cuales se pretende responder a las necesidades de un mundo que induce al uso de la información en forma rápida y precisa, pero al mismo tiempo requiere de elementos y espacios para que el individuo pueda generar e integrar conocimientos que le permitan no sólo explicarse lo que sucede en el mundo, sino que contribuyan a enriquecer su concepción individual del mismo.

En las últimas décadas ha prevalecido la idea de sustentar el análisis y representación de documentos en la concepción del almacenamiento y recuperación de la información, cuya intencionalidad esencial radica en la construcción de técnicas que sean funcionales para hacer eficiente el proceso de búsqueda de información. De esta manera, se hace hincapié en un servicio destinado a proveer de información al usuario, y la organización de colecciones de documentos así como los instrumentos para el acceso a los mismos se dirigen a la localización de información y datos espe-

cíficos, relegando a un segundo plano el establecer y mostrar las relaciones que existen entre diferentes documentos.

La funcionalidad y procesos de captura y conversión de los documentos digitales, los convierte en volátiles. Algunos de ellos tienden a estar en constante transformación, como es el caso del web. Ante esta situación es conveniente impulsar una taxonomía que facilite comprender la naturaleza de este tipo de documentos y por tanto permita generar principios generales para cualificar el acceso y los espacios temáticos de éstos, en donde la modificación constante de los contenidos informativos son parte de su razón de ser. La práctica de fijar los textos que contienen los temas tratados en los documentos, derivada del uso de la cultura sustentada en la imprenta, no es esencialmente provechosa para comprender y explicar el fenómeno social que envuelve la creación y uso de los documentos digitales.

¿Cómo podemos comprender las propiedades de los documentos digitales? podríamos empezar por nombrarlos y describirlos con acciones verbales más que con sustantivos y adjetivos y averiguar las acciones y causas que motivan su presencia en el seno de una comunidad. Puesto que la terminología sustentada en sustantivos lo aísla del contexto social en que se produce y utiliza, por ejemplo cd-rom, nombra una tecnología que no denota las relaciones sociales, las prácticas y las formas de organización del trabajo que implica cuando se utiliza como un instrumento para el registro de diversos tipos de documentos, con finalidades tan diferentes como puede ser el archivar los documentos contables de una empresa, o el difundir los resultados de una investigación.

Las relaciones y prácticas sociales que propician la aparición y difusión de documentos digitales pueden ser mejor representadas por verbos, que indiquen el tipo de proceso que se trata de denotar, esto puede ayudar a llamar la atención acerca de los fenómenos sociales implicados en la creación y uso de documentos digitales. Esto significaría que la simple idea de una taxonomía que únicamente refleje el orden del fenómeno, sería inapropiada al no enfatizar las relaciones con las actividades que motivan al empleo de un documento digital. Es fundamental dar una mayor importancia al contexto para significar al documento digital.

Las reglas para construir una taxonomía de los documentos digitales puede sustentarse en los siguientes principios:

- * Es necesario diseñar un esquema de clasificación que organice los conceptos disponibles, por medio de categorías, que nos sirvan a los propósitos de formar grupos de conceptos bajo un solo nombre con objeto de identificar la redundancia de términos.
- * Las categorías que se diseñen deben arreglarse conforme a un orden primario, preferido. Este orden preferido debe reflejar las relaciones signifi-

cativas entre las características y funciones que motivan la creación y uso de documentos digitales, o la subordinación de un concepto a otro. Se entiende la jerarquía como una forma de ordenar por rangos de algún tipo, en donde el tipo superior incluya a los que se encuentren bajo él en la escala, únicamente se pretende la colocación de conceptos de clase en orden lógico. Lo que se coloca en rangos distintos son entidades, niveles de organización o de significación.

La descripción de la forma, funcionalidad y contenidos temáticos del documento digital implica cuestiones relacionadas con el análisis de términos y conceptos, pero además conlleva implícito un análisis que requiere la comprensión del uso y objetivos a cumplir por este tipo de documentos. Esta comprensión será factible, a través del conocimiento y clarificación de las funciones que cumplen en el ámbito de una comunidad, y del establecimiento de términos para nombrar determinada categoría o tipo de documento por medio de palabras o frases cuyos significados sean comprensibles, por lo que un buen principio, sería la elaboración y publicación de diccionarios digitales de tipología documental regularmente revisados y actualizados.

De acuerdo con lo hasta aquí señalado, puede verse cómo los aspectos morfológico, semántico y pragmático guardan estrecha relación al momento de realizar una taxonomía de los documentos digitales. Generalmente existe una tradición en cuanto al uso de términos de frecuente utilización sin corroborar que significan o si corresponden a ideas o conceptos ya nombrados por medio de otros términos.

Tradicionalmente, las bibliotecas han sido un medio de creación y organización documental en donde el bibliotecario dispone y clasifica acervos con objeto de facilitar la disponibilidad de los documentos, desbrozando los documentos por medio de esquemas y políticas de selección y clasificación documental. De esta manera, suministra toda la información relevante, en razón del orden de clasificación creado, que ofrece los medios para comprender la naturaleza de la organización documental, y por tanto, el tipo de documentos que se encuentran reunidos. Todo ello preñado del subjetivismo propio de una actividad de selección, no obstante estar sujeta a políticas y objetivos de organización y servicios documentales, que desde luego no deben implicar la censura. Eso nunca ha sido malo y distingue la competencia y cualidades de la actividad profesional de los bibliotecarios. En otro sentido, debemos distinguirla de la libertad de información, puesto que esta no está en mantener la asepsia informativa, sino en ofrecer al lector los medios para poder escoger.

¿Cuál es la alternativa en el uso de los documentos digitales? Diversas campañas reivindican y reclaman la libertad de expresión en Internet, cuando quizá estén defendiendo el postulado neoliberal: “toda la información al alcance del lector, sin

intermediarios”. Esta idea que parece tan progresista, inocente, democrática y avanzada, oculta e incluye la falsa, manipulada, tendenciosa, y aparentemente veraz *opinión* camuflada de suceso, la *opinión* fundada sobre hechos falsos, la mala interpretación de los hechos, el rumor.

Hay un grupo profesional que libera al lector de vivir en un marasmo de información. Hace tiempo que sabemos que mucha información no es buena información. Y, desde hace siglos los profesionales de la organización bibliotecaria, han estado intentando informar verazmente por el único medio posible: dedicar tiempo y esfuerzo en acumular toda la información, filtrarla, clasificarla y diseñar servicios para facilitar su libre disponibilidad. La *revolución de la información en Internet* ha sido retrotraer al lector a sucumbir ante una montaña de información. La libertad del lector frente a un soporte digital y multidireccional como Internet, ya no consiste únicamente en disponer de múltiples fuentes de información y poder discriminarlas con conocimiento de causa, sino que la libertad más amplia del lector se consigue cuando éste también puede aspirar a publicar contenidos, a estrenarse en esta actividad que ha sido perseguida sistemáticamente por el poder del momento en la mayor parte de momentos de la historia de la humanidad.

Bibliotecas y tecnologías: una propuesta de integración

ALVARO QUIJANO SOLÍS

Biblioteca Daniel Cosío Villegas,

El Colegio de México

quijano@colmex.mx

INTRODUCCIÓN

El 11 de mayo de este año, en la Gaceta UNAM, el rector de nuestra casa de estudios dio a conocer su Plan de Desarrollo 1997-2000. El análisis del entorno realizado para este plan destaca las variables que se han vuelto lugar común en la literatura sobre el futuro de las universidades: globalización, competencia, tele-educación, etc. El análisis apunta una visión de futuro que cristaliza en una misión para la Universidad Nacional reflejando las aspiraciones de esta institución en el umbral del nuevo milenio.

Podemos o no compartir la visión de futuro del planeador, tal vez hasta podamos disentir acerca de las líneas programáticas planteadas como puente hacia ese futuro, pero no podremos dejar de reconocer que el modelo educativo planteado por la UNAM es, por decir lo menos, ambicioso. Dos citas textuales pueden ilustrar este punto:

“...deben ser capaces de *aprender*, no sólo durante los años que asisten a la escuela, sino *a lo largo de toda su vida activa*...”

En: *Las perspectivas para el próximo siglo*. El papel de las nuevas generaciones universitarias. UNAM: Plan de Desarrollo 1997-2000

“...La pedagogía necesita orientarse hacia la promoción de las habilidades de razonamiento y cuestionamiento; al desarrollo de competencias para seleccionar, organizar y procesar sistemas diversos de información...”

En: *Los programas estratégicos*. La renovación de los procesos educativos. UNAM: Plan de Desarrollo 1997-2000.

De la primera de estas citas se desprende una concepción del proceso no solamente de aprender, sino del proceso mismo mediante el cual se aprende. Es decir, necesitamos entender y controlar el proceso de apropiación del conocimiento, necesitamos formar estudiantes que aprendan a aprender, única garantía de que repetirán este proceso conscientemente a lo largo de toda su vida, mas allá de las aulas.

En la segunda cita, más cercana a nuestros oídos profesionales, la implicación sobre las habilidades cognoscitivas es clara: los docentes debemos enseñar a razonar, cuestionar, seleccionar, organizar y procesar información.

El modelo esbozado se mueve en un entorno de alta tecnologización mundial que implica cambios dramáticos en la información, a velocidades nunca antes vistas.

En contraste, el modelo también se mueve en un país de profundas desigualdades pero, más importante, está destinado a una población que proviene de un modelo educativo tradicional, en donde el maestro y el libro de texto son las fuentes casi únicas de conocimiento, en donde no existe el hábito de la lectura, ni bibliotecas escolares que apoyen los tempranos procesos de enseñanza-aprendizaje y los de autosuficiencia en el uso de la información que permitan al alumno aprender por sí mismo a aprender para adaptarse a los cambios tecnológicos.¹

El modelo planteado confronta, pues, la ausencia en los estudiantes de las capacidades para reconocer sus necesidades de información, para localizarla, evaluarla y usarla.²

Más central a este coloquio, antes que señalar las carencias, es, sin duda, ubicar el papel que las bibliotecas universitarias jugarán en este modelo futurista que no es privativo de las aspiraciones de la UNAM, sino que se ha convertido en el *leit-motiv* de la reforma universitaria mexicana, lo mismo en el sector público que en el privado.

Parece, pues, que la biblioteca universitaria deberá buscar su propia transformación para participar más activamente en el proceso tardío de alfabetizar a sus usuarios en el uso de la información. Como consecuencia de la mencionada carencia de bibliotecas escolares, los usuarios que la biblioteca universitaria mexicana pueda formar no representarán, de tener éxito el proceso, más del 1% de la población del país; es decir, aquella fracción que tendrá acceso a la educación superior en los próximos años.

1 Mayor, J., A. Suengas y J. González Marqués. *Estrategias metacognitivas: aprender a aprender y aprender a pensar*. Madrid: Síntesis, 1995.

2 Varlejs, J. "Information literacy: just another buzzword." En: *Information literacy: learning how to learn*. Jefferson, N.C: MacFarland, c 1991, pp. 1-11.

Es necesario investigar mas sobre estos fenómenos de empobrecimiento informativo-educativo en un mundo que, de suyo, ya ha creado demasiados contrastes sociales, económicos y políticos. También es necesario investigar sobre los diferentes retos que, en este contexto, se presentan ante la biblioteca universitaria y que *grosso modo* pueden enunciarse de la siguiente manera:

- * La necesidad de detectar cuales son las habilidades informativas que necesita el personal profesional bibliotecario para enfrentar la creciente demanda de formar usuarios autosuficientes en el uso de información.³
- * La racionalización del proceso de aceptación tecnológica en función del apoyo que la tecnología preste al valor que el trabajo intelectual de los profesionales agrega a la información.⁴
- * Los procesos de reingeniería necesarios para adaptarse al cambiante entorno tecnológico y a la globalización informativa.⁵
- * La necesidad de una planeación prospectiva de los procesos y servicios bibliotecarios que permita el avance hacia una cultura de organización y administración de conocimiento, mas que hacia la prevaleciente cultura de organización y administración de información.⁶

El presente trabajo propone una primera aproximación a un modelo sistémico para enfocar el punto 3 ya mencionado, a partir de la experiencia que la Biblioteca Daniel Cosío Villegas de El Colegio de México ha tenido en la administración del cambio tecnológico en la última década.

Aunque todos los factores mencionados influyen en la visión que tenemos del futuro de la biblioteca universitaria, este documento se centrará en algunos de los aspectos organizacionales que ésta confronta de cara a la acelerada transformación tecnológica que ha experimentado en los últimos años.

3 - Quijano, A. y M. G. Vega Díaz. "Valor agregado y desarrollo de habilidades informativas: un enfoque sistémico." En: *XXXII Reunión Nacional de Bibliotecarios*. Buenos Aires: ABGRA, 1998.

- Quijano, A. y V. Herrero. "Informatización y perfil bibliotecario: el caso de la Biblioteca de El Colegio de México." En: *Conferencia Internacional de Bibliotecas de Educación Superior*. Santiago de Chile: Colegio de Bibliotecarios de Chile, 1996.

4 Quijano A. "Aceptación tecnológica y valor agregado: el caso de la biblioteca de El Colegio de México." En: *Quinto Congreso Nacional de Bibliotecología*. Medellín, Colombia: ASCOLBI, 1998.

5 Lau, J. "La reingeniería administrativa y una experiencia sui generis." En: *Quinto Congreso Nacional de Bibliotecología*. Medellín, Colombia: ASCOLBI, 1998.

6 Tenkasi, R. y R. Boland. "Exploring knowledge diversity in knowledge intensive firms; a new role for information systems." En: *Journal of Systemic Knowledge Management*. January, 1998 (<http://www.free-press.com/journals/knowledge/issue1/article4.htm>)

EL IMPACTO DE LAS NUEVAS TECNOLOGÍAS EN LA BIBLIOTECA UNIVERSITARIA

Una constante reciente en el mundo de las bibliotecas académicas es la de escuchar acerca del impacto que las nuevas tecnologías están teniendo sobre la forma en que aquéllas venían ofreciendo sus servicios y realizando sus procesos. Los nuevos retos planteados van desde una amenaza de convertir estas bibliotecas en entes “virtuales”, hasta los mas realistas que presionan su transformación hacia un mayor aprovechamiento de los recursos electrónicos ahora disponibles. El lugar común en el que se ubica esta presión por una nueva biblioteca académica es, como el de muchas otras áreas de la actividad humana, la mayor globalización y la competencia por los recursos financieros.

Las bibliotecas, y no sólo las académicas, son organizaciones en donde la calidad de los productos que se generan está determinada por la aplicación intensiva del intelecto humano, particularmente al agregar valor a los diferentes insumos informativos y convertirlos en productos intelectuales diferenciados que contribuyen a la generación de conocimiento.

En el contexto descrito, los recursos tecnológicos juegan un papel importante al apoyar las labores bibliotecológicas intelectuales asociadas con los procesos de valor agregado, que están, a su vez, orientados a lograr la misión y los objetivos de las propias bibliotecas a las que apoyan.

De hecho, las bibliotecas académicas actualmente basan muchas de sus propuestas de servicios en las tecnologías de información existentes. Puede decirse, en suma, que existe una estrecha relación entre la capacidad intelectual bibliotecológica para resolver problemas y la asimilación de tecnología para viabilizar soluciones.

Como un ejemplo de esta interrelación, se mantiene una expectativa razonable de que las tecnologías de la información podrían contribuir a un mayor acercamiento entre las propuestas de ordenamiento temático diseñados por bibliotecarios académicos y la dinámica del lenguaje natural de sus demandantes o usuarios. En el logro de esta estructuración semántica, se manifiesta una creciente necesidad de un conocimiento temático cada vez más especializado e interdisciplinario entre los campos científicos y tecnológicos. En este contexto las tecnologías de la información facilitan este acercamiento a través de la manipulación de las bases de datos que se generan; pero ciertamente no son suficientes ya que el problema del ordenamiento temático requiere de un marco teórico adecuado en donde están en juego la diversidad de las representaciones semánticas así como los procesos cognoscitivos y los de validación académica. En suma, se requiere un mayor trabajo intelectual. Este ejemplo ilustra el por qué se hace necesario un proceso de

asimilación tecnológica para adaptar soluciones a los problemas planteados por la diversidad semántica y cognoscitiva de las comunidades académicas.⁷

Puede decirse que, sin limitarla, la creatividad intelectual de los bibliotecólogos para plantearse y resolver problemas relacionados con aumentar el valor y la pertinencia de sus productos, depende en un grado cada vez mayor de lo que las tecnologías de información pueden hacer para resolver esos problemas. Los mejores productos y servicios informativos, en nuestros días, tienen una alta componente tecnológica, asociada con una aguda capacidad profesional para utilizar racionalmente estos apoyos en el proceso de agregar valor a productos y servicios.

Desde la perspectiva de los sistemas abiertos, la tecnología ha venido a reforzar los mecanismos con los que las bibliotecas académicas buscan adaptarse a un entorno heterogéneo y cambiante, para “sobrevivir”. La mayor capacidad de adaptación tecnológica se convierte, así, en una ventaja competitiva al estabilizar el sistema ante los retos del medio ambiente.⁸

En otro orden de cosas, el cambio tecnológico se apoya en el elemento humano, de tal suerte que el centro contemporáneo de las adaptaciones a los retos planteados a las bibliotecas académicas por la modernidad, son los conocimientos y habilidades humanas para manipular la tecnología, particularmente aquella que es ahora tan común para los usuarios de este tipo de bibliotecas. Siguiendo este razonamiento, puede decirse que, junto a, y más importante que la inversión tecnológica, la inversión en el desarrollo de los conocimientos y habilidades para manejar información y tecnología constituye el principal mecanismo para aumentar la productividad y, por ende, la calidad de los servicios que las bibliotecas académicas proporcionan a sus usuarios.

Estos conocimientos y habilidades se vuelven particularmente importantes a medida que la biblioteca abandona su papel reactivo y asume como suyos los procesos mismos de enseñanza-aprendizaje y de investigación. Y el futuro no presenta otra

7 Tenkasi y Boland, (*op.cit.*), sostienen que la no solución de estos conflictos semánticos y terminológicos se refleja en la incapacidad de los sistemas de información para proveer información pertinente a los procesos de innovación tecnológica, en empresas de uso intensivo de conocimiento (farmacéutica, por ejemplo), en donde se requiere integrar diferentes campos de conocimientos a un producto específico.

8 Katz, D. y R.L. Kahn. *The social psychology of organizations*, (Chapter Z: “Common characteristics of open systems”). New York: Wiley, 1966.

alternativa en un contexto en que la tendencia en el uso directo de las fuentes de información apunta hacia la formación de usuarios cada vez más autosuficientes, si de verdad se pretende que las universidades formen generaciones capaces de aprender de por vida, tal como lo prevé el mencionado plan de desarrollo de la UNAM.

LA REINGENIERÍA DE PROCESOS EN LAS BIBLIOTECAS

En los noventas, el sector de las bibliotecas de universidades públicas mexicanas logró consolidar los apoyos que había venido otorgando el Gobierno Federal desde la década anterior y que permitieron extender a más universidades los efectos de la automatización que ya eran visibles en la Universidad Nacional Autónoma de México y en la Universidad de Colima. Puede decirse que las bibliotecas universitarias mexicanas del sector público constituyen el grupo pionero del sistema bibliotecario mexicano, desde el punto de vista del valor agregado a las tecnologías de información, como consecuencia de los apoyos señalados y de una mayor conciencia profesional para el empleo racional de estas tecnologías.

La irrupción de las tecnologías mencionadas, junto con la popularización de INTERNET, ha constituido un conjunto de retos a la administración de las bibliotecas mexicanas, toda vez que los cambios de los contextos, insumos, procesos y productos de la información, han sido por demás vertiginosos.

Podría decirse que también confluyen en este proceso otras tendencias “modernizantes” emanadas de los nuevos análisis que se hacen del entorno de los negocios. Así, el resurgimiento de la planeación estratégica parece proporcionar un modo de enfrentar el problema de la sobrevivencia en un medio cambiante, para afinar las ventajas competitivas de las organizaciones. En este ámbito, y frente a una incorporación intensiva de tecnología que propicia una mayor dependencia de las bibliotecas mexicanas hacia ésta, es cada vez más frecuente que debamos agregar a nuestro léxico profesional, términos como calidad total, mejora continua, valor agregado y “benchmarking,” entre nosotros.

Independientemente de la pertinencia coyuntural de estos términos, se puede afirmar que existe una mayor conciencia profesional, no solamente sobre la necesidad de racionalizar el empleo de las tecnologías de la información, sino sobre las condiciones necesarias para desarrollar nuevos conocimientos y habilidades del personal profesional de las bibliotecas académicas mexicanas, bajo el convencimiento de que la tecnología, por sí sola, no es capaz de producir

combinaciones inteligentes de los diferentes insumos que resulten en productos diferenciados con alto valor agregado.⁹

Las implicaciones de todo este entorno han generado, en el caso de la biblioteca de El Colegio, un suelo fértil para los procesos de reorganización administrativa, alrededor de una práctica profesional que experimenta cambios que, en nuestra opinión, pueden agruparse en seis aspectos:

- * Participación de personal profesional en la planeación, incluyendo una tendencia a la organización matricial entre departamentos y proyectos, y la toma de decisiones más colegiada, que engloba la autoplaneación de las actividades.
- * Administración de la calidad del control bibliográfico, particularmente en lo que se refiere a una mayor normatividad para la transferencia de información y el uso de lenguajes documentales más precisos.
- * Ajustes en los procesos de control administrativo para enfocar la calidad desde el punto de vista del valor agregado a la tecnología de información, a través de capacitación selectiva.
- * Incrementos en la calidad de procesos y servicios a través de una mejor calidad de vida en el trabajo, involucrando una mayor racionalización de las actividades y del tiempo dedicado a ellas.
- * Especialización temática del personal y multifuncionalidad del mismo para unificar los recursos informativos, independientemente de su formato, en colecciones desarrolladas y procesadas integralmente por personal calificado (esto es: bibliotecarios especialistas en un área que seleccionan, catalogan, clasifican y dan servicio).
- * Incorporación del usuario en la evaluación de los procesos y servicios existentes y en la creación de nuevos, incluyendo un desarrollo más eficaz de sus habilidades informativas.

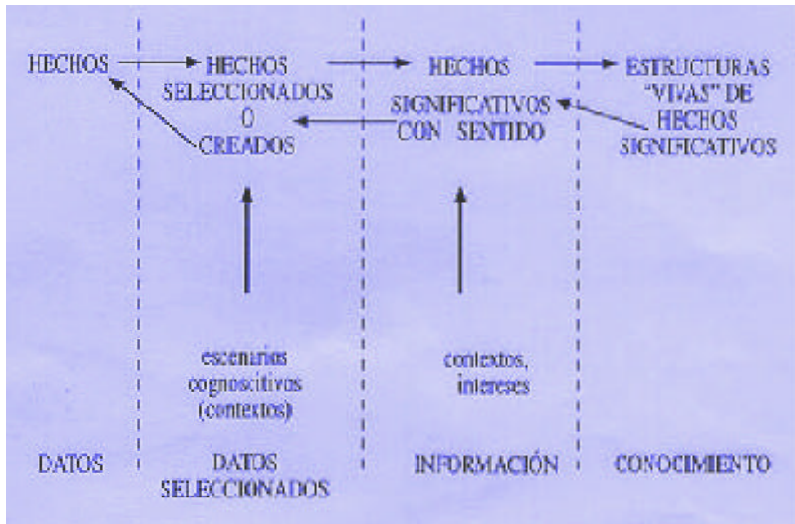
Estos esfuerzos reorganizativos parten del supuesto de que nuestra biblioteca constituye un sistema socio-técnico¹⁰ en donde el factor humano es el recurso más importante para integrarse con éxito a los procesos de investigación y de enseñanza-aprendizaje de El Colegio de México, en un entorno de “globalización” académica. Con estos esfuerzos, en nuestra hipótesis, estaríamos reaccionando al medio

9 Hace un par de años, por ejemplo, era vigente la controversia sobre si los bibliotecarios debían o no aprender instrumentos tales como HTML. Con la evolución de los productos informáticos hacia un manejo más simple y hacia una mayor popularización, resulta imprescindible para el bibliotecario el mayor entendimiento de estas herramientas.

10 Emery, F.E. y E.L. Trist. “Socio-technical systems.” En: *C.W. Churchman y M. Veshulst (eds.) Management science, models and techniques*, vol. 2. London: Pergamon, 1960.

ambiente para mantener un estado estable, corrigiendo “disfunciones” y obligándonos a adaptarnos a los cambios impuestos desde afuera, estableciendo mecanismos autopoyéticos que propicien, por ejemplo, la aceptación de la tecnología, con el fin de regular los insumos y productos que se reciben y se entregan a nuestro medio ambiente, transformados. El fin último de estos procesos es el de agregar valor a la transmisión de información transformable en conocimientos (*Figura 1*), a partir del desarrollo de conocimientos y habilidades en bibliotecarios y usuarios.

Figura 1. Proceso de Valor Agregado



La biblioteca académica, de esta manera, requiere “sobrevivir” al papel meramente reactivo ante las demandas de sus usuarios y asumir un protagonismo dinámico que actúe proactivamente buscando adelantarse a aquellas, mediante la mayor creatividad en la investigación de instrumentos de apoyos a las labores académicas de la institución y en la promoción de los servicios creados para satisfacer la demanda futura, que generarán los sujetos de los nuevos procesos educativos. Esta creatividad deberá darse lo mismo en el WWW que en el mostrador de préstamo.

EL VALOR AGREGADO

El valor agregado está asociado con el tiempo profesional y con la tecnología dedicados a transformar materias primas en productos que se entregan al usuario final. En nuestro contexto, nos referimos al valor que el bibliotecario incorpora a los datos para transformarlos en conocimiento, con la ayuda creciente de una tecnología de información que es cada vez más abundante.

Aunque el concepto es originalmente económico, no podemos divorciarlo de otros factores intangibles como pertinencia, calidad, seguridad y confianza. Así, en la biblioteca académica el trabajo intelectual del bibliotecario debe buscar el crecimiento del valor de los productos, a través de la determinación de la naturaleza de la información, sus usuarios finales y sus posibles aplicaciones en los procesos de investigación y de enseñanza-aprendizaje.

Visto en esta perspectiva, el valor agregado a la información supone dos instancias sobre las cuales construir una estructura para dimensionarlo. La primera se refiere a los procesos bibliotecarios que van añadiendo una diferenciación en la información y engloban, en forma agregada, la selección que se hace en el mundo de la información disponible, la organización bibliográfica en bases de datos de la información seleccionada y la puesta en servicio de estas piezas de información localizables a través de un sistema que permita al usuario encontrarlas, independientemente del campo semántico que éste utilice. El valor de estos procesos se da independientemente de si la información es utilizada o no y, por tanto, es susceptible de contabilizarse y, eventualmente, de ubicarse en términos monetarios.

Una segunda instancia de valoración del valor agregado estaría, por supuesto, relacionada con la que establece el usuario, a partir de su propio entorno, y de sus propios filtros cognocitivos. Valorar lo que la información significa para un usuario supone paradigmas personales que resultan muy difíciles de ser previstos por quien desarrolla productos de información, un tema largamente tratado en la literatura y del que queda la sensación de que la satisfacción del usuario, desde el punto de vista de la pertinencia (o relevancia) de la información, es un factor marginal a la efectividad de las bibliotecas. En otras palabras, la disponibilidad de las colecciones y de productos con valor agregado, pero estandarizados para comunidades amplias, tienen un mayor impacto sobre la satisfacción del usuario que las precisiones en términos de pertinencia, utilidad visible, etc.

De cualquier manera, para lograr esta valoración, es evidente la falta de una axiología común a los investigadores. Mientras haya una tendencia hacia la efectividad, otros enfoques apuntan hacia una teoría del valor intrínseco.¹¹ Todo parece indicar que hace falta mucha investigación para determinar con precisión la satisfacción de un usuario en una biblioteca.

El enfoque que aquí se presenta, supone que el valor agregado se desprende de la cuantificación del tiempo profesional que se aplica a cada etapa de los procesos que van desde el desarrollo de colecciones hasta la creación de servicios, pasando por la organización del material. En este sentido, el valor agregado proviene de centros de costo internos de la biblioteca y, al ser medible en tiempo profesional, puede cuantificarse en dinero. Obviamente, a mayor especialización del personal, mayor será el valor agregado a los productos y, de ahí, la importancia de desarrollar las habilidades informativas.

El borrador de modelo que se propone para esquematizar el valor agregado y su relación con las habilidades informativas profesionales, se presenta en la (*Figura 2*) y parte de un enfoque sistémico¹² que adapta la metodología de los sistemas suaves¹³ y la metodología de la contabilización de valor agregado en sistemas de información aplicados a la toma de decisiones en las empresas.¹⁴ Sin embargo, el enfoque de este documento presenta un mayor detalle en la estructuración de los llamados dominios. Se plantea la existencia de cuatro: el humano, el profesional, el informático y el tecnológico. Los cuatro componentes descansan uno sobre el otro, y muchos de sus procesos integrantes son paralelos e interdependientes, de un componente a otro.

En el dominio humano, los procesos cognoscitivos referidos a la selección, evaluación, búsqueda, análisis, juicio, síntesis, etc. se hacen presente para sintetizar el proceso esquematizado en la *figura 1* y convertir los insumos primarios de hechos y datos en productos susceptibles de ser convertidos en conocimiento. Sobre este componente, el valor agregado se expresa en el tiempo que se asigna a las labores que van desde la selección hasta el servicio en las bibliotecas académicas.

11- Rappaport, A. *Creating shareholder value: the new standard for business performance*. New York: Free Press, 1968.

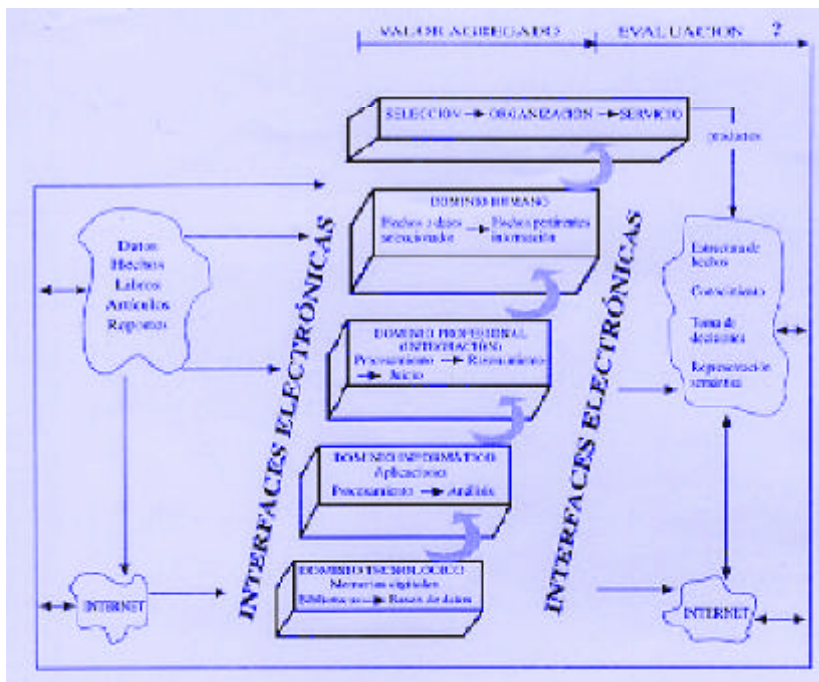
- Habayeb, A. R. *System effectiveness*. London: Pergamon, 1987.

12 Quijano, A. y M. G. Vega Díaz, *op. cit.*

13 Checkland, P. y S. Holwell. *Information, systems and information systems*. London: Wiley, 1998.

14 McPherson, P. K. "Accounting for the value of information." *En: ASLIB Proceedings*, 46(9), September 1974, pp. 203-215.

Figura 2. Modelo Sistemático del Valor Agregado



En el dominio profesional, los procesos cognoscitivos integran las aplicaciones informáticas para apoyar las labores que se ejecutan en el dominio humano. En esta instancia, nos interesan particularmente las habilidades cognoscitivas que permiten procesar información para emitir juicios de valor sobre ella.

En el dominio informático nos referimos básicamente a aquellas actividades humanas, propias de especialistas en el área de informática, cuya principal labor es utilizar y adaptar aplicaciones informáticas ya existentes.

En el dominio tecnológico, mas allá de toda actividad humana, el centro se establece en la utilización de memorias digitales para convertir colecciones de bibliotecas en bases de datos susceptibles de manejarse a través de aplicaciones informáticas y de integrarse profesionalmente al proceso de valor agregado bibliotecario.

Se resalta en la cabeza del modelo los procesos a los que se agrega valor a lo largo del sistema, y que pueden o no ser tangibles. En todo caso, estos procesos constituyen un todo integrado que deriva en los productos diseñados para el usuario. Se trata, en resumen, de realizar las operaciones semánticas a lo largo del proceso bibliotecológico, para transformar los datos en información que luego se convierte en conocimientos socialmente significativos y legitimados, como se había esquematizado en la *figura 1*.

Mediante la continua repetición de estos procesos es que se llega a conformar el cuerpo del conocimiento, mismo que requiere del manejo de mecanismos de comunicación de la información y de los conocimientos, para los cuales es necesaria la existencia de representaciones lingüísticas y documentales apropiadas y consensadas. Estas representaciones tienen, entre otras funciones, las de hacer perceptible el valor semántico e informativo de los textos, reducir los tiempos y la incertidumbre en la búsqueda de información, y organizar el cuerpo de conocimientos de manera lógica y pertinente.

La creación de representaciones documentales utilizando tecnologías de la información, exige la aplicación de diversas habilidades cognitivas. En el modelo presentado, estas habilidades son interdependientes entre los diferentes dominios.

Al implantarse una nueva tecnología, se modifican, eliminan o insertan actividades y aplicaciones en la organización, que pasan, de ser novedosas, a ser percibidas como rutinarias, conforme, se van incorporando a los procesos cotidianos. El impacto de las nuevas tecnologías se estabiliza, entre otros aspectos, por el conocimiento previo de los usuarios, por el entrenamiento, por la frecuencia de uso y la percepción de la utilidad. Hay un ciclo, que puede esquematizarse como la (*Figura 3*), y que se refiere a las diferentes instancias en el proceso del cambio tecnológico. En la (*Figura 4*), se presenta el modelo propuesto por Saga para precisar las interrelaciones existentes entre las diferentes variables asociadas con este proceso de cambio tecnológico.¹⁵ La variable “frecuencia de uso” es la que se propone como indicador cuantitativo de la aceptación de nuevas tecnologías introducidas en el dominio tecnológico del modelo propuesto y que es la clave para agregar valores valor en los tres dominios restantes, toda vez que es la expresión medible de las habilidades informativas adquiridas.

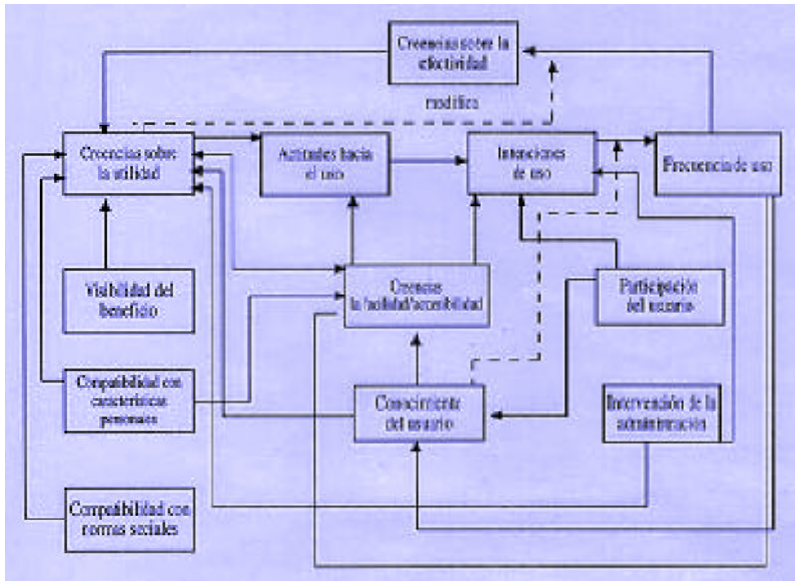
15 Saga, V. I. y R. W. Zmud. “The nature and determinants of IT acceptance, routinization, and infusion.” En: L. Levine (ed.), *Diffusion transfer and implementation of information technology*. Amsterdam: North Holland, 1994.

Figura 3. Proceso de Cambio Tecnológico



Las creencias sobre la utilidad y accesibilidad de la tecnología son otras de las variables relacionadas con el grado de conocimientos de los usuarios. Las expectativas sobre la utilidad de una tecnología están condicionadas, a su vez, por la percepción de los beneficios visibles, la compatibilidad con las características personales y con las normas sociales, y por la accesibilidad, definida como el grado en el cual una tecnología se percibe como fácil de usar y aprender. Estas creencias afectan o benefician las actitudes y las intenciones de uso de las tecnologías y, finalmente, se traducen en frecuencia de uso. Estas creencias pueden ser manipuladas favorablemente, hacia la aceptación de tecnología, si se establecen mecanismos para desarrollar las habilidades cognitivas dentro de las organizaciones base de la reingeniería que se necesitan para garantizar la “sobrevivencia.”

Figura 4. Modelo de Zaga Zmud para el Proceso de Aceptación de Tecnología



RECAPITULACIÓN

Es claro que el desarrollo de los conocimientos y habilidades de los bibliotecarios contribuirá a aceptar el cambio tecnológico, sin el cual difícilmente la biblioteca podría trascender hacia un mundo en donde ésta ya no es el único lugar en el que se deposita la información académica. Invertir en capacitación selectiva de personal bibliotecario para obtener un servicio más personalizado a los usuarios, reduciendo verificaciones y controles, constituye la mejor preparación para el futuro de la universidad, al tiempo que se aumenta el valor que se agrega a la información.

Puede decirse que existen algunos principios que nos proporcionan un marco de referencia para enfrentar los retos del entorno tecnológico y que nos permita observar y analizar los problemas en un contexto mas general, para buscar soluciones de una manera sistemática, y no meramente casuística. Estos principios podrían resumirse de la siguiente manera:

- ✧ El valor que se agrega a los datos, para transformarlos en conocimiento, puede ser analizado sistemáticamente.
- ✧ El valor que se agrega a la tecnología se incrementa en la medida en que es aceptada por el personal y por los usuarios, y esta aceptación es una función de otras variables que pueden ser entendidas y controladas, particularmente de la frecuencia de uso, resultado del desarrollo de habilidades informativas.
- ✧ El incremento en las habilidades informativas del personal y de los usuarios constituye así el objetivo principal de la integración tecnológica.

Estos principios podrían guiar la reingeniería de procesos que, como ya se dijo, parece ser inevitable para la sobrevivencia de la biblioteca académica, a condición de que la planeación del futuro busque modelarlo y alcanzarlo, no solamente adaptarse a los retos que nos presenta.


Los flujos de información electrónica y la educación del futuro

MARGARITA ALMADA DE ASCENCIO

Centro Universitario de Investigaciones

Bibliotecológicas de la UNAM

almada@cuib.unam.mx

ómo se presentará la oferta educativa en las próximas décadas? Seguramente los programas de educación a distancia, educación virtual y universidad virtual tendrán una participación muy importante en el escenario educativo.

Los flujos electrónicos de información y formas de acceso y recuperación pertinentes sobre insumos prioritarios que deben considerarse desde el diseño de todo programa de educación. El papel de los profesionistas investigadores/docentes en bibliotecología y ciencias de la información son y serán cada vez más, actores indispensables en los grupos multi- e interdisciplinarios de trabajo que diseña y desarrollan nuevos programas educativos acordes a las nuevas políticas y demandas de educación en todos los niveles y modalidades. Las teorías y aplicaciones surgidas de estos campos del conocimiento: bibliotecología, ciencia de la información, comunicación y educación y pedagogía deben agregar innovaciones a la organización del WEB y de los flujos de contenidos...

Koltay¹ indica que “...en la era de materiales digitalizados, multimedia y el “web” el enfoque de la biblioteca tradicional es ya anacrónico. Tener acervos de libros de texto, procesar artículos para reserva y ofrecer cursos sobre índices bibliográficos no satisfacen las necesidades y demandas del personal académico que quiere mejorar la enseñanza utilizando el poder de las tecnologías actuales”.

1 Koltay, Zsuzsa, Trelease, Ben, and Davis, Phillip M. (1996). Technologies for learning: instructional support at Cornell's Albert R. Mann Library. Library Hi Tech 14 (4) consecutive issue 56:83-98.

Como siempre, los “fatalistas” al enfrentar el advenimiento de nuevas tecnologías, parecen gozar en asustar a los colegas vaticinando el cercano fin de las profesiones actuales o de la ocupación profesional. Es verdad también, que la humanidad ha aprendido a lo largo de los siglos a adaptarse a cambios que afectan y transforman su vida cotidiana. Sin embargo estos complejos cambios no se efectúan de manera instantánea, sino a través de una evolución continua. Sin estar totalmente de acuerdo con Koltay, si concuerdo en que hay que estar alerta a cambios que deben efectuar las bibliotecas para satisfacer una oferta ampliada de posibilidades educativas que ya están surgiendo en los programas y proyectos de educación a distancia y virtual tanto para diferentes niveles educativos básicos y medios, como para la universidad virtual.

Las bibliotecas se enfrentan a un conjunto de opciones con recursos generalmente limitados y su reto es hacer “lo correcto y adecuado” para su entorno, pero teniendo presente los cambios mundiales. Cada biblioteca irá definiendo su siguiente etapa según las circunstancias. En la mayoría de los casos no serán cambios gigantescos con enormes riesgos, sino cambios continuos que se dan paso a paso. Así los riesgos son menores y los costos sufragables.

Los metadatos representan primeros cambios que pueden ser útiles al organizar y catalogar recursos de información en Internet. Las viejas formas de trabajar sólo son buenas en tanto sigan teniendo sentido; ello podría aplicarse a la catalogación.² Recordemos algunos de los cambios se han implantado en las últimas décadas y qué actividades dejaron de tener sentido debido a esos cambios, como por ejemplo, la impresión de tarjetas catalográficas.

Es indispensable que los profesionistas en bibliotecología e información se involucren en los proyectos y programas de los nuevos modelos educativos para ir preparando la manera en que la biblioteca puede mejor coadyuvar en dicho proceso. Las colecciones digitalizadas y la organización de los flujos electrónicos de información para su fácil integración, acceso y recuperación son aspectos muy importantes que, a pesar los avances tecnológicos, presentan múltiples escollos.

Más que nunca los especialistas en bibliotecología, documentación y ciencia de la información deben colaborar de manera interdisciplinaria con otros actores que participan o deberían participar en los procesos de educación a distancia y educación virtual. Hago énfasis en la palabra interdisciplinaria, porque cada uno de los participantes en un proyecto semejante tendrá que comprender la perspectiva de los especialistas de las otras disciplinas, con objeto de tener una visión integrada

2 Jul, Erick (1997). “Cataloging internet resources : survey and prospectus.” *Bulletin of the American Society for Information Science* 24 (1): 6-8

del problema y de las necesidades de los maestros e investigadores, los alumnos y los administradores.

La biblioteca virtual debe ser para la educación a distancia y la universidad virtual, la unidad de administración y organización de la información y el conocimiento de la institución. Las bibliotecas en la educación virtual serán centros de información y conocimiento que distribuyan los contenidos de información electrónica, las colecciones digitalizadas y los demás recursos informativos aprobados por los maestros, así como orientadores de otros sitios electrónicos relevantes a los temas de los cursos. Los bibliotecólogos distribuirán oportunamente, a través de los intranet y la internet, los flujos de información relevantes a los cursos que requieren tanto la comunidad presencial como la virtual.

Antes de contar con los materiales bibliográficos y educativos digitalizados que requieran un curso virtual, es indispensable formar a los docentes para acceder y recuperar información electrónica, incluyendo planes y guías de enseñanza para la educación a distancia y virtual. Un ejemplo de ello es el proyecto GEM-Gateway to Educational Materials- Puente de Acceso a Materiales Educativos- que actualmente se desarrolla en los Estados Unidos, que tiene como objetivo buscar un marco operativo útil para los maestros de educación básica de ese país, a través de un acceso sencillo a miles de planes de cursos, unidades curriculares y otros recursos educativos montados en internet. Tiene como objetivo mejorar la organización, acceso y recuperación de acervos importantes de este tipo de materiales que ya están disponibles en sitios en internet, de agencias federales, estatales, locales, comerciales, académicas, etc. pero que en general son de difícil o laborioso acceso.³

¿Cuál es la educación a distancia y la educación virtual que amplíe las posibilidades de una oferta educativa de calidad? Los planteamientos, ciertamente interesantes y seguramente aplicables, tienen múltiples problemas y riesgos que ya han surgido en muchos de los programas de educación a distancia que se ofrecen en la actualidad; entre ellos destaco:

- * Distanciamiento: se diluye la misión educadora del docente si no logra el contacto virtual e interacción con los alumnos y entre los alumnos.
- * No debe confundirse el papel de los maestros con el de los facilitadores, instructores y supervisores, que son personal de apoyo, aunque en muchas ocasiones el maestro tendrá que asumir también el papel del personal de apoyo cuando este no existe.

3 Sutton, Stuart A., Oh, Sam G. (1997). "GEM: using metadata to enhance internet retrieval by K-12 teachers." *Bulletin of the American Society for Information Science* (24) 1: 21-27.

- ✱ Cerrazón en lugar de apertura para acceder a la información. En vez de ampliarse las posibilidades de acceso y recuperación como podría pensarse al usar el internet, o al usar recursos informativos digitalizados, se restringe por el mismo caos actual de organización de información en internet o porque los cursos a distancia sólo cuentan con contenidos de información digitalizados proporcionados por el maestro y no se promueve ni se facilita al alumno distante a buscar información sobre el tema y otros tipos de recursos informativos y materiales disponibles en Internet o en colecciones digitalizadas de las bibliotecas ad hoc y disponibles a través de intranets de la institución oferente o en Internet.
- ✱ Las posibilidades de contar con la información pertinente y suficiente podría reducirse por la ausencia de recursos electrónicos disponibles o por falta de un control bibliográfico adecuado. Al no tener éxito en búsqueda, el alumno recurre sólo a lo que el maestro coloque en el web, Así se perdería la esencia del cambio educativo que busca que el alumno “aprenda a aprender” y a buscar información que le ayude a formarse un criterio para dar respuesta a sus dudas e inquietudes. La educación debe orientarse a que el alumno aprenda a resolver los problemas que se lo presentan, buscando diversas opciones, y por ello, diversas fuentes y recursos de información y de conocimiento registrado. Debe aprender a buscar, acceder y recuperar información, relacionada con su problema y encontrar opciones para su resolución.

Al desarrollar sistemas de información digitalizada para la universidad virtual, estamos obligados a reconocer al autor/usuario; es decir, que tanto el docente como el alumno son, a su vez, oferente y receptor de información. Los sistemas deben ser organizados para el acceso coherente, amplio, oportuno y especialmente pertinente para los participantes del proceso educativo.

Si aceptamos que el proceso de enseñanza/aprendizaje debe diseñarse para que cada individuo aproveche al máximo sus capacidades, el concepto de educación debe ampliarse en tiempo y espacio para que realmente sea *un proceso de vida, una educación permanente*.

El concepto de educación permanente es la construcción continua de la persona, de su saber y de sus aptitudes, de su facultad crítica, sus actitudes y su capacidad de actuar; de habilitarnos a desarrollar una conciencia crítica sobre nosotros mismos y fomentar nuestra plena participación en el trabajo y en la sociedad. De ahí la importancia de dirigir los planes educativos hacia una “sociedad educadora y educada” para el próximo siglo. Ello no se alcanzará si no se logra ampliar la oferta educativa de calidad y con equidad y con el apoyo en las llamadas nuevas tecnologías de información.

El concepto de educación continua y educación a lo largo de la vida tiene las ventajas de su flexibilidad, su diversidad y su disponibilidad para las personas, en tiempo y espacio. La educación continua, la educación a distancia y el aprovechamiento de las tecnologías de comunicación e información, podrán ofrecer mayores oportunidades para el desarrollo del proyecto personal de cada individuo.

El gran reto de la política educativa en el mundo es lograr el doble objetivo de enseñanza de calidad con equidad. Es necesario para ello, adecuar los cursos, métodos y contenidos educativos, así como contar con las condiciones necesarias para su eficacia.

Si se entiende que la información surge como resultado del proceso cognoscitivo y como tal se comunica, será obvia su importancia y relevancia para los procesos de enseñanza/aprendizaje y por ello la necesidad de comprender los procesos de los flujos de información y su organización. Para que una persona se eduque requiere adquirir información que le enriquezca sus conocimientos y a través de ellos, adquirir nuevos conocimientos, habilidades y actitudes que le facilitan tomar decisiones para su desarrollo personal y colectivo. Deberá tener acceso a la información y conocimiento registrado, relevante para su aprendizaje.

Sin duda estamos en la transición hacia una sociedad cognoscitiva; es decir, dirigida por el conocimiento. Una sociedad del conocimiento es aquella que esta educada para acceder y recuperar la información pertinente y relevante para su propia toma de decisiones y resolución de problemas y para darle a esa información un valor agregado al sumarse a sus conocimientos y generar nuevo conocimiento.

Los primeros desarrollos tecnológicos para la transferencia de información electrónica prácticamente no tomaron en cuenta la necesidad de la participación de especialistas en la organización de contenidos de información. Sin embargo, ante la demanda de facilitar el acceso y mejorar la recuperación de la información, ha sido necesario incorporar, cada vez con una mayor participación, a los profesionales de información expertos en la organización, acceso o recuperación de información. Así han surgido nuevas tecnologías para apoyar la organización del conocimiento: estructuras y relaciones en el ambiente en línea, aplicaciones de estructuras tradicionales y no tradicionales a los sistemas digitalizados de recuperación de información (motores de búsqueda, bases de conocimiento distribuidos y multilingües, etc.).

Se dice que la recuperación de la información debe ser oportuna y relevante a las necesidades del usuario de información. El concepto de relevancia es subjetivo; dependerá de cada individuo, de sus conocimientos y de su juicio de selección. ¿Cómo se logra entonces consistencia en la recuperación? Dependerá de la organización de la información y de los modelos matemáticos de recuperación que se apliquen para agregar un factor cualitativo, así como la preparación del profesio-

nal de la información sobre la organización del conocimiento y teorías de recuperación y sobre su percepción y comprensión del tema y su terminología.

Se requieren modelos conceptuales que sirvan de base para el desarrollo de los flujos y contenidos de información, un análisis pertinente para conocer el perfil de los usuarios y el desarrollo o aplicación de los soportes tecnológicos, hardware y software de cómputo y de redes, adecuados a las necesidades de los flujos de información para su acceso y recuperación.

El diseño de sistemas de información para la educación virtual requiere tomar en cuenta: la estructura y las relaciones entre los lenguajes de indización y su recuperación, el diseño de los vocabularios controlados, la construcción de terminología y herramientas de extracción para la recuperación (tesauros y metatesauros, tesauros multilingües), todo ello soportado en la normalización que considera los problemas de compatibilidad.

El enfoque cognoscitivo de la organización del conocimiento; es decir el conceptual y sus relaciones inter-conceptuales, la formación de categorías, el uso o desuso de las clasificaciones tradicionales y no tradicionales para la organización de la información y su representación conceptual en sistemas basados en el conocimiento y en el análisis orientado por objetos, son temas actuales de discusión e investigación por expertos y serán de suma importancia al aplicarse a los flujos de información y conocimiento en el proceso educativo apoyado en tecnologías de información.

Las estructuras que surgen en la llamada organización del conocimiento deberán considerar su forma de representación en la mente humana y en las herramientas para el manejo de información, como son: los esquemas de clasificación, los tesauros y las herramientas de indización para los sistemas “inteligentes” llamado así porque contiene en sí mismos el valor agregado de experiencias y conocimientos que les permiten resolver preguntas. Las teorías de organización del conocimiento, la historia y la filosofía, las actuales divisiones de la ciencia y sus paradigmas, deben, asimismo considerarse.

Tradicionalmente la fortaleza de las profesiones que involucran a la cadena informativa se identifican principalmente con relación a las habilidades de aplicación. Sin embargo, no son meramente adjuntos que dan acceso que dan servicio a los procesos de investigación y generación de nuevos conocimientos sino que trabajan y piensan dentro de un paradigma intelectual que promueve la observación y participación dentro de otros paradigmas.⁴

4 Johanson, Graeme (1997). “Information, knowledge and research.” *Journal of Information Science* 23 (2): 103-109.

Las ciencias de la bibliotecología, documentación e información, las ciencias de la computación y las ciencias cognitivas, la cibernética y la inteligencia artificial tienen sus fundamentos y estructuras teóricas que ahora se reconocen como indispensables para participar en los modelos de universidad virtual.

En algunas universidades en el mundo se incluye el estudio de las teorías de información y recuperación en programas de posgrado en bibliotecología y ciencias de la información, en ingeniería-investigación de operaciones, telecomunicaciones -, cibernética, ciencias administrativas y procesos de toma de decisiones y en ciencias cognitivas. Sin embargo, todavía son pocas las instituciones que estudian estas teorías y aún queda mucho por investigar en estos campos y mucho que revisar en los programas curriculares para mejor entender el complejo fenómeno interdisciplinario de la transferencia de información y conocimiento; en este último, habría que agregar a las neurociencias y a las ciencias del comportamiento.

Los proyectos de universidad en línea requieren de grupos interdisciplinarios de desarrollo y apoyados en investigaciones interdisciplinarias, en los que participen, entre otros, investigadores en educación y pedagogía, cómputo y cibernética, informática y telecomunicaciones, psicología, lingüística, bibliotecología, ciencia de la información y de la comunicación. Reitero la importancia de incluir en los equipos de trabajo a especialistas en bibliotecología e información, con objeto de: apoyar el mejor flujo y organización de contenidos sobre soportes electrónicos para su integración a distancia, incluyendo los materiales para los cursos, las bibliotecas digitalizadas y el software para la interacción individual y colectiva.

Si entendemos como universidad virtual la posibilidad de continuar la misión universitaria de docencia, investigación y extensión de la cultura, con una interacción a través del soporte de las tecnologías de información, diremos que en este contexto de convergencia digital de la universidad continúa viva y se fortalece dentro y fuera de su propio campus, agregando la interacción presencial, la interacción virtual. Así entendida la universidad virtual incluye y amplía la temática y posibilidades de una universidad abierta, de la educación continua y de la educación formal, de la educación a distancia, y, lo más importante, vincula estas actividades de manera estrecha, casi indivisible, con la universidad presencial y la vida intramuros de la universidad. La universidad virtual incluye tanto a la comunidad presencial como a la comunidad extramuros.⁵

La universidad virtual debe ser una parte esencial de la transformación de la estructura universitaria en este fin de siglo. Para desarrollar la universidad virtual an-

5 Almada de Ascencio, M. (1997). "Políticas de información y universidad virtual." En *Memorias del Seminario Internacional "Investigación en información para la universidad virtual"*. México. UNAM, Centro Universitario de Investigaciones Bibliotecológicas (en prensa).

tes descrita se requiere de un extraordinario esfuerzo por parte de todos los actores del proceso universitario: los docentes y los investigadores; la administración académica universitaria; los cuerpos colegiados; los educadores, pedagogos, psicólogos, lingüistas, administradores, comunicadores, y profesionales de información e informática, incluyendo estos últimos a los bibliotecólogos, documentalistas, archiveros, especialistas en flujos de información, editores, computólogos, electrónicos, profesionales en telecomunicaciones; con el objetivo primordial de participar en el proceso educativo formal y el continuo, en la investigación y difusión de la cultura y en la administración académica, con flujos de información electrónica que faciliten la comunicación de la información y la transferencia del conocimiento.

Sin embargo aunque los principios que rigen los flujos de información electrónica nos sugieren las amplias posibilidades para la educación a distancia y la universidad virtual, la planeación de los sistemas en red y las ventajas o no de los sistemas integrados de información deben aun analizarse. En ocasiones, las administraciones universitarias privilegian las tecnologías de información sobre los recursos y servicios de información. Sin una adecuada inversión de recursos humanos y económicos, en la planeación y organización de los flujos de información para la universidad virtual, esta no podrá funcionar como tal.

En encuestas hechas a personal académico sobre el uso de tecnologías de información, indican que si bien ayudan, les ocupa un tiempo importante en la búsqueda y organización de su información, a veces con la percepción de no haber sido exhaustivos. También es cierto que las dos características cualitativas principales para el uso de tecnologías de información por personal académico es que satisfagan las necesidades de apoyo para la tarea de desempeñar y los patrones de hábito de uso más que realizar una evaluación continua de las tecnologías disponibles. Mucho de los agentes causales determinan tanto la utilización como los hábitos de uso.⁶

El cambio paradigmático no está en la tecnología sino en la relación entre estudiantes y profesores y entre éstos y sus instituciones. Las escuelas virtuales, con su propósito de establecer una relación estratégica de aprendizaje a través de la vida acerca más, vincula más a los usuarios de la universidad, transformando radicalmente el “mercado de la educación.”

A pesar de los múltiples programas de educación a distancia y de las instituciones de nivel nacional e internacional que reportan haber creado programas de educación virtual, instituciones virtuales, universidad virtual, realmente son sólo

6 Lending, D., Straub, D.W. (1997). “Impacts of a integrated information center on faculty end-users: a qualitative assessment.” *Journal of the American Society for Information Science*. 48 (5), 466-471.

etapas ciertamente importantes, pero aún primitivas. Falta contar con suficientes y adecuadas fuentes de información digitalizada, más útiles, que realmente satisfagan nuestras necesidades cognoscitivas con diversos niveles de información. También están mejorando las versiones de los diversos softwares que apoyan la organización electrónica de información y documentos, el acceso y la recuperación precisa de información y la integración de tecnologías de telefonía y periféricas que faciliten aún más la interconectividad y la facilidad de navegar simultáneamente con varias tecnologías y fuentes de información.

Entre los insumos de información mínimos que deberían aparecer en la página web de un curso a distancia se mencionan a continuación los siguientes:

- * objetivos y propósitos del curso
- * descripción del curso
- * notas de seminario
- * publicaciones
- * artículos y referencias bibliográficas
- * formatos de examen
- * vinculaciones a otros materiales relevantes en el web
- * secciones de interés general
- * acceso a colecciones digitalizadas y biblioteca virtual

Desde su planeación y diseño tanto la página web como los cursos de formación, capacitación y actualización a distancia, requieren la participación del bibliotecólogo y del informático y otros especialistas en información. Las bibliotecas digitales son esenciales para el adecuado desarrollo de este tipo de cursos. Por ello es necesario invertir más recursos para su desarrollo, mejor control y seguridad en el medio cibernético, además de alcanzar acuerdos viables en torno a la problemática de la propiedad intelectual que incluye el derecho de autor y el derecho de copia.

La universidad virtual debe facilitar la interacción entre maestros y alumnos y su acceso a un conjunto de documentos electrónicos para el curso, así como la posibilidad de hacer sus propias búsquedas sobre el tema o temas afines en bibliotecas digitalizadas y otras fuentes de información que les sean útiles para enriquecer su conocimiento, para resolver problemas, para ser más competitivos.

El maestro debe identificar los materiales necesarios para el curso y debe ser creativo para propiciar la interacción y la participación de alumnos. Debe ofrecer la posibilidad de que los alumnos tengan la interacción vía electrónica con maestros y con otros alumnos; es decir, una verdadera comunidad virtual. Los alumnos usan los medios electrónicos para presentar sus trabajos. Todo ello requiere un análisis retrospectivo y prospectivo del proceso de enseñanza/aprendizaje, del papel del maestro y del tipo de materiales, métodos y contenidos que deben fluir electrónicamente.

La capacitación y actualización a distancia representan principalmente un tipo de enseñanza “justo a tiempo” basada en la demanda más que una enseñanza “por si acaso”. Es primordial que la biblioteca que apoye esta modalidad refleje esta necesidad con acervos “justo a tiempo”, que para estos casos se requiere en vez de colecciones “por si acaso”.

En la actualidad, la educación a distancia se esta transformando de una actividad aislada y solitaria a un aprendizaje interactivo, gracias a la integración y evolución de tecnologías diversas y a su planeación interdisciplinaria. Requiere aún de la transformación en los comportamientos cotidianos de maestros y de los demás actores del proceso educativo.

La universidad del siglo XXI requiere una verdadera inclusión participativa importante, estratégica y protagónica de la biblioteca digitalizada como referente de estos materiales. Los acervos de las bibliotecas actuales y la información y los documentos digitalizados en sistemas de información que circulan en Internet y los intranet constituyen la biblioteca virtual que es capaz de ofrecer documentos e información empaquetada según las necesidades del curso o de cada individuo.

Una universidad del futuro e ser totalmente virtual o una combinación de presencial y virtual. Lpuedo que prevalecerá será la institución educativa y académica sin programas en línea y virtuales. Mucho se ha hablado del modelo de biblioteca convencional como el gran almacén del conocimiento impreso y los cambios que implica el almacenamiento o bien el acceso a documentos electrónicos. La biblioteca, entendida por sus efectos de transmisión organizada de información, pueden “construir o destruir” el edificio, pero nunca su función. La biblioteca tradicional empezó a transformarse hace algunas décadas y en la actualidad, cambia con mayor rapidez su forma de trabajo y algunos paradigmas, más no su esencia que aumenta en tanto el ser humano tenga la necesidad cognoscitiva de encontrar respuestas a sus preguntas.