



**CATALOGACIÓN DE RECURSOS EN INTERNET
EL TRATAMIENTO DE LA INFORMACIÓN ELECTRÓNICA**

ALFREDO YEGROS YEGROS
Universidad Politécnica de Valencia
Proyecto Final de Carrera dirigido por Antonia Ferrer Sopena
Valencia, 5 de julio de 2002

INDICE

1. INTRODUCCION	4
2 MATERIAL Y MÉTODOS	6
3. DIRECTORIOS CLASIFICATORIOS	7
3.1 EJEMPLOS DE DIRECTORIOS	9
4. MOTORES DE BÚSQUEDA	16
4.1 OTRAS CARACTERÍSTICAS DE LOS BUSCADORES	18
4.1.1 EL LENGUAJE DE INTERROGACIÓN	18
3.1.2 EL INTERFAZ DE USUARIO	20
4.1.3 ORDENACIÓN DE LOS RESULTADOS	21
4.2 BUSCADORES ESPECIALIZADOS.....	22
5. METABUSCADORES	31
6. VENTAJAS E INCOVENIENTES ENTRE ROBOTS, BUSCADORES Y LISTAS DE ENLACES	33
7. BIBIOTECA DIGITAL	35
7.1 DEFINICIONES Y CARACTERÍSTICAS	35
6.2 BIBLIOTECA DIGITAL: MITOS	40
7.3 BIBLIOTECA DIGITAL: RETOS.....	46
7.4 CÓMO CREAR UNA BIBLIOTECA DIGITAL.....	50
7.5 EJEMPLOS DE BIBLIOTECAS DIGITALES	53
7.6 LA COLECCIÓN DE UNA BIBLIOTECA DIGITAL	60
8. LA PROPIEDAD INTELECTUAL Y LA OBRA DIGITAL.....	61
9. ORGANIZACIÓN DE LOS RECURSOS ELECTRÓNICOS	70
9.1 TRATAMIENTO TRADICIONAL DE LOS RECURSOS ELECTRÓNICOS	70

9.2 MODIFICACIONES INTRODUCIDAS EN LAS HERRAMIENTAS DE CATALOGACIÓN PARA ACOGER LOS RECURSOS INTERNET.....	73
9.2.1 AACR2.....	73
9.2.2 Formato USMARC.....	76
9.2.3 ISBD (CF)	83
9.2.4 ANÁLISIS COMPARATIVO DE LAS NORMAS ISBD (CF) E ISBD (ER)	85
9.3 PROPUESTAS PARA MEJORAR EL ACCESO A RECURSOS ELECTRÓNICOS Y FACILITAR SU ORGANIZACIÓN.....	110
9.3.1 URI (UNIFORM RESOURCE IDENTIFIERS)	110
10. FORMATOS PARA LA DESCRIPCIÓN DE RECURSOS ELECTRÓNICO: ESTÁNDARES DE METADATOS	112
10.1 TEXT ENCODING INITIATIVE. TEI.....	113
10.1.1 SGML y TEI.....	114
10.1.2 El encabezado de TEI	116
10.1.3 Análisis de un encabezado TEI.....	116
10.2 PICS	118
10.3 DUBLIN CORE METADATA INITIATIVE.....	120
10.3.1Dublin Core	120
10.3.2 Elementos de Dublin Core:.....	123
10.4 PROYECTOS DE CATALOGACIÓN DE RECURSOS ELECTRÓNICOS	127
10.4.1 El proyecto CORC	127
10.4.2 Berkeley Digital Library	131
11 RDF	136
11.1 ANTECEDENTES	136
11.2 MODELO DE DATOS DE RDF	137
11.3 LA SINTAXIS DE RDF	138
15. CONCLUSIONES	139
BIBLIOGRAFÍA ISO 690 E ISO 690-2 (RECURSOS ELECTRÓNICOS).....	140

1. INTRODUCCION

La gran cantidad de información disponible en Internet, la falta de organización y la poca eficiencia de las herramientas de recuperación disponibles hasta el momento (principalmente buscadores y directorios) contribuyen a la ineficacia en la recuperación de recursos en Internet. A pesar de que los directorios ofrecen algún tipo de categorización, sus categorías temáticas a menudo carecen de referencias cruzadas y los listados temáticos pueden llegar a ser demasiado largos y poco manejables. La creación de categorías temáticas resulta una tarea costosa para los humanos.

Debido a que la creación y mantenimiento de los directorios implican la intervención humana, es comprensible que cubran una parte incluso menor de Internet que los buscadores.

Existen valiosos recursos de información que exigen métodos más efectivos y sistemáticos para identificarlos, describirlos y recuperarlos.

En este sentido existen algunas discusiones y debates entre los profesionales de la información sobre diferentes aproximaciones para la recuperación de recursos en Internet. Entre estas aproximaciones cabe destacar tanto la creación de bibliotecas virtuales o pasarelas temáticas y el uso de metadatos para la descripción de recursos.

Por todo esto, con el presente Proyecto Final de Carrera (en adelante Proyecto) se pretenden alcanzar los siguientes objetivos:

1. Estudiar el estado de la cuestión sobre la recuperación de información en Internet, haciendo un breve repaso sobre herramientas como buscadores, metabuscadores y directorios.
2. Analizar las nuevas tendencias y proyectos en cuanto a la recopilación y descripción de recursos electrónicos.

Para la realización del Proyecto se han utilizado no solo fuentes bibliográficas tradicionales (artículos de revista, monografías, etc.) sino que también se ha utilizado la información que algunas organizaciones (Dublin Core Metadata Initiative, OCLC, etc.)

tienen disponible en Internet, con el fin de obtener una panorámica lo más actualizada posible.

2 MATERIAL Y MÉTODOS

La metodología utilizada en el desarrollo del Proyecto ha sido la siguiente:

1. Búsquedas en bases de datos bibliográficas
 - Library and Information Science Abstracts (LISA)
 - Information Science Abstracts (ISA)
 - Current Contents
 - Documents in information Science (DOIS)
 - Base de datos ISOC
2. Obtención de los documentos originales correspondientes a las referencias pertinentes, obtenidas tras las búsquedas en las bases de datos anteriormente mencionadas.
3. Búsqueda en Internet, utilizando diferentes herramientas (directorios, buscadores) para localizar información sobre instituciones y proyectos relacionados con el tema de nuestro Proyecto.
4. Elaboración del Proyecto, tratando de integrar toda la información recopilada a través de las distintas fuentes de información, con el objetivo de proporcionar una exhaustiva y actualizada panorámica sobre el estado de la cuestión.

3. DIRECTORIOS CLASIFICATORIOS

Los directorios son una de las herramientas existentes para localizar información en Internet. Son catálogos de páginas Web organizados por categorías temáticas. Cada Web es clasificada manualmente por personas. El directorio clasificatorio busca en sus recursos pero no en toda la Web. Los directorios conducen a la página inicial de un sitio Web y desde allí se explora hasta encontrar lo buscado.

Hay dos modos de inclusión de una página Web en un directorio temático:

Los interesados en figurar envían un formulario en el que la principal dificultad está en clasificarlo según su estructura, ellos lo miran y deciden si incluirla o no. Es decir, los autores de las páginas Web clasifican ellos mismos los temas de los que trata la página cuando se inscriben en el buscador, y posteriormente se busca en esas clasificaciones realizadas previamente. Con esto, la indización no será homogénea porque no la realizará siempre la misma persona y porque cada autor que analice una página puede tener unos criterios diferentes a los de otros autores.

El segundo modo es mediante recopilación automática, lo cual supone un retraso en el ritmo de actualización de la red. Una vez realizada la recopilación automática, se estudia si se incorpora o no la dicha página a Internet, y posteriormente se clasifica con la intervención de un especialista.

En los directorios que no existe la poli jerarquía, es decir la posibilidad de más de una clasificación, cada página Web está clasificada en un único epígrafe clasificatorio y no siempre está claro qué epígrafe asignar a la página, lo cual constituye un problema a la hora de buscar. Por otra parte, también se puede añadir otro defecto relativo al análisis documental, esto es, la descripción se realiza a partir de un resumen que se añade en los formularios de alta. Además no incluyen en su base de datos todos los términos que forman parte de una Web con todos sus niveles. Por ejemplo, pueden llegar a la página de la Universitat Politècnica de València pero no a las páginas interiores. La ventaja de los directorios es que es fácil encontrar la información.

Se recomienda emplear directorios en las búsquedas en los siguientes casos:

- Cuando se tenga una idea vaga de lo que se quiere y por tanto se agradecería la ayuda
- Cuando se quiera localizar sitios por temas amplios (música, cine, literatura...)
- para localizar información de lo cual apenas tenemos conocimiento, exactamente igual que cuando empleamos las páginas amarillas
- Cuando se quiera curiosear por categorías temáticas y ver que hay en la red.

3.1 EJEMPLOS DE DIRECTORIOS

YAHOO! <http://www.yahoo.com/>

Contenido: cobertura internacional con nodos locales que sirven a sus respectivos ámbitos geográficos en: Alemania, Asia, Australia, Canadá, Corea, China, Dinamarca, España, Estados Unidos, Francia, Italia, Japón, Noruega, Reino Unido y Suecia.

Recogida de Páginas: la mayoría de los ‘sites’ son sugeridos por los propios usuarios, y más tarde son visitados y evaluados por un equipo de Yahoo!. Existe un formulario de sugerencias donde se solicita la URL, el título y la descripción del recurso a incluir. Yahoo! está organizado en categorías temáticas, por lo que el solicitante además debe indicar la(s) categoría(s) en las que piensa debe enmarcarse el recurso sugerido.

Búsqueda: se puede buscar siguiendo dos procedimientos: o bien a través de la clasificación temática o bien utilizando los formularios de búsqueda. En el primer caso se puede descender en la jerarquía hasta donde se considere oportuno y visualizar los recursos de interés. En el segundo, se teclean los términos en una ventana debiéndose elegir entre visualizar los epígrafes clasificatorios o los recursos que contengan dichos términos. Si el resultado no es satisfactorio, se tiene la posibilidad de lanzar la pregunta en las páginas indizadas por el motor de búsqueda Inktomi o conectar con otro buscador (Altavista, Hotbot, Infoseek, etc ...). El sistema admite trincar, buscar por términos compuestos y el uso de los operadores booleanos AND (+), NOT (-) y OR. La opción ‘Search Options’ permite hacer una acotación en función del tiempo transcurrido desde el alta del recurso, preguntar por una URL determinada o buscar los términos en el título del recurso.

Resultados: no se puede formatear el resultado de una búsqueda. De cada ‘site’ se visualiza el título con el enlace, una pequeña descripción y el epígrafe clasificatorio al que pertenece, en el caso de que la búsqueda se haya realizado en la base de Yahoo!. Si la búsqueda se ha lanzado en Inktomi la descripción se

corresponde con las primeras líneas de las páginas. Alcanzarán un mayor rango en la presentación aquellos documentos que contengan más palabras clave, que estas estén en determinadas partes del documento (se prima el título) o que el recurso pertenezca a una categoría que esté situada en una posición elevada, es decir más general.

Ayudas: las ayudas detallan la forma de realizar las búsquedas con ejemplos, los operadores a utilizar y los criterios de relevancia que se aplican en la presentación de resultados.

Otros: a las nuevas incorporaciones de les añade la etiqueta de ‘nuevo’ y se las posiciona al principio de cada epígrafe durante la primera semana.

Positivo: 1) El alta de recursos se realiza a través de un formulario donde se solicita la URL, el título y la descripción del mismo. 2) Existe una clasificación temática. 3) Posibilidades de búsqueda satisfactorias ya que se pueden utilizar los operadores AND, OR y NOT, se puede truncar y localizar términos compuestos. 4) Se pueden acotar las búsquedas a la URL y al título. Las búsquedas en texto libre se realizan sobre la descripción del recurso. 5) Se recuperan unidades de información y no páginas. 6) Las ayudas son satisfactorias. 7) Se pueden detectar las nuevas incorporaciones al acotar por fecha de alta del recurso. 8) Tiene versión en español.

Negativo: 1) Posibilidades de búsqueda medianamente satisfactorias, ya que aunque el sistema permite truncar, localizar términos compuestos, y utilizar los operadores AND, OR y NOT, no es posible utilizar paréntesis. 2) No se controla el vocabulario. 3) No se pueden diseñar los formatos de presentación. 4) No se pueden mantener abiertas paralelamente la lista de resultados de la búsqueda y una de las direcciones recuperadas.

WEBCRAWLER <http://www.webcrawler.com>

Contenido: cobertura internacional.

Recogida de páginas: los usuarios pueden sugerir nuevas páginas a partir de un formulario en el que únicamente deben consignar la URL de la página a incluir. Las páginas son visitadas y dadas de alta definitivamente por un equipo de expertos de Webcrawler.

Búsqueda: Webcrawler está organizado en categorías en las que se encuentran enmarcadas las páginas visitadas y evaluadas por los expertos. Existen dos procedimientos para realizar las búsquedas: o bien descendiendo por las categorías de su clasificación hasta llegar al epígrafe adecuado o bien tecleando los términos deseados sobre un formulario de búsqueda. En este formulario se debe especificar el conjunto de documento sobre el que se quiere buscar: documentos revisados por Webcrawler, documentos con luz verde o en la base de datos de Excite (The Entire Web). No funciona el truncado. Las búsquedas pueden realizarse a través de unitérminos o de términos compuestos que pueden ser combinados entre sí con operadores booleanos (AND, OR y NOT). Se pueden utilizar paréntesis. No se puede limitar por campos. Al igual que en Excite se pueden pedir documentos con las mismas características de aquel o aquellos que más se ajustan a las preferencias del solicitante (Find Similar). También como en Excite en la presentación se indica que el sistema realiza las búsquedas por conceptos, es decir que ante una pregunta el sistema identifica los conceptos representados por las palabras utilizadas y es capaz de localizar estos mismos conceptos en otras páginas. Lo que no se explica es el procedimiento exacto para realizar esta búsqueda inteligente.

Resultados: los resultados se presentan según un orden de relevancia establecido por el sistema, que no se explica en las ayudas. De cada una de las páginas seleccionadas aparece el título con el enlace, un pequeño resumen que sirve como descripción de la página, la URL y el % de relevancia.

Ayudas: en las ayudas se detalla convenientemente como realizar una búsqueda presentándose ejemplos.

Positivo: 1) La selección de páginas la hace un equipo de expertos del propio Webcrawler. 2) Todos los recursos de información existentes en la base están clasificados temáticamente. 3) Existe cierto control del vocabulario. 4) Se pueden pedir documentos de iguales características a uno designado. 5) Se recuperan unidades de información y no páginas. 6) La descripción del recurso se hace a través de un pequeño resumen. 7) Los recursos se ordenan utilizando criterios de relevancia. 8) Las ayudas a la búsqueda son completas, pero no se explican ni los criterios de indización ni los de relevancia para establecer el orden de presentación de recursos.

Negativo: 1) Las posibilidades de búsqueda no son completas, pues no se puede trincar. 2) No se puede buscar por campos concretos. 3) No se pueden seleccionar formatos. 4) No hay forma de detectar las novedades. 5) No existe versión en español.

LOOKSMART <http://www.looksmart.com/>

Contenido: cobertura internacional con versiones locales en Australia y el Reino Unido.

Recogida de Páginas: existe un formulario de alta donde los interesados pueden sugerir nuevos recursos. En dicho formulario se demanda título, URL y descripción como datos obligatorios, y palabras-clave y fecha en la que el recurso fue puesto en la red como datos opcionales. Los datos deben de ser rellenados en el epígrafe clasificatorio en el que se piensa debe de estar enclavado el recurso. Esta información es revisada por un equipo de expertos de Looksmart quienes dan el alta definitiva.

Búsqueda: las búsquedas pueden realizarse o bien descendiendo por la clasificación temática o utilizando un formulario. El formulario es muy simple, pues tan sólo cuenta con una ventana donde han de teclearse los términos. En el formulario de búsqueda las palabras tecleadas son por defecto unidas con el operador AND. No se pueden utilizar ni operadores de búsqueda ni trincar ni localizar términos compuestos. No se puede acotar por campos. Existe la posibilidad de enlazar con el formulario para búsquedas avanzadas de Hotbot y lanzar la interrogación en este sistema.

Resultados: cuando se desciende por la clasificación lo que se obtiene son los recursos dados de alta en Looksmart. Si se utiliza el formulario de búsqueda, en primer lugar aparecerán los recursos existentes en Looksmart y a continuación las páginas obtenidas en Altavista. En el primer caso se obtiene el título con el enlace, la descripción y la categoría de la clasificación temática a la que pertenece con un enlace activo. En el segundo los datos proporcionados por Altavista. En el caso de Looksmart no se aplican criterios de relevancia en el orden de presentación de los recursos.

Ayudas: las ayudas son escasas en cuanto a la forma de realizar las búsquedas.

Positivo: 1) El alta de recursos se realiza a través de un formulario donde se solicita título, URL, clasificación, descripción, palabras clave y fecha de puesta en marcha del recurso. 2) Todos los recursos están clasificados siguiendo su índice temático. 3) Se recuperan recursos y no páginas. 4) Los recursos van acompañados de un resumen descriptivo del contenido. 4) Las ayudas son aceptables.

Negativo: 1) No se puede utilizar ningún operador de búsqueda ni trancar ni localizar términos compuestos. 2) No se puede acotar por campos. 3) No se controla el vocabulario. 4) No se puede seleccionar formatos. 5) No se aplica ningún criterio de relevancia en el orden de presentación de recursos. 6) No se pueden detectar las novedades.

OLE <http://www.ole.es/>

Contenido: Webs en español.

Recogida de Páginas: el método utilizado para localizar e incluir recursos es el alta por parte de los responsables a partir de un formulario. Antes de introducir los datos es necesario posicionarse en aquella categoría de la clasificación general en la que se piensa debe ubicarse el nuevo recursos. En el formulario se solicitan los siguientes datos: título, URL, país, idioma, ámbito, zona de influencia, tipo de servicio, descripción, palabras-clave, e-mail de contacto técnico y de contacto de contenido.

Búsqueda: se puede acceder a la información de dos modos: o descendiendo por el árbol de sectores o realizando una búsqueda desde cualquier punto de este mismo árbol utilizando un formulario desde donde se definen las búsquedas debiendo decidir si se truncan las palabras, si se enlazan con AND o con OR y en que campo se desea sean localizados los términos: título, URL, descripción, palabras clave, zona de influencia, sector, idioma, ámbito, tipo de servicio y país de origen. No se puede utilizar el operador NOT ni buscar términos compuestos ni usar paréntesis. En su opción "Temas" ofrece la posibilidad de visualizar por orden

alfabético todas las categorías de su clasificación y en la de "Arbol" de desplegar la clasificación por apartados.

Resultados: no existe posibilidad de elección de formato de salida, apareciendo los recursos seleccionados agrupados por categorías y dentro de cada una de ellas por orden alfabético. De cada recurso se presenta el nombre con el enlace y la descripción. Se destacan las nuevas incorporaciones colocándolas por delante una "N".

Ayudas: en la ayuda se orienta muy someramente sin ejemplos sobre como se debe realizar una búsqueda.

Positivo: 1) El alta de recursos se realiza a través de un formulario donde se solicitan los siguientes datos: título, URL, país, idioma, ámbito, zona de influencia, tipo de servicio, descripción, palabras-clave, e-mails de contacto técnico y de contacto de contenido. El alta debe de hacerse desde el epígrafe clasificatorio adecuado. 2) Los recursos están organizados según una clasificación temática. 3) Se puede limitar por los siguientes campos: título, URL, descripción, palabras clave, zona de influencia, sector, idioma, ámbito, tipo de servicio y país de origen. 4) Se recuperan recursos de información y no páginas. 5) En la visualización de resultados se reproduce la descripción proporcionada por el responsable del recurso. 6) Las novedades pueden ser detectadas, pues las nuevas incorporaciones se destacan del resto durante 15 días a partir de su incorporación.

Negativo: 1) Las posibilidades de búsqueda son limitadas pues se puede truncar y utilizar los operadores AND y OR, pero no se puede utilizar el operador NOT ni el paréntesis ni localizar términos compuestos. 3) No se controla el vocabulario. 4) No se puede elegir formato de salida. 5) No se utilizan criterios de relevancia en la presentación de recursos. 6) Las ayudas no son muy completas y no presentan ejemplos de búsquedas.

4. MOTORES DE BÚSQUEDA

Los buscadores de la W3 presentan una estructura constituida por: un robot o araña, es decir, un programa que cruza la W3 moviéndose de un documento a otro, descendiendo progresivamente a través de los hiperenlaces; un programa de indización que indiza la información de los millones de páginas Web ubicadas en servidores conectados a la red y enormes bases de datos a las que acceden los usuarios a través de la interfaz del buscador. Por tanto, los buscadores no sólo deben facilitar la localización de los recursos incluidos en sus bases de datos sino que, además, deben compilarlos.

Salvo en el caso de los directorios temáticos como Yahoo, Olé y otros, la indización automática es el método predominante utilizado por las herramientas de consulta de la W3. No obstante, pese a la ventaja de la rapidez, este sistema, que conlleva necesariamente el uso de robots automáticos, también cuenta con detractores los cuáles identifican los siguientes problemas:

- El uso de robots automáticos de búsqueda de manera indiscriminada produce un incremento del tráfico en la red, una sobrecarga de los servidores y otros problemas de uso de las infraestructuras.
- La justificación de sistemas de indización mediante robots sería difícil si la red cambiara hacia un uso no gratuito de los recursos.
- El tipo de datos recogidos por los robots no es útil, ya que aun hoy las arañas presentan un funcionamiento demasiado simple.
- Los sistemas de comprensión del lenguaje natural no están lo suficientemente avanzados como para extraer el significado de los recursos.
- La mayor parte de los buscadores sólo reconocen texto, por lo que se hace más difícil generar una identificación automática de las características de recursos tales como las imágenes digitales o los diferentes ficheros multimedia.

- La indización automática tiende a una perspectiva simplista, poco selectiva que provoca que la localización de recursos en la red y la recuperación de la información (RI) solicitada llegue a ser cada vez menos factible.
- Frente a los indicadores humanos, los programas automatizados tienen dificultades para identificar características de un documento Web como el contexto temático general en la que se engloba, y el género, por ejemplo una comunicación científica, información profesional o informal, al que ese recurso pertenece.
- La W3 carece de reglas para facilitar la indización automática. Los documentos no están estructurados de forma que los programas puedan obtener de modo fiable la información, conocida como metadatos (autor, título, longitud del texto, materia) que un indicador humano detectaría fácilmente tras una rápida revisión.
- Los editores y/o creadores de estos recursos a veces abusan del carácter indiscriminado de la indización con el fin de atraer la atención de los usuarios, mediante la repetición en el documento de una palabra como sex, muy utilizada en las búsquedas, aunque sea otro su contenido.
- El indicador profesional puede describir los componentes de páginas individuales de distinto tipo y puede aclarar de qué forma han de incluirse esas partes en una base de datos de información.
- La información cambia con frecuencia y las arañas únicamente actualizan las bases de datos de los buscadores con cierta periodicidad. Muchas páginas Web no son ficheros estáticos sino que recogen información dinámica, perecedera y en constante cambio lo cuál dificulta el que puedan ser analizadas e indizadas por los programas.

Se recomienda emplear un buscador en los casos siguientes:

- Cuando se quiera llegar rápidamente a una información
- Para buscar o localizar un producto en particular y se sepa el nombre

- Cuando se busque a una persona por su nombre y apellidos
- Para encontrar artículos y textos concretos
- Cuando se necesite información concreta sobre un tema, un producto

4.1 OTRAS CARACTERÍSTICAS DE LOS BUSCADORES

4.1.1 EL LENGUAJE DE INTERROGACIÓN

Dado que los lenguajes de recuperación cambian de un buscador a otro es siempre recomendable acceder a los archivos de ayuda para ver las posibilidades que ofrecen.

La búsqueda de palabras simples: es la forma de búsqueda más habitual que reciben los motores de búsqueda.

La búsqueda de referencias compuestas: la mayoría de buscadores ofrecerán ruido ante este tipo de búsqueda y ofrecerán documentos en los que estén ambos términos pero no tengan nada que ver con lo pedido, lo recomendable para estas búsquedas es emplear siempre comillas.

Uso de operadores: para emplear operadores en algunos casos hay que ir a la denominada búsqueda avanzada. Para conocer que operadores emplea un buscador hay que leer la ayuda.

La búsqueda con operadores booleanos: el empleo de la lógica booleana es la más extendida entre los buscadores. Los operadores más comunes son OR, AND y NOT. Los buscadores que emplean el operador OR (por ejemplo Lycos) por defecto son los que peor resultado ofrecen pues el índice de ruido es muy elevado. Para evitar este ruido muchos motores han puesto por defecto el operador AND, que reduce el ruido aunque no lo elimina. En Alta Vista el operador AND puede sustituirse por el punto y coma, ejemplo Java; Database. El operador NOT es mucho más eficaz para recuperar información. Cuando se emplea, se está ordenando al motor que encuentre las páginas que contienen un término determinado y que omita otro. Yahoo y Lycos no pueden procesar el operador NOT.

La búsqueda con operadores posicionales (with, near, adj): no son muy comunes en los motores actuales. Un recurso equiparable a los operadores son las “””. Cuando se buscan varios términos entrecomillados el motor sabe que está buscando una frase y no las palabras por separado.

La búsqueda con operadores relacionales: son los operadores de comparación, muy útiles pero poco extendidos. Sirven para acotar el resultado de la búsqueda limitando el resultado temporalmente o por idioma o por tipo de fichero.

La búsqueda cualificada: permite buscar delimitando por campos. Sirve para limitar la búsqueda de uno o varios términos a uno o varios campos. Poco utilizado por los usuarios en aquellos motores, como Altavista, en que se permite. Se puede acotar para que busque las palabras que representen los temas buscados sólo en el campo título (etique Title) o en que busque sólo en las etiquetas META, o acotar la búsqueda a los servidores existentes en España (domain:es)

La búsqueda con truncamiento: los truncamientos suelen realizarse siempre por la derecha. Algunos motores emplean esta opción sin necesidad de indicárselo, buscan los términos introducidos a partir de la raíz del mismo.

La búsqueda distinguiendo entre mayúsculas y minúsculas: algunos motores distinguen entre mayúsculas y minúsculas por lo que el resultado es distinto. Por ello como norma se recomienda emplear las palabras en minúsculas ya que entonces se recuperan todas las palabras estén como estén escritas.

La búsqueda teniendo en cuenta un fichero de palabras vacías: algunos motores no indexan palabras sin significado como preposiciones o artículos. Emplean listas de palabras vacías (stopwords). Esto mejora los tiempos de indexación y de búsqueda en la base de datos. Cada servidor tiene su propia lista y se puede consultar.

Refinamiento de la búsqueda: algunos buscadores (Google, Buscopio) permiten refinar la búsqueda bien mediante búsqueda sobre los resultados o bien mediante búsqueda por patrones de documentos

Algunos buscadores ofrecen sugerencias de términos al realizar una búsqueda. La opción Live Topics de Alta Vista presenta, tras introducir los términos a buscar y pinchar en Live Topics, una página de las posibles palabras clave relacionadas en un formato de verificación. Se pueden seleccionar las deseadas.

Algunos motores de búsqueda han agregado a sus paquetes la posibilidad de buscar imágenes. Por ejemplo Alta Vista permite buscar imágenes mediante la sintaxis “image:” como prefijo para una búsqueda o bien seleccionado la opción image. También Google ha incorporado esta posibilidad con muy buenos resultados. Yahoo dedica secciones enteras de su directorio a diseño gráfico. Otra posibilidad, en buenos buscadores, es buscar el tema y añadir el tipo de fichero correspondiente a imagen (jpeg, gif).

3.1.2 EL INTERFAZ DE USUARIO

Se refiere tanto al modo en que el sistema permite al usuario introducir sus búsquedas como a la forma de presentación de los resultados. Los interfaces de búsqueda son bastante homogéneos basados en formularios muy simples son una sola casilla en donde se escriben los términos a buscar. Otros presentan formularios más complejos, con dos o más recuadros para escribir, botones, listas de opciones para seleccionar idioma o fechas.

En cuanto al formato de visualización de resultados se suelen ofrecer en una pantalla unos 10 ó 20 resultados para no ralentizar la descarga cuando el resultado es excesivamente amplio.

Prácticamente todos muestran el título del documento, otros ofrecen un breve resumen a partir de los términos encontrados en el texto y resaltan los hits (expresiones que coinciden con las formuladas en el perfil), el tamaño de la página, la fecha de modificación de la página, la URL.

4.1.3 ORDENACIÓN DE LOS RESULTADOS

La mayoría de buscadores recurren a algoritmos clásicos de las teorías de recuperación de información, como la de Salton y la de Blair y fórmulas de frecuencia de aparición de un término que suelen tener en cuenta el número de veces que aparecen las palabras, si las palabras están juntas, si están más cerca del título o no, etc. Y ofrecen un ranking de relevancia a partir de estas reglas.

Este sistema no evita el ruido y tampoco asegura que se recuperen documentos relevantes (un documento muy relevante puede tener el término o términos buscados pocas veces en el texto).

Además surgen problemas nuevos como el fenómeno denominado “spamming”, por medio del cual los constructores de páginas Web se aseguran que su página se va a recuperar incluyendo muchas veces el mismo término o incluso introduciendo muchas veces términos muy buscados, como sex, mp3 y otros, que no tienen que ver con su contenido pero hacen que se recupere y consiguen que su Web salga siempre y se conozca, dándose así publicidad. Otro problema que ha surgido es que determinadas Web pagan a los buscadores para que sea lo que sea lo que se busca salir ellos en primer lugar.

Otro método que ha revolucionado la recuperación en Internet es el PageRank empleado por Google. Google fue creado por Sergey Brin y Lawrence Page de la Universidad de Stanford, y el resultado de sus investigaciones ha sido uno de los mejores motores de recuperación que existe actualmente.

Para calcular la relevancia y ordenar los resultados, Google hace uso de la conectividad de la Web para calcular el grado de calidad de cada página, esta graduación es el PageRank.

El PageRank asume que el número de enlaces (que puede compararse con las citas) hacia una página tiene que ver con la calidad de la misma. Cuantas más páginas apunten a otra el índice de PageRank es mayor.

El proyecto Cléber (<http://www.almaden.ibm.com/cs/k53/clever.html>), en el que trabajaba IBM, también tenía en cuenta el PageRank. Este proyecto no se terminó puesto que los creadores de Google ofrecieron antes el descubrimiento y su motor de búsqueda.

Se han realizado numerosos estudios comparativos entre buscadores de los que se puede afirmar que el buscador más exhaustivo es Alta Vista, Infoseek da mayor relevancia y Hot Bot da buenos resultados de búsqueda.

Sin embargo han ido apareciendo buscadores de nueva generación como fase, especialmente Google, que es sin duda el mejor buscador en la actualidad. Y desde septiembre de 2001 ya están operativos dos nuevos buscadores, Wisenut (<http://www.wisenut.com>) y Teoma (<http://www.teoma.com>). Wisenut tiene en cuenta el método PageRank para ordenar los resultados de las búsquedas.

4.2 BUSCADORES ESPECIALIZADOS

Para efectuar una búsqueda más precisa se puede recurrir a un motor de búsqueda especializado cuya función consiste en recopilar todos los sitios relevantes de un tema en particular.

Una buena recopilación de buscadores especializados es:

- **Buscopio** <http://www.buscopio.com> que nació primero como un buscador de buscadores y ahora es también un buscador. Está organizado por temas.
- **FindLaw** <http://www.findlaw.com> especializado en legislación
- **HealtzAtoz** <http://www.healthatoz.com> indiza sitios relacionados con la salud y la medicina

EJEMPLOS DE BUSCADORES

ALTAVISTA <http://altavista.com/>

Contenido: cobertura internacional con nodos locales que sirven a sus respectivos ámbitos geográficos y lingüísticos en los siguientes países: Australia, Canadá, España, Estados Unidos, Japón y Suecia.

Recogida de Páginas: las páginas son recogidas por un robot que es capaz de indizarlas y de saltar a aquellas otras enlazadas con ellas. Las altas se realizan a partir de un formulario donde únicamente se solicita la dirección de la página a dar de alta. El sistema es capaz de reconocer las etiquetas META de título, descripción y palabras-clave y extraer la información que contienen.

Búsqueda: presenta dos formularios de búsqueda con casi las mismas prestaciones, ya que se puede acotar por idioma, buscar en campos concretos (dominios, servidores, URL's, links, imágenes, título e idioma), utilizar operadores booleanos (AND, OR, NOT), de proximidad y paréntesis, localizar términos compuestos y truncar. Los dos también contemplan la posibilidad de 'Afinar' la búsqueda, es decir hacer una segunda selección entre las páginas recuperadas originalmente en función de la elección de nuevos términos propuestos por el sistema por el grado de coincidencia con los demandados en un principio. La diferencia entre los dos formularios reside en que en el simple las páginas seleccionadas son presentadas por un orden de relevancia establecido por el propio sistema y en el avanzado es el usuario el que decide el orden de presentación haciendo uso de la opción 'Ranking', indicando el término que deben contener las páginas presentadas en primer lugar. En el formulario para búsquedas avanzadas se puede seleccionar documentos que hayan sido publicados o modificados por última vez en un periodo de tiempo establecido.

Resultados: no se puede formatear el resultado de una búsqueda. De cada página se visualiza: el título con el enlace, la URL y las primeras palabras en el caso de que no contenga etiqueta META de descripción, sí es así aparece el contenido de

la misma. Además, aparece el tamaño de la página (bits), el idioma y la fecha de la última modificación. Los criterios de relevancia para la ordenación de las páginas de una búsqueda simple son: si las palabras o frases demandadas se encuentran en las primeras líneas de una página; la frecuencia de ocurrencia; si las palabras o frases se encuentran cerca una de otra en el documento. Se puede mantener abiertas simultáneamente la página de resultados y uno de los documentos presentados.

Ayudas: las ayudas detallan con ejemplos la forma de realizar las búsquedas, los operadores a utilizar y los criterios de relevancia que se aplican en la presentación de resultados.

Otros: Altavista conecta desde su página principal con 'Looksmart', índice de páginas clasificados en categorías.

Positivo: 1) Reconoce las etiquetas META y utiliza su información para indizar las páginas y como criterio de relevancia a la hora de presentación de resultados. 2) Las posibilidades de búsqueda son bastante completas: truncado, términos completos, operadores booleanos (AND, OR, NOT) y paréntesis. 3) Posibilidad de buscar por campos concretos: dominios, servidores, URL's, links, imágenes, título e idioma. 4) Posibilidad de "afinar" la búsqueda utilizando términos propuestos por el sistema. 5) Posibilidad de mantener una página abierta a la vez que se visualiza el listado de resultados. 6) Las ayudas son bastante completas. 7) Se puede recuperar páginas publicadas o revisadas en un periodo de tiempo dado. 8) Versión en español.

Negativo: 1) Carece de clasificación temática. 2) No se puede recuperar información utilizando los campos de descripción y palabras clave. 3) No se controla el vocabulario. 4) Se recuperan páginas y no recursos de información que además no pueden ser agrupadas por servidores. 5) No se pueden diseñar los formatos de presentación.

EXCITE <http://www.excite.com>

Contenido: cobertura internacional con versiones locales en Alemania, Francia, Holanda, Suecia, Reino Unido, Japón, China y Australia.

Recogida de páginas: las páginas que contiene su base de datos son localizadas e indizadas por un robot, aunque también se pueden dar de alta usando un formulario existente para tal fin. El robot desestima toda la información proveniente de las etiquetas META.

Búsqueda: existen dos formularios de búsqueda: uno con una única ventana y otro más completo. Las búsquedas pueden realizarse a través de palabras o de términos compuestos que pueden ser combinados entre sí con operadores booleanos (AND, OR y NOT). Se pueden utilizar paréntesis y trincar. No se puede limitar por campos. En la búsqueda aparecen dos funciones interesantes, por una parte la sugerencia de palabras asociadas a los términos solicitados (Search Wizard) y por otra la posibilidad de pedir documentos con las mismas características de aquel o aquellos que más se ajustan a las preferencias del solicitante (More Documents like this one). En la presentación y ayudas se indica que Excite realiza las búsquedas por conceptos, es decir que ante una búsqueda, el sistema identifica los conceptos representados por las palabras utilizadas y es capaz de localizar esos mismos conceptos. Lo que no explica las ayudas es el procedimiento exacto para realizar esta búsqueda inteligente.

Resultados: los resultados se presentan según un orden de relevancia establecido por el sistema. De cada una de las páginas seleccionadas se puede elegir entre visualizar únicamente el título y el enlace o además la descripción (primeras palabras de la página), la URL y el % de relevancia. Las páginas pueden ser agrupadas por servidores (List by Web Site).

Ayudas: en las ayudas se detalla convenientemente como realizar una búsqueda y los criterios de presentación de recursos.

Otros: también presenta una selección de páginas clasificadas por temas

Positivo: 1) Se pueden utilizar los operadores AND, OR y NOT, utilizar los paréntesis, trincar y localizar términos compuestos. 2) Posibilidad de ‘afinar’ la búsqueda utilizando términos asociados propuestos por el propio sistema. 3) Se pueden pedir documentos con las mismas características a uno anteriormente seleccionado. 4) Realiza un cierto control del vocabulario en los términos de búsqueda. 5) A la hora de presentación de resultados se puede elegir entre un formato reducido (sólo título) y uno con mayor información. 6) Las páginas recuperadas pueden ser agrupadas por servidores. 7) Las ayudas proporcionan suficiente información.

Negativo: 1) Las páginas clasificadas temáticamente representan una pequeña parte del total. 2) No reconoce las etiquetas META. 3) No se puede buscar por campos, salvo por idioma. 4) La descripción de las páginas se corresponde con las primeras palabras de la página. 5) Se recuperan páginas y no recursos, aunque estas puedan ser agrupadas por servidor. 6) No se pueden mantener varias páginas abiertas a la vez que se visualiza la lista de resultados. 7) No hay forma de detectar las páginas nuevas. 8) No existe versión en español.

LYCOS <http://www.lycos.com/>

Contenido: cobertura internacional con nodos locales en: Alemania, Bélgica, España, Estados Unidos, Francia, Holanda, Italia, Reino Unido, Suecia y Suiza.

Recogida de Páginas: las páginas son localizadas por un robot a partir de las URL's dadas de alta en el formulario que existe para tal fin, donde sólo se solicita la URL y el e-mail de la persona que realiza el alta. Ignora las etiquetas META de descripción y palabras clave.

Búsqueda: cuenta con un formulario simple y otro más complejo (Lycos Pro). En ambos se puede acotar por idioma y utilizar: operadores booleanos (AND, OR y NOT), operadores de proximidad y paréntesis. Sin embargo, no se admite el truncado. Además en el formulario complejo se puede elegir entre buscar los términos de búsqueda en todo el documento, el título o la URL. También se pueden hacer búsquedas en un Web determinado existente en Lycos, indicando su dominio en el formulario de búsqueda. La opción "más resultados" permite visualizar las páginas de iguales características que una seleccionada. Se pueden localizar imágenes y sonidos.

Resultados: se puede elegir entre visualizar las páginas seleccionadas agrupadas por dominios o hacerlo individualizadamente. De cada página se visualiza: el título con el enlace, la URL, las primeras palabras a modo de resumen, el tamaño en bits de la página y el grado de pertinencia. El usuario puede elegir entre varios el criterio de relevancia a utilizar para el orden de presentación de las páginas.

Ayudas: las ayudas detallan la forma de realizar las búsquedas, los operadores a utilizar y los criterios de relevancia a aplicar en la presentación de resultados.

Otros: además de la base de datos confeccionada a partir del motor de búsqueda, Lycos contiene lo que denomina Webguide, que no es más que un índice

de recursos seleccionados clasificados por temas. No tiene motor interno de búsqueda.

Positivo: 1) Se puede buscar por los siguientes campos: idioma, título y URL. 2) Se pueden hacer búsquedas dentro de un Web determinado existente en Lycos. 3) Visualización de páginas con igual características que una seleccionada. 4) Se pueden localizar imágenes y sonidos. 5) Se recuperan páginas, pero pueden ser agrupadas por dominios. 6) El usuario puede elegir los criterios de relevancia en el orden de presentación de los resultados. 7) Las ayudas son bastante completas. 8) Existe versión en español.

Negativo: 1) No reconoce las etiquetas META de palabras claves ni de descripción, por lo tanto su contenido no puede ser utilizado ni en la indización ni en la presentación de resultados. 2) Sólo están clasificadas parte de las páginas. 3) Las posibilidades de búsqueda son incompletas, pues no admite el truncado. 4) No se controla el vocabulario. 5) En los resultados la descripción de la página se realiza a través de las primeras líneas del texto de la misma. 6) No se pueden detectar las novedades.

INFOSEEK <http://infoseek.go.com/>

Contenido: cobertura internacional con versiones locales en España, Alemania, Francia, Italia, Dinamarca, Suecia, Reino Unido y Japón.

Recogida de Páginas: las páginas que contiene su base de datos son localizadas e indizadas por un robot, también se pueden dar de alta usando un cuestionario que el propio sistema presenta donde exclusivamente se solicita la URL de la página. Las páginas se dan de alta individualizadamente. El robot es capaz de identificar las etiquetas META de palabras clave ("keywords") y de descripción existentes en las páginas utilizando la primera para la indización de la página y la segunda como resumen del contenido en la visualización de resultados.

Búsqueda: la búsqueda se puede realizar utilizando cualquiera de los dos formularios existentes: uno para búsquedas simples y otro para las más complejas. El sistema permite localizar palabras y frases que pueden ser combinadas con los operadores AND, OR y NOT, pero no se puede utilizar paréntesis. Se pueden truncar términos. Las búsquedas pueden hacerse en texto libre o bien limitarse a los siguientes campos: URL, "hyperlinks", título y lugar geográfico.

Resultados: los resultados son presentados siguiendo un orden de relevancia primándose aquellas páginas en las que los términos requeridos aparecen en el título o cerca del inicio, contienen un mayor número de ocurrencias o los términos solicitados son considerados relevantes (palabras poco comunes en una base de datos). De cada página seleccionada aparece el título con el enlace, la URL, la descripción (la que aparece en la etiqueta META "description", en el caso de que exista, o las primeras líneas que aparecen en la página propiamente dicha), el tamaño de la página (bits) y la fecha de alta en el sistema.

Ayudas: en las ayudas se detalla convenientemente como realizar una búsqueda y los criterios de presentación de recursos.

Otros: también presenta una sección de páginas clasificadas directamente por un equipo de personas de Infoseek.

Positivo: 1) Reconoce las etiquetas META de palabras-clave y de descripción. 2) Posibilidades de búsqueda bastante completas: truncado, búsqueda de términos compuestos y operadores booleanos (AND, OR, NOT). 3) Posibilidad de buscar por los siguientes campos: URL, "hiperlink", título y lugar geográfico. 4) Las páginas recuperadas se presentan siguiendo un orden de relevancia. 5) Se utiliza el texto de la etiqueta META de descripción para resumir el contenido de cada página recuperada. 6) Detección de novedades por la fecha de alta en el sistema. 7) Las ayudas son bastante completas. 8) Versión en español.

Negativo: 1) El alta debe hacerse página a página. 2) Las páginas se encuentran parcialmente clasificadas. 3) No permite el uso del paréntesis. 4) No se puede buscar ni en el campo de palabras-clave ni en el de descripción. 5) No se controla el vocabulario. 6) Existe un único formato de presentación. 7) Se recuperan páginas y no recursos que además no pueden ser agrupadas por servidores. 8) Se conoce la fecha de alta de las páginas, pero no se puede acotar por este criterio.

5. METABUSCADORES

Los metabuscadores realizan búsquedas paralelas en varios buscadores a la vez. Ejemplos: Metacrawler, Metabusa, Inference, Altheweb, Dogpile, Search.com, Starting Point... (el orden de los resultados es un problema en los metabuscadores ya que cada buscador emplea el suyo propio).

Metacrawler busca a la vez en los siguientes buscadores: LookSmart, Findwhat, Direct Hit, Atavista, Internet Keyword.

La gran cantidad de información y el notable aumento de buscadores accesibles en la Red acabó en la posibilidad de realizar diferentes búsquedas con diversos buscadores simultáneamente con una única estrategia de búsqueda. Este hecho generó bastante expectación, ofreciendo una gran cantidad de posibilidades interesantes para los que querían introducirse en una materia. Con una búsqueda no muy precisa y sin pretensiones de selectividad en el resultado, estas herramientas permiten no tener que realizar en diversas ocasiones el mismo proceso con diversos buscadores y con direcciones diferentes. A pesar de esto, este concepto de exhaustividad no es infalible ni perfecto; cada buscador tiene sus características específicas y esto dificulta su integración. El metabuscador basa su búsqueda en los buscadores que tiene disponibles y que previamente han acordado y solicitado el permiso pertinente. Por lo tanto, la simplicidad de estos metabuscadores contará con una variable importante, que es el popular ruido documental en relación con el resultado final. El tiempo de espera también es un factor importante que algunos permiten precisar. Hay que tener en cuenta que si el tiempo definido es demasiado corto el resultado seguramente no será tan exhaustivo.

Por esta razón, también se creó y diseñó otra tipología de metabuscadores con otra filosofía: el acceso directo a diversos buscadores simultáneamente pero con un único interfaz. Algunos ofrecen a partir de una previa selección de índices temáticos, unos motores especializados que se adaptan al perfil de la consulta, mejorando la eficacia de la búsqueda.

A continuación se ofrece una lista de los metabuscadores más populares:

All-in-One Search Page <http://www.allonesearch.com/all1srch.html>

Metacrawler <http://www.metacrawler.com/index.html>

Dogpile <http://www.dogpile.com>

All4one <http://www.all4one.com>

6. VENTAJAS E INCOVENIENTES ENTRE ROBOTS, BUSCADORES Y LISTAS DE ENLACES

En el caso de las listas de enlaces el primer problema es el de elegir a cuál de ellas recurrir dada su extrema popularidad. Prácticamente cada departamento universitario, o incluso investigadores a título individual, ofrecen listas de enlaces recomendados a su área de investigación. En general, cuando contienen una elaboración cuidada y actualizada, suelen ser de una alta especialización, en campos muy concretos, con enlaces a los recursos claves de esa materia. Su gran inconveniente es que ofrecen una información muy limitada sobre el recurso, normalmente su título, no acostumbran a disponer de sistemas de búsqueda o clasificación de los recursos, y su mantenimiento habitualmente no es constante ni depurado.

En cuanto a los buscadores su principal inconveniente es la calidad de los ítems recuperados. Aunque en general los métodos de ordenación de los resultados siguen un valor de relevancia, el usuario se enfrenta a listados de miles de documentos. Su presentación consiste en una mezcla del texto extraído de determinadas etiquetas del código fuente html del recurso, sin ningún tipo de coherencia, y de los metadatos, cuando, en el mejor de los casos, el autor lo ha descrito, que no siempre es así.

En este último caso surge el problema de la autenticidad y valor de los datos aportados. Se ha detectado el uso por parte de algunos autores de aquellas palabras que más utilizan los usuarios de robots en las búsquedas, tales como sex, Microsoft, etc., para describir sus documentos, aunque no tengan ninguna relación de contenido temático, con el único propósito de atraer visitantes a sus páginas.

Por otro lado, los robots indizan páginas HTML, no recursos. La diferencia entre unas y otras es importante: un recurso, como por ejemplo un artículo en una revista electrónica, puede consistir en una serie de páginas HTML enlazadas entre sí a través de una página de contenidos. El robot no indizará la unidad lógica que es el artículo, sino la unidad física que es la página, o, en algunos casos, una muestra aleatoria de páginas.

Los buscadores y metabuscadores no son la solución a los problemas que plantea el control de la información en la red, necesario, por otro lado, para la identificación y localización. Las páginas Web tienen una vida media muy corta (de 44 a 75 días), proporcionan una localización volátil (mudanza de servidor), se producen cambios con frecuencia (el 30% de ellas lo hace cada 10 días. El HTML carece de la expresividad para la representación de la información estructurada.

Conscientes de los problemas planteados, algunas bibliotecas han comenzado a aplicar su experiencia en el campo de la descripción de recursos electrónicos mediante la construcción de las denominadas bibliotecas virtuales o pasarelas temáticas (subject gateways), que no son otra cosa que descripciones detalladas, y siguiendo unos estándares internacionales, de los recursos existentes en una disciplina dada.

7. BIBIOTECA DIGITAL

7.1 DEFINICIONES Y CARACTERÍSTICAS

En los últimos años estamos asistiendo a una verdadera “revolución tecnológica” a todos los niveles. Cada día “la explosión de la información” nos sorprende con nuevos cambios y avances que solo unos meses antes hubiéramos considerado imposibles. De ahí que, en el mundo de las bibliotecas, tengamos que plantearnos nuevos conceptos y metas, entre ellos la aparición de lo que se conoce como la biblioteca digital.

Por su novedad, el término “biblioteca digital” provoca una impresión diferente en cada lector. Para algunos es simplemente la automatización de los catálogos bibliográficos de bibliotecas convencionales. Para otros es una colección de imágenes que resultaron de la digitalización de documentos en papel. En muchos artículos publicados sobre el tema se usan como si fueran sinónimos, los términos Biblioteca electrónica y biblioteca virtual (<http://sunsite.berkeley.edu/mydefinitions.html>). Esta confusión se debe, en parte, a que, la investigación sobre el uso de las nuevas tecnologías en bibliotecas ha experimentado un enorme crecimiento en años recientes.

Hemos encontrado multitud de definiciones acerca de lo que es una biblioteca digital, desde las que se pierden en precisiones técnicas y elaboradas bajo una perspectiva más informática, hasta las que buscan un carácter más “bibliotecario”, de entre todas ellas podríamos destacar las siguientes:

- “Una biblioteca digital es una biblioteca que tiene materiales de forma digitalizada”
- “una biblioteca digital, es un ensamblaje de cómputo digital, mecanismo de comunicación y software necesario para reproducir, emular y extender los servicios propuestos por bibliotecas convencionales como son la colección, catalogación, búsqueda y disseminación de la información”

- “Una biblioteca digital es una biblioteca que mantiene toda o una parte importante de su colección en un formato que puede ser procesado en forma automatizada como una alternativa, suplemento o complemento a los materiales tradicionales en papel y micropelícula que existen en la biblioteca”
- “Bibliotecas digitales son organizaciones que proveen los recursos, incluyendo personal especializado, para escoger, estructurar, ofrecer acceso intelectual, interpretar, distribuir, preservar la integridad y asegurar la perdurabilidad a través del tiempo de obras digitalizadas (imágenes) de forma que estén disponibles para su uso de forma fácil y económica por una comunidad o un grupo de comunidades de usuarios”

Como se puede observar existe cierta disparidad de criterios; ¿no sería mejor hablar de Colecciones digitales?, de acuerdo con la Association of Research Library (ARL) de Estados Unidos, la mayoría de las definiciones de biblioteca digital coinciden en los siguientes elementos:

- Por lo general no es un ente aislado, sino que está integrado por diferentes colecciones de fondos creados y administrados por diferentes organizaciones.
- Requiere tecnologías específicas para compartir enlazar recursos dispersos
- Los enlaces entre diversos fondos y servicios de información deben ser transparentes para el usuario
- Los fondos contienen objetos digitales no necesariamente distribuibles en formato impreso
- El acceso universal a sus fondos y servicios de información son una meta obligada

Ante las múltiples posibilidades de definición de lo que es una biblioteca nos decantaremos por una bien sencilla, una biblioteca es un sistema diseñado para recopilar, almacenar, organizar y preservar información de diversos tipos y formatos (microfilm,

papel, etc.), para ponerlos al alcance de una comunidad. Una biblioteca digital es por tanto una biblioteca donde la información es almacenada y procesada en formato digital pero donde también cabe la posibilidad de que la biblioteca ofrezca recursos externos en formato electrónico, independientemente de quien los genere. De esta forma la biblioteca digital sería quien los unifique y lo ofrezca mediante una interfaz adecuada a sus usuarios.

Una biblioteca digital es vista popularmente como una versión electrónica de una biblioteca tradicional, pero la sustitución del papel por recursos en formato electrónico, trae consigo grandes diferencias como pueden ser:

- El almacenamiento y recuperación de información en libros, revistas, periódicos, discos, videos... etc., en forma digital
- La interacción directa por parte del usuario con la información mediante una interfaz adecuado, con el fin de obtener el material deseado.
- La copia del material en versión original
- El uso y disseminación de la información por parte del usuario y todos los problemas (Copyright, licencias...) que conlleva
- El intercambio de puntos de vista, experiencias e información entre usuarios en forma electrónica mediante foros de discusión, debates, etc.

Algunas características que distinguen a las bibliotecas digitales de las convencionales son las siguientes:

- Proporcionan información en cualquier tiempo y en cualquier lugar
- Proporcionan acceso a colecciones de información multimedia, construidas mediante la integración de texto, imágenes, gráficas, audio y vídeo
- Permiten el ingreso y la información personalizada de la información, incluyendo soporte para recuperar sólo información relevante
- Mejoran las actividades de colaboración, tales como la investigación, el aprendizaje y el diseño, al reducir las barreras del tiempo y la distancia

Las ventajas de una biblioteca digital sobre una biblioteca física, son similares a aquellas que cualquier base de datos digital tiene sobre una en papel:

- Rápido acceso a la colección de datos
- Mejoramiento en la funcionalidad de revisión y búsqueda
- Rápida distribución desde puntos de creación y los puntos de almacenamiento y los puntos de uso

Como se puede observar las diferentes definiciones fluctúan desde interpretaciones más o menos sencillas, hasta otras mucho más complejas: Pero antes de seguir adelante con las bibliotecas digitales sería interesante intentar aclarar algunas cuestiones relativas a algunos conceptos que surgen cuando se habla de este tipo de bibliotecas.

Según Claudia Paz Yanes (2001) las bibliotecas digitales participan de las siguientes características:

- **Mundialización.** Cualquiera en cualquier momento puede acceder a un catálogo, a una base de datos o a una página institucional desde un punto geográfico distinto del de la ubicación original. Salvado cuestiones como licencias de uso, derechos de autor etc., que el usuario generalmente ignora, una biblioteca digital es una ventana a la información de este mundo. Sobre las implicaciones sociales, educativas, económicas, etc., no hay consenso y se están estudiando. La homogeneización resultante de este contexto mundial no se hace esperar: predominio del inglés como lengua de gran difusión, productores y distribuidores de los servicios digitales de origen anglosajón, etc.
- **Integración.** Los recursos digitales ofrecidos por las bibliotecas tienden cada vez más a su integración, que no es lo mismo que concentración. No se trata de juntar elementos varios sin nexos de unión, sino de que todos ellos formen parte de un conjunto íntegro. Así, una biblioteca digital proporciona información al usuario de sus préstamos a través del catálogo; permite acceder al artículo o al libro y remite a otros muchos recursos relacionados.

- **Transparencia.** Se trata de un concepto polémico. La transparencia presume la visibilidad de la información. Así pues, las bibliotecas digitales dan por hecho que la información proporcionada es pública, es visible, es transparente a sus usuarios. La biblioteca digital cuenta con un gran aliado o enemigo: el buen o mal uso de la tecnología que la sustenta. De ese correcto uso se derivará una mayor transparencia o no.
- **Instantaneidad.** En el mundo de la rapidez y la velocidad actual en que vivimos, las bibliotecas digitales son uno de los mayores ejemplos de este principio. La tendencia a la integración y mundialización de los recursos, se ve coronada por un acceso rápido, resultados de esta era de las telecomunicaciones.
- **Ubicuidad** o capacidad de estar en varios sitios a la vez. Se trata de una ubicuidad virtual: en vez de que el usuario sea el que se mueva, éste acerca la biblioteca a su lado. La hace presente en la máquina a través de la tecnología, seleccionando la información a la que tiene acceso de una manera personalizada donde archiva, refunde, copia y adapta lo que su interés le marque.
- **Virtualidad.** La virtualidad provoca la accesibilidad documental sin barreras.

Merlo y Sorli proponen una clasificación de las bibliotecas digitales en función de determinadas características:

- Según el almacenamiento de la información: un único servidor, distribuida
- Según la temática: multidisciplinarios, especializadas
- Según el ámbito geográfico: generales, locales
- Según la institución: privadas, públicas, mixtas. Universitarias
- Según la institución: privadas, públicas, mixtas, universitarias
- Según el acceso: público, restringido, semipúblico

Los dos modelos de bibliotecas digitales más importantes son la universitaria y la pública.

La biblioteca digital universitaria, tiene bien acotado el perfil de usuarios para el que dispone de servicios. Así, debe implicarse en el proceso de enseñanza-aprendizaje y jugar un papel, cada vez mayor, no sólo entre profesores y alumnos sino también con cada uno de ellos por separado. Para ello ha de facilitar la conversión de la información en conocimiento (gestión del conocimiento) y, ofertar una serie de servicios concretos: SOD (Servicio de Obtención de Documentos); DSI (Difusión Selectiva de la Información); formación de usuarios, alerta documental; consulta de catálogos; acceso a bases de datos; enlaces a recursos Web...

Respecto a la biblioteca pública digital, debe servir a cualquier usuario haciendo especial incidencia en el usuario local. Las funciones que asume son:

- Como centro local proveedor de información: guías de Internet, guías de lectura, recursos en Internet, guías divulgativas.
- Como centro generador y proveedor de información local y especializada
- Servicios vinculados a la biblioteca pública: servicio virtual de obtención de documentos, servicios virtuales de información bibliográfica, acceso a los catálogos de la biblioteca

6.2 BIBLIOTECA DIGITAL: MITOS

El deslumbramiento causado por las nuevas tecnologías y la carencia de una reflexión crítica ha conducido a la aparición, en torno a las bibliotecas digitales, de una serie de mitos que a menudo se toman por futuras realidades: la gratuidad del acceso a la información, la desaparición del formato papel como soporte de la información o la creencia de que todo está en Internet son algunos de ellos. Según Ángel Borrego (2001) nos encontramos con una serie de mitos, entre los cuales destacan:

- **Gratuidad del acceso a la información**

Tal y como están las cosas no parece que la información digital vaya a ser, no ya gratuita, sino ni siquiera más barata de lo que lo era en papel. La información digital que está accesible gratuitamente - básicamente la ofrecida por entidades sin ánimo de lucro y administraciones públicas - es la misma que ya se ofrecía sin costes en formato papel. Además, algunas de las iniciativas de las asociaciones de editores apuntan a un futuro que nos lleva al “pago por visión”, es decir, al desembolso de una cantidad de dinero cada vez que se consulta un documento digital. De esta forma la información no sólo no se abarataría sino que su consulta devendría cada vez más cara y podría darse la paradoja de que el nuevo medio que debería ampliar las posibilidades de consulta de la información, en realidad limitase severamente el acceso a la misma.

En el futuro se mantendrán los elevados costes que en la actualidad requiere el mantenimiento de las bibliotecas físicas, que serán complementadas por las colecciones digitales pero no sustituidas por ellas. Así, en un contexto en el que los presupuestos bibliotecarios apenas se incrementan, además de hacer frente a las partidas tradicionales, habrá que destinar cada vez más recursos a las nuevas tecnologías. No hay que olvidar los elevados costes en infraestructuras, equipos y programas - cuya actualización requiera un desembolso constante - a los que tendrán que hacer frente las bibliotecas para afrontar el futuro digital. La elevada cuantía de estas inversiones reduce la posibilidad de adquirir documentos convencionales y limita severamente las posibilidades de acceder a la información tanto en nuestro entorno - donde las conexiones a menudo son todavía escasas, caras y lentas - como especialmente, en los países en vías de desarrollo.

- **Desaparición del formato papel como soporte de la información**

Tampoco parece factible que el formato electrónico vaya a sustituir, y ni siquiera a desplazar parcialmente, al papel. De hecho, resulta curiosa la

paradoja que supone la enorme cantidad de textos escritos que ha generado el debate sobre el fin del libro en su formato actual. Sin embargo, productos como el libro son tecnológicamente muy avanzados: su lectura no requiere de ningún tipo de fuente de energía, ofrece una elevada resolución, es fácilmente manejable etc. Por esta razón no resulta probable que sea sustituido, al menos en un futuro inmediato, por las nuevas versiones digitales que, en distintas variantes, aparecen cada poco tiempo. Resulta que lo nuevo no siempre es mejor, de la misma forma que un producto relativamente antiguo, como el libro impreso, continua cumpliendo plenamente su labor como vehículo transmisor del conocimiento.

- **Todo está en Internet**

Otro de los mitos más extendidos en torno a las bibliotecas digitales es el de que todo está en Internet. El empleo de fuentes electrónicas ha aumentado las expectativas de los usuarios, hasta el punto de que muchos de ellos llegan a creer que cualquier información que necesiten la encontrarán en la red. Sin embargo, no es cierto que todo el conocimiento esté en Internet ni es probable que lo esté en un futuro inmediato. Las razones estriban en la lentitud y elevados costes de los procesos de digitalización y en el hecho de que los documentos impresos que hoy manejamos están sujetos a unos derechos de autor que no pueden ser obviados. Sin embargo, la creencia de que en Internet se encuentran las respuestas a todas nuestras preguntas ha llevado a que se reduzca la consulta de fuentes en papel – puesto que se entiende que no aportarán nada nuevo a la información disponible en línea – e, incluso a que se obvie la posibilidad de acudir a fuentes electrónicas que no suministran el texto completo de los documentos.

- **Las bibliotecas digitales son universalmente accesibles, desde cualquier lugar, en cualquier momento**

Queda mucho por hacer convertir este mito en realidad. Podemos aceptar que una red global de ordenadores – Internet o alguno de sus sucesores – va a ser en el futuro el mecanismo principal de distribución de información

digital, pero es difícil admitir que en la actualidad se da ya un acceso igual para todos, pues se olvida con frecuencia que Internet no está tan extendido. Hay relativamente pocas conexiones fuera de las zonas más densamente pobladas, los costes de acceso todavía son altos y, para la inmensa mayoría de la población mundial que vive en los países en vías de desarrollo, disponer de un acceso fácil a Internet es como un sueño. Es más, las conexiones que realmente existen son lentas para la mayor parte de los usuarios.

Puede que en el futuro los complejos recursos y sistemas multimedia exijan una maquinaria (hardware) y unos programas (software) tan especializados que sólo un número limitado de terminales estén preparados para el acceso a la información. Los límites en el ancho de banda de la red y las velocidades lentas de transmisión podrían suponer problemas importantes para el acceso eficaz a la información por parte de muchos usuarios.

Si consideramos las inmensas trabas técnicas y legales que deben solucionarse, resulta evidente que aún es problemático ese acceso equitativo a los recursos de las colecciones y los servicios digitales. La reforma de los derechos de autor va a ser un proceso lento, y nadie puede asegurar que no se vaya a concretar en un sentido que haga imposible las bibliotecas digitales. Las tecnologías que conectan los ordenadores y las que permiten el almacenamiento y tratamiento de la información son variables dinámicas. Lo que sí es seguro es que la tecnología para la gestión de las bibliotecas digitales se está volviendo más compleja, como también la administración de las autorizaciones y del acceso de los usuarios. Todos estos factores pueden afectar negativamente la igualdad en el acceso.

- **El supuesto fin de la función intermediadora del bibliotecario**

Otro mito que se ha desarrollado alrededor de la biblioteca digital es el de la desaparición de las funciones intermediadoras del bibliotecario, especialmente el de referencia, en el entorno digital.

El desarrollo de las nuevas tecnologías ha supuesto importantes cambios, tanto cuantitativos como cualitativos, en los servicios bibliotecarios de referencia. Cuantitativos porque han puesto a nuestra disposición la posibilidad de acceder a una enorme cantidad de recursos informativos en una variada gama de temas. Y cualitativos porque han permitido que el acceso a todos estos conocimientos no sea patrimonio exclusivo de los bibliotecarios o profesionales de la información, que durante mucho tiempo han detentado el acceso a ésta, sino que han facilitado que los usuarios finales se hayan lanzado a la búsqueda de las informaciones que les resultan de interés, con diferentes resultados en función de sus distintas habilidades.

La facilidad de acceso a los recursos informativos ha llevado a los usuarios a obviar al bibliotecario, cuyo papel como intermediario entre el usuario y la información parece haber quedado dañado. Se estaría en camino de llegar a la situación de que cada usuario sea su propio bibliotecario de referencia. Parece que el modelo al que estamos irremediabilmente abocados es el de un usuario autosuficiente que, gracias al elevado grado de amigabilidad de los sistemas y la disponibilidad de sistemas de ayuda que puede utilizar cuando encuentre dificultades, puede acceder a la información que necesita sin ningún tipo de intermediario. Obviamente, es bueno que los usuarios adquieran habilidades en la búsqueda de información que les permitan independizarse del bibliotecario. Sin embargo, la aparición de nuevas herramientas de búsqueda y la mejora de las habilidades del usuario en la obtención de información no tienen porque implicar una desaparición de la labor intermediadora del bibliotecario.

Para entender por qué es necesario que exista una persona detrás de un servicio de referencia es conveniente, en primer lugar, comprender la naturaleza de este servicio. En muchas ocasiones el verdadero objetivo del bibliotecario de referencia no es tanto el de proporcionar respuesta a las preguntas planteadas por los usuarios, como el de eliminar el bloqueo que les impide solventar sus necesidades de información. En efecto, a menudo el usuario llega sin saber qué es lo que realmente desea y la labor del

bibliotecario consiste en guiarle en el proceso de determinar cuál es su necesidad de información, una tarea que difícilmente puede ser llevada a cabo por una máquina. Cuando el usuario no es capaz de describir acertadamente su necesidad de información resulta difícil que pueda dar unas órdenes medianamente coherentes al sistema de recuperación de la información. En última instancia, los sistemas automatizados podrían ser válidos para atender a aquellos usuarios que plantean cuestiones simples o incluso de una dificultad media. Sin embargo, parecen poco eficaces a la hora de atender consultas más complejas por la forma confusa en que son planteadas.

Por otro lado, la dificultad a la hora de dar respuesta a una demanda de información no depende únicamente de la sencillez o complejidad de ésta, sino que se ve influenciada en gran medida por las circunstancias del usuario. Por esta razón es posible que una misma necesidad informativa, planteada por diferentes personas, dé lugar a respuestas diferentes. Difícilmente podrá el usuario final plantear cuestiones de este tipo a una máquina, del mismo modo que resulta extremadamente complicado creer que ésta será capaz de interpretarlas.

La posibilidad que tiene el usuario final de acceder a fuentes de información electrónicas no sólo no ha eliminado la exigencia de que siga existiendo la figura del bibliotecario de referencia sino que ha acrecentado la necesidad de que desarrolle algunas de sus competencias, especialmente en el terreno de la formación de usuarios. En definitiva, no sólo no desaparece la necesidad del bibliotecario, sino que ésta se acentúa ya que la cantidad de información aumenta y los sistemas que permiten su recuperación no se desarrollan a la misma velocidad, por lo que encontrar información pertinente deviene cada vez más complicado. Al mismo tiempo, es necesario tener en cuenta que el trabajo del bibliotecario de referencia no se limita a la búsqueda de datos y que sus tareas en los ámbitos de la orientación y la formación ganan en importancia. Por un lado debe ayudar al usuario a determinar la naturaleza de su necesidad en función de sus circunstancias. Al mismo tiempo, es

necesario que forme a los usuarios en el manejo de las herramientas de búsqueda y en la valoración crítica de las informaciones obtenidas. La existencia de profesionales que lleven a cabo estas tareas continúa siendo primordial.

7.3 BIBLIOTECA DIGITAL: RETOS

La creación de bibliotecas digitales que sean “eficaces” plantea serios retos a la tecnología existente y futura. La integración de los medios digitales en las colecciones tradicionales no será directa, como ocurrió con los precedentes medios nuevos (por ejemplo, el vídeo o las cintas de audio), y ello debido a la naturaleza singular de la información digital: es poco fija y fácil de copiar; múltiples usuarios pueden acceder a ella simultáneamente desde lugares bien remotos. Procesos de la biblioteca tradicional tales como la referencia y el desarrollo de la colección, aunque constituyen una base potencial para el trabajo de la “biblioteca digital”, tendrán que ser revisados y mejorados de modo que puedan dejar espacio para estas diferencias (Kuny, Terry y Cleveland, Gary (2001))

Retos:

- **Descubrimiento de recursos**

Si por algo se caracteriza la información en Internet es por el hecho de que los documentos digitales pueden tener varios formatos y estar posiblemente en diferentes versiones y en lugares que aún no son fijos. Puede ocurrir que un documento o recurso que se encuentra en un lugar de la red hoy, mañana haya desaparecido.

Las personas encargadas del desarrollo de los servicios de indización se esfuerzan por mejorar sus capacidades, pues de otro modo no pueden hacer frente a la creciente masa de información que hay en la red, pero deben encontrar para ello nuevas fuentes de financiación. En la actualidad confían en la publicidad, pero no hay garantías de que en el futuro un sistema como éste, que provee un servicio de mucha calidad para el público en general, sea rentable. Es razonable esperar que se establezcan niveles de servicio. En el

nivel básico, todo el mundo accede a la base de datos sin cargo alguno, pero los resultados son de utilidad limitada y la buena información se ve desbordada por una riada de otra mala. La búsqueda no resultará eficaz. En los niveles superiores, los usuarios que pueden pagar por mejores servicios se benefician de unas posibilidades de búsqueda muy acrecentadas, algoritmos que jerarquizan mejor la información o una velocidad de acceso superior.

Hay también cuestiones importantes del tipo de si las tecnologías existentes pueden adaptarse suficientemente al crecimiento en el número de usuarios y en el volumen de la información que circula por Internet. Los usuarios están descubriendo que la información de calidad que consiguen recuperar en las grandes bases de datos puede quedar perdida en medio de una riada de resultados irrelevantes.

Los bibliotecarios organizan el conocimiento a través de los procesos de clasificación y catalogación y crean así información sobre la información, algo que recibe también el nombre de “metadatos”. Resta aquí mucho por hacer si se quieren desarrollar métodos para identificar y recuperar la información que sean consistentes y únicos, y que sirvan para cualquier información, con independencia del formato que tenga o de dónde se encuentre.

Los bibliotecarios tienen otro papel en la asistencia a quien busca recursos, sin apenas competencia por parte de las tecnologías digitales, pues sólo ellos son capaces de proveer servicios “de confianza”. Los bibliotecarios son poseedores de confianza y autoridad porque hacen elecciones y evalúan la información como parte del desarrollo de una colección, siempre con una comprensión completa de lo que los usuarios necesitan.

- **Desarrollo de la colección digital**

Los bibliotecarios tienen una experiencia considerable en digitalización, aunque la profesión ha tendido a llamarla de otro modo. La “conversión retrospectiva” de las fichas de biblioteca impresas en asientos de un

catálogo que pudieran leer los ordenadores, representa un esfuerzo temprano de digitalización. La conversión del papel a un formato digital es cara y costosa.

Algunos soportes son más fáciles de digitalizar que otros. Las colecciones de fotografías, los recursos bibliográficos, las series estadísticas e incluso las publicaciones periódicas admiten la digitalización, mientras que otros materiales, por ejemplo los mapas y los libros, lo harían en menor medida. Nos queda mucho por conocer todavía sobre lo que necesitan los usuarios, sobre cómo usan la información y sobre si los formatos digitales responden eficazmente a sus necesidades. Llevar a cabo grandes iniciativas de digitalización sin una comprensión fundamental de las necesidades de los usuarios es como comenzar la casa por el tejado. Lo digital y la utilidad no siempre coinciden.

- **Conservación**

Los medios digitales son “frágiles” y tienen una vida limitada. La información digital que se almacena en estos medios, incluso aunque perdure, puede resultar ilegible debido a la obsolescencia de la tecnología, pues no hay que olvidar que ésta evoluciona con rapidez y que los sistemas antiguos, al desaparecer, se llevan consigo la posibilidad de leer la información que gestionaban.

Las bibliotecas digitales, para conservar sus fondos digitales, se verán constantemente obligadas a ir convirtiendo la información a las nuevas configuraciones de la maquinaria y los programas existentes. La obsolescencia tecnológica, la necesidad de convertir la información digital y algunos problemas legales y de organización están marcando “los límites de la tecnología digital”. No hay patrones de conservación para la información digital.

- **La administración de la biblioteca digital**

La gestión de la infraestructura técnica para los servicios de la biblioteca digital se constituirá en un obstáculo de importancia para muchas bibliotecas, en especial a medida que los presupuestos se restrinjan mientras que los costes de desarrollo y mantenimiento de las colecciones van incrementándose.

- **Derechos de autor y autorizaciones**

Si las bibliotecas comienzan de verdad a recoger sistemáticamente información digital en gran escala, no va a ser empeño fácil asegurar un acceso eficaz a los recursos. De hecho, los derechos de autor pueden acabar por hacer que a muchas bibliotecas se les quite la idea de proveer un acceso abierto a la información digital que recopilan. Ciertamente hay que encontrar una solución para la creación y distribución de la información digital a través de las bibliotecas digitales, pero esa solución debe ser vista como justa y razonable tanto por parte de los productores de la información como por parte de sus consumidores. Los derechos de autor podrían convertirse en una barrera infranqueable para el desarrollo de las colecciones digitales.

Los miembros de la comunidad bibliotecaria se encuentran cada día más preocupados ante la idea de que los cambios en los derechos de propiedad afecten adversamente las capacidades de las bibliotecas para proveer colecciones y servicios digitales.

- **Coste**

La información no ha sido nunca realmente gratis. Hay siempre un coste en su creación, producción y diseminación. Las bibliotecas públicas de libre acceso, financiadas con los impuestos, ocultan en buena medida a los usuarios los costes reales de la información, lo cual es correcto allí donde las bibliotecas son consideradas un bien público. Las bibliotecas digitales introducen nuevas e inciertas realidades económicas y relaciones en el interior de las bibliotecas. En el sistema anterior, se ocultaban a los

“usuarios” los costes del acceso a la información; la era digital, por contraste, necesita consumidores que paguen cuando deseen acceder a los fondos y los servicios digitales.

Las bibliotecas digitales podrían convertirse en fondos y servicios de propiedad privada accesibles mediante suscripción, pago por uso, tarifas, etc.

La dura realidad económica es que el desarrollo de fondos digitales implica costes de implementación, autorizaciones, promoción y desarrollo y mantenimiento de una infraestructura técnica.

7.4 CÓMO CREAR UNA BIBLIOTECA DIGITAL

Según Nuria Lloret (1999) crear y organizar una biblioteca digital es un trabajo de gran envergadura, que implica pensar en diversos aspectos de carácter económico, técnico, administrativo, legal, de seguridad, entre otros. Se requiere primero conocer las necesidades de nuestros usuarios y analizar luego nuestras posibilidades para realizar el trabajo arduo y permanente que significa crear la biblioteca digital.

Tanto Nuria Lloret (1999) como Hípola, Vargas-Quesada y Senso (2000) coinciden en mencionar que hay que tener en claro el trabajo y costos que significan poner en marcha una biblioteca digital, ya sea que la biblioteca exista y por lo tanto hay que realizar el proceso de reconversión y digitalización, o que la biblioteca “nazca” en cuyo caso hay que considerar la adquisición de documentos digitales para formarla.

De acuerdo a Nuria Lloret (1999), es necesario considerar los siguientes aspectos para tomar decisiones concretas:

- **Aspectos técnicos:** entre los que se incluye la colección (en soporte electrónico o digitalización de las existentes), la tecnología para la digitalización, la capacidad de almacenamiento del sistema, las necesidades de hardware y software, y los sistemas de seguridad tanto para la información como para el usuario.
- **Aspectos económico-legales:** en donde hay que considerar los diversos costos en los que se va a incurrir, como la infraestructura, la digitalización,

las suscripciones electrónicas, la actualización del sistema, así como diversas fuentes de financiamiento y la posibilidad de conformar consorcios para compartir gastos. Mientras que en los aspectos legales hay que tomar en cuenta los temas de copyright y derecho de autor, seguridad de la información, acceso libre o pago, etc.

- **Aspectos de accesibilidad del usuario:** donde se incluyen consideraciones relacionadas con el diseño de la interfaz del usuario, la normalización de los datos para facilitar su recuperación, la prueba del sistema y opciones de retroalimentación, y la capacitación o autocapacitación del usuario.

Metodología para la creación de una biblioteca digital contemplaría las siguientes fases:

- **Planificación de la biblioteca digital:** se inicia el proyecto estableciendo con detalle los objetivos y requerimientos de la biblioteca digital así como las expectativas de la misma. Es decir, se requiere la elaboración de un proyecto específico que incorpore requerimientos técnicos, económicos, legales, costos y cronograma; y la preparación de una propuesta detallada con especificaciones de personal, hardware y software necesarios, almacenamiento, procesos, costos, etc., además de características de contenido, estructura y servicios que ofrecerá la biblioteca virtual que se pretende formar. La planificación del proyecto debe comenzar con la selección de un equipo de trabajo conformado tanto por bibliotecarios como informáticos, que elaboren un diagnóstico de la realidad y preparen una propuesta.
- **Puesta en marcha de la biblioteca digital:** en esta etapa se delimitarán las características específicas de la biblioteca digital, como la estructura, contenidos, servicios a ofrecer, interfaces de acceso y búsqueda, etc., además de señalar los procesos a realizar para poner en marcha la biblioteca virtual. Entre estos procesos se cuentan los de desarrollo de colecciones (selección, adquisición de material), digitalización, procesamiento técnico (indización),

adecuación de material existente, así como servicios específicos de búsqueda bibliográfica, préstamo interbibliotecario, disseminación selectiva, etc. Es probablemente la etapa que ocupará la mayor parte del tiempo previsto en el cronograma, especialmente en el caso de una biblioteca digital que se “hace”, pues existe material bibliográfico, bases de datos, servicios, etc. Físicos o tangibles que deberán adaptarse a los requerimientos de la biblioteca digital.

- **Lanzamiento y mantenimiento de la biblioteca digital:** esta última parte del proyecto se concentra en los aspectos externos y el desarrollo futuro de la biblioteca digital. Es decir, la etapa del diseño o forma de la biblioteca virtual, incluyendo apariencia interna y externa de las pantallas, texto, imágenes, etc. Igualmente se crearán los formatos que el usuario necesitará para solicitar los servicios que ofrece la biblioteca, así como facilidades de capacitación o autocapacitación, a fin de que el usuario obtenga el mejor provecho de ella. Con relación al mantenimiento, se recomienda continuar con el equipo de trabajo inicialmente considerado para el proyecto, incorporando adicionalmente artistas gráficos, publicistas, psicólogos, y otro personal que se requiera, tanto para el diseño como para la actualización y creación de nuevos servicios en la biblioteca digital. Será necesario contar igualmente con mecanismos de retroalimentación que permitan modificaciones acordes con los requerimientos de los usuarios. Es conveniente una encuesta u otro medio formal de evaluación inmediatamente después del lanzamiento de la biblioteca digital y otro medio permanente de evaluación y comentarios que aseguren un mantenimiento adecuado de la misma.

7.5 EJEMPLOS DE BIBLIOTECAS DIGITALES

American Memory. Biblioteca Digital Nacional de Estados Unidos
<http://memory.loc.gov>

Principal fuente de materiales relacionados con la historia de estados Unidos. Tiene más de 7 millones de documentos digitalizados procedentes de más de 100 colecciones históricas. Proporciona acceso a textos, vídeos, audio, mapas, imágenes, partituras. Proporciona acceso a los programas necesarios para su visualización o su audición en caso de no tenerlo instalado en el ordenador desde donde se consulta.

Biblioteca Digital Dioscórides (UCM) <http://www.ucm.es/cgi-bin/show-libros>

Biblioteca Digital en Ciencias de la Salud. Proporciona acceso a un fondo histórico biomédico, de los siglos XV al XVIII, existente en las Bibliotecas de la Universidad Complutense de Madrid.

Sus objetivos son: conservar el patrimonio bibliográfico biomédico; acceso a este patrimonio para el mayor número de usuarios posibles; proporcionar una herramienta avanzada para mejorar la actividad investigadora.

Biblioteca Digital Joseph Laporte <http://www.fbjoseplaporte.org>

Ofrece acceso a bases de datos médicas y a una colección de revistas a texto completo. Servicio de DSI (Difusión Selectiva de la Información); boletín electrónico; atención al usuario; lista de preguntas frecuentes. En el futuro contará con un centro de documentación que ha sido acreditado como Internacional Docline Center de la National Library de los Estados Unidos para España

Biblioteca **Pública** **Digital** (Argentina)

http://www.educ.ar/educar/superior/biblioteca_digital/

Ofrece acceso público a una extensa colección de textos completos de todas las disciplinas académicas, en formatos PDF y HTML

Biblioteca “Virtual” Miguel de Cervantes <http://www.cervantesvirtual.com>

Proyecto de la Universidad de Alicante con el objetivo de digitalizar obras clásicas, de carácter hispanoamericano, a través de la aplicación de nuevos métodos de edición electrónica para grandes masas de información documental

Biblioteca Digital de Proyectos Fin de Carrera (Universidad Politécnica de Valencia) <http://bdpfc.inf.upv.es>

Proporciona acceso al texto completo de los proyectos fin de carrera que se realizan en la Universidad Politécnica de Valencia

Proyecto Gutenberg (Biblioteca Virtual) <http://promo.net/pg>

La idea de almacenar en ordenadores y poder recuperar de alguna manera todo el conocimiento universal, almacenado hasta ese momento en las bibliotecas la tuvo Michael Hart en 1971 en los estados Unidos.

La base de su teoría es que cualquier persona, en cualquier lugar del mundo, pueda acceder a un libro almacenado en un ordenador, teniendo la mínima cantidad de recursos disponibles.

Esta idea se llevó a la práctica cuando Hart copió la Declaración de la Independencia de EEUU en un archivo de un ordenador y dio origen a su Proyecto Gutenberg. En este proyecto trabajan cientos de voluntarios copinado, o escaseando páginas y páginas de material literario.

La mayor parte de esos “e-texts” (electronic texts – libros electrónicos) se han archivado en sitios FTP, en formatos de fácil recuperación, y mayoritariamente escritos en inglés. Los únicos textos, por el momento, en castellano son El Quijote y La Celestina.

Los textos tienen que descargarse al ordenador personal, lo cual es una desventaja.

La biblioteca del Proyecto Gutenberg está formada por una colección de libros electrónicos. Son ediciones sencillas, que únicamente incluyen texto, que pueden ser leídas incluso por los ordenadores más antiguos y que generalmente permiten que el contenido de un libro quepa en un único disquete de ordenador. De esta manera, la copia y distribución libre de disquetes, alentadas por el proyecto, permiten pensar en un público potencial compuesto por todos aquellos que tienen acceso a un ordenador. Otro punto importante es la gratuidad de la biblioteca. Y lo es porque se nutre de aquellos libros cuyo copyright ha expirado, por tanto pueden ser reproducidos libremente, y porque centenares de voluntarios repartidos por todo el mundo se encargan de investigar el estado del copyright de las obras y, llegado el caso, de realizar la edición electrónica de las mismas.

7.6 LA COLECCIÓN DE UNA BIBLIOTECA DIGITAL

Cuando nos planteamos como es el fondo de una biblioteca tradicional, tendemos a pensar en una ubicación física y concreta, en un lugar único en que se encuentra el documento de nuestro interés y al que podremos acceder mediante una serie de mecanismo como el préstamo, la consulta en un terminal...etc. Sin embargo la biblioteca digital será una suma de recursos con ubicaciones, almacenamientos y accesos que pueden encontrarse en puntos muy alejados e independientes.

Las bibliotecas digitales no residen en una construcción física, ni manejan objetos físicos como libros, catálogos o archivos. De hecho las bibliotecas digitales, deben trasladar los elementos físicos y conceptuales de la biblioteca física y adaptarlos al ambiente digital e incluso se deben añadir nuevos elementos digitales que no aparecen en la biblioteca física.

Desde un punto de vista computacional, la biblioteca digital está construida de componentes simples conocidos como objetos digitales. Para representar la complejidad de la información en las bibliotecas digitales, se pueden agrupar varios objetos. Este es el llamado conjunto de objetos digitales. Los diferentes tipos de materiales en una biblioteca digital pueden estar divididos en categorías. Según algunos autores la arquitectura de la información está basada en conceptos simples como: tipos de datos, metadatos estructurales, y meta-objetos (metadatos):

Un tipo de datos describe las propiedades técnicas de los datos, tales como el formato o método de procesamiento

Un metadato es un objeto que proporciona referencias para un conjunto de objetos digitales

Un metadato estructural es un metadato que describe los tipos, versiones, relaciones y otras características de los materiales digitales.

8. LA PROPIEDAD INTELECTUAL Y LA OBRA DIGITAL

Existe una gran incertidumbre sobre el futuro de los derechos de autor y derechos afines en el nuevo entorno tecnológico. Una de las cuestiones que más preocupan en todas las instancias es la de las repercusiones de la aplicación de la tecnología digital y la evolución hacia la convergencia, sobre el derecho de reproducción y más concretamente, sobre la regulación de la copia privada. Por ello, se hace patente una permanente actualización y una mayor armonización de la normativa sobre la materia, que se ha plasmado en numerosos documentos comunitarios y en los tratados internacionales más importantes sobre la materia. Por ello destacan los dos nuevos Tratados de la OMPI (Organización Mundial de la Propiedad Intelectual), el Tratado de la OMPI sobre derechos de autor y el Tratado de la OMPI sobre interpretación o ejecución y fonogramas, que muestran especial atención a la protección de los derechos de autor y derechos afines frente a los usos ilícitos en todo el mundo.

La adopción de una normativa adecuada resulta indispensable para crear un entorno favorable a la creación intelectual y a la inversión en la industria de la cultura. Esta normativa procurará la protección de la propiedad intelectual frente a la piratería. Para la creación de un entorno de seguridad jurídica tanto dentro como fuera de la Comunidad Europea, resulta de gran utilidad la regulación de la materia entre los países miembros de la Comunidad Europea. Si las condiciones de protección son equiparables, efectivas y transparentes en todos los países miembros, la seguridad jurídica que ello genere permitirá la explotación transfronteriza de los productos y servicios protegidos por derechos de autor. Estas iniciativas europeas se vienen caracterizando por mantener el elevado nivel de la tradicional protección de los derechos de autor en su ámbito a través de la adecuación de la regulación al entorno en el que se crean y explotan las obras.

En cuanto al instrumento adecuado, la Comisión Europea ha considerado que la forma más adecuada es una Directiva. Mediante la Directiva se establece un denominador común respetuoso con las diversas vías de adaptación que las diferencias legales y culturales van a marcar. No obstante, es en campos como el referente a la copia privada

donde se prevé que la Comisión Europea deba controlar en mayor medida los efectos del desarrollo tecnológico con vistas a iniciar acciones complementarias si fuera necesario.

Si incidimos en algunos aspectos concretos puede constatarse que desde el Libro Verde y la Comunicación de 20 de noviembre de 1996 se viene observando una fuerte oposición a una regulación por la que se determinara como país de origen el país donde el acto de transmisión estuviese teniendo lugar, así como su legislación nacional la única aplicable y ello por dos motivos: la propia naturaleza técnica de la transmisión digital por la que es difícil determinar un único lugar de origen de la misma y el riesgo que supone la aplicación del principio “país de origen” consistente en dejar sin protección a los titulares de derechos, fundamentalmente en los casos en que la transmisión se origina en terceros países.

No puede dejar de mencionarse por su importancia en el ámbito de la Comunidad el Libro Verde de la Comisión Europea de 19 de julio de 1995, sobre los derechos de autor y los derechos afines en la Sociedad de la Información, así como la comunicación elaborada por la Comisión relativa al seguimiento del Libro Verde. Este se compone de dos partes, la primera de las cuales se refiere al funcionamiento teórico de la sociedad de la información, destacándose la importancia del desarrollo de la misma para la Comunidad y para la consecución del mercado interior, al mismo tiempo que trata de determinar los retos que plantea. En la segunda parte, se intentan fijar los temas prioritarios para la protección de los derechos de autor y derechos afines en el marco de la sociedad de la información. Llama la atención sobre el hecho de que la mayoría de los servicios y productos que se creen podrán beneficiarse plenamente de las autopistas de la información, precisándose una marco normativo coherente a nivel nacional, comunitario e internacional. Tales servicios y productos, o bien recurrirán a obras preexistentes, o bien conducirán a la creación de obras nuevas, debiendo adaptarse las obras preexistentes al entorno digital antes de su transmisión. Por lo que respecta a las obras nuevas y la creación de nuevos servicios se prevé que conlleven cuantiosas inversiones, ya que de lo contrario el contenido de los mismos sería muy limitado.

El Libro Verde ya tenía en cuenta la importancia de los esfuerzos y de las inversiones en los nuevos servicios, y que éstas sólo van a realizarse si en el entorno digital

la protección a los derechos de autor y derechos afines resulta ser la adecuada. También tomaba en consideración el principal problema en lo que respecta a la protección, es decir, que una vez prestado el servicio en la red es muy difícil garantizar que la obra o la prestación no sea copiada, transformada o explotada sin el conocimiento de los titulares de los derechos y en su perjuicio. Así pues, el Libro Verde sentó las bases de la progresiva regulación en el marco comunitario y de su necesaria armonización, tanto de los aspectos horizontales – el derecho aplicable y el agotamiento de ciertos derechos – como respecto a derechos específicos.

En cuanto a los derechos específicos, puede hablarse del derecho de reproducción, el concepto de público, o el análisis de determinados derechos concretos que podrían otorgarse a los diferentes tipos de transmisiones digitales, distinguiéndose así entre el derecho de difusión digital y el derecho de radiodifusión digital. También es objeto de atención el problema de derecho moral y de las cuestiones relativas a la gestión de los derechos y los anejos a los sistemas de identificación y de protección técnica.

Es necesario encuadrar la evolución que ha experimentado la protección internacional de los derechos de autor y derechos afines. Esta evolución se ha plasmado en tres grandes acuerdos multilaterales: el Convenio de Berna para la protección de las obras literarias y artísticas (Acta de París, 1971), la Convención de Roma sobre la protección de los artistas intérpretes o ejecutantes, los productores de fonogramas y los organismos de radiodifusión (1961) y el Acuerdo sobre aspectos de los derechos de propiedad intelectual relacionados con el comercio (Acuerdo ADPIC, 1995). Más recientemente, la Conferencia Diplomática de la OMPI de diciembre de 1996 dio lugar a la adopción de los dos nuevos Tratados ya mencionados de la OMPI sobre interpretación o ejecución y fonogramas, que adquieren una gran trascendencia y constituirán la base de condiciones internacionales equiparables para los derechos de autor en la era de la digitalización.

El Tratado de la OMPI sobre derechos de autor, concertado en Ginebra el 20 de diciembre de 1996, constituye uno de los principales pilares de la regulación de esta materia. Incluye dentro de la protección del derecho de autor los programas de ordenador y las bases de datos. En cuanto a los derechos de autor aborda tres: el derecho de distribución, el derecho de alquiler y el derecho de comunicación al público, cada uno de los cuales

constituye un derecho exclusivo sujeto a ciertas limitaciones o excepciones. El Tratado obliga a las partes contratantes a prever recursos jurídicos contra la acción de eludir las medidas tecnológicas – como la codificación – utilizadas por los autores en relación con el ejercicio de sus derechos y contra la supresión o alteración de información consistentes en ciertos datos que identifican a la obra o a sus autores y que son necesarios para la gestión de sus derechos. Asimismo obliga a cada parte a adoptar, de conformidad con su sistema jurídico, las medidas necesarias para garantizar su aplicación, y en particular a asegurarse de que en su legislación nacional se establezcan procedimientos de ejercicio de los derechos protegidos que permitan la adopción de medidas eficaces contra cualquier acción infractora de los mismos, incluyendo todo recurso ágil para prevenir las infracciones.

Por otra parte y como es sabido los códigos para representar la información pueden analógicos y digitales. Mientras la forma analógica consiste en recoger y transmitir variaciones continuas en el fenómeno sobre el que versa el mensaje, en la representación digital de la información se parte de reducir el universo de casos posibles a un conjunto numerable y de asignar a cada uno de ellos uno o varios símbolos elegidos de entre un repertorio finito. Cualquier información es susceptible de ser digitalizada ya que cualquier código por el que se exprese o se conserve puede traducirse a otro tipo de números, de aquí el nombre de digitalización o conversión a dígitos. En definitiva, toda información, cualquiera que sea su naturaleza puede convertirse en una secuencia de bits. La digitalización junto con los programas de ordenador y las redes de transmisión de banda ancha han dado lugar al nacimiento de las autopistas de la información mediante las cuales cualquier usuario de red, a través de su ordenador personal puede recibir toda clase de obras casi instantáneamente, desde cualquier punto del planeta. Hasta ahora las copias privadas no tenían relevancia económica dado el trabajo que era necesario emplear para hacerlas o simplemente porque era imposible la reproducción con los medio técnicos disponibles por los particulares, como ocurriría con las grabaciones audiovisuales. Ello se reflejaba en la regulación legislativa, pues lo cierto es que tradicionalmente el derecho exclusivo de reproducción tenía por objeto la multiplicación de ejemplares para su distribución en el mercado o entre terceros ajenos al ámbito privado doméstico de quién hacía las copias.

Los derechos de autor son susceptible de verse modificados por los nuevos avances tecnológicos, lo que supone otro elemento de inestabilidad. Esta es la razón por la que actualmente deban ser redefinidos. Como se ha indicado anteriormente hoy en día cualquier información puede ser digitalizada y distribuida instantáneamente por el mundo. La aparición de Internet hace necesaria la revisión de muchos conceptos inherentes a los hasta ahora vigentes derechos de autor y la creación de otros nuevos. Entre ellos cabe destacar, el concepto de propiedad y autoría, el de equilibrio y el de transmisión. En efecto, Internet ha integrado dos diferentes aspectos de comunicación: comunicación Interpersonal y comunicación de masas, a lo que hay que añadir las posibilidades que ofrece de “intertextualizar” y compilar información de múltiples fuentes formando un nuevo producto poniendo en controversia los conceptos de autoría y propiedad.

El uso de Internet genera una serie de operaciones que se repiten constantemente y suscitan interrogantes referentes al derecho de autor. A este respecto resultan muy clarificadoras las conclusiones a que llega Ribas. Siguiendo a este autor pueden reconocerse las siguientes operaciones: búsqueda, visualización, introducción en memoria caché, descarga (download o almacenamiento), impresión, transformación, publicación en Internet, cesión y distribución, y cita automática.

En cuanto a la búsqueda que reproduce el título de una página HTML que cumpla los criterios de búsqueda introducidos por el usuario, así como la URL, el tamaño y la fecha y un fragmento de la obra, se trata de una reproducción mínima limitada a los elementos que resultan necesarios para identificar la información y poder apreciar el interés de su contenido, por lo que no precisa el previo consentimiento del autor ya que puede entenderse como una cita de la obra, exenta de autorización. Este autor señala que cuando la búsqueda la efectúa el propio usuario, utilizando programas específicos para ello, la trascendencia jurídica de esta actividad de rastro es mínima, pues se entiende limitada a las zonas de libre acceso de la red.

La visualización consiste en la representación en pantalla de una obra, lo que implica una transferencia de la misma desde el servicio hasta el ordenador donde reside el programa navegador o cliente. Ello constituye un acto implícitamente autorizado por el autor que ha introducido su obra en la red, especialmente si ha sido el creador de la página

Web en que se reproduce dicho contenido, existiendo la posibilidad de limitar la comunicación pública mediante claves, sistemas de cifrado y otros medios que garanticen el cobro de las correspondientes tarifas. En cuanto a la introducción en memoria caché por la cual los servidores, navegadores y routers almacenan en un área del disco duro los datos o archivos consultados por el usuario, al tratarse de un acto automático, ajeno a la voluntad del mismo, se excluye cualquier intencionalidad de éste a no ser que el usuario disponga de los medios para administrar su contenido en cuyo caso el uso de las obras almacenadas equivaldrá al de un download. Cuando se habla de download se hace referencia a la transferencia de determinada información disponible en la red al disco duro del ordenador personal del usuario, lo que se considera también como una actividad inherente al uso de Internet. Para Ribas, el autor que introduce una obra en un servicio conectado a Internet de forma abierta, está autorizando, ya sea implícitamente o de forma expresa la descarga de la misma y su almacenamiento en el disco duro, e incluso puede incluir una opción para la descarga del fichero. Debe distinguirse entre los contenidos visualizados en tiempo real y los contenidos que se hallan en un formato no visualizable por el programa navegador, puesto que la apariencia de titularidad puede ser distinta en ambos supuestos. Cuando se visita una página Web se presume que el texto que aparece en pantalla ha sido redactado por el que aparece como autor y en caso de que notoriamente no se así, la descarga de la obra puede constituir una reproducción no autorizada. Si la obra transferida está en un fichero que no puede ser visualizado en tiempo real, lo prudente es comprobar su titularidad en el momento de la instalación. La comprobación de las facultades del autor que cede una obra a través de Internet es conveniente para evitar infracciones al derecho de autor. Esta comprobación se ve facilitada por la implantación de dispositivos identificadores que confieren información sobre la titularidad de la obra, actos autorizados, modalidades de explotación, etc.

La impresión supone el paso a papel de una obra, lo que constituye también una reproducción en otro soporte y salvo prohibición expresa u otra indicación, ha de entenderse también autorizada en el momento en que un autor introduce la obra en Internet. En cuanto a la transformación no se precisará autorización si se destina a un uso privado, no así cuando se publiquen o distribuyan las transformaciones realizadas en cuyo caso se precisa autorización del titular. La cesión y distribución de la obra a terceros no es un acto

privado exigiendo autorización expresa del titular de los derechos, pues la introducción de una obra en Internet no faculta a su redistribución. Asimismo la introducción de una obra en una base de datos o en una página Web accesible a través de Internet constituye un acto de comunicación pública y precisa la autorización expresa del autor o titular de los derechos de explotación. Por último y en referencia a la cita automática, constituye un uso aceptado por los usuarios de grupos de noticias y listas de distribución y puede entenderse autorizada por el simple hecho de participar en estas formas de debate.

De lo dicho anteriormente se desprende la necesidad de llegar a un acuerdo en cuanto a la definición del acto de reproducción y a la cuestión de las limitaciones y excepciones. Las medidas habrán de ser a escala europea y no nacional.

Como ventajas se esperan grandes beneficios desde el punto de vista económico, social y cultural, nuevas perspectivas de desarrollo y explotación de la creación intelectual, aumento de la competitividad de la industria europea y de la cuota de mercado. No obstante, estos buenos augurios encuentran un serio oponente en el evidente riesgo de piratería a gran escala de la propiedad intelectual que ya se produce, en cierta medida, en el mercado fuera de línea, sobre todo, en lo referente a programas de ordenador, obras audiovisuales y grabaciones musicales. Dada la gran disponibilidad de obras protegidas en formatos digitales en línea se teme que, en caso de no tomarse las medidas adecuadas, las ventas lícitas en esta área se vean muy disminuidas. Es por ello que la protección de los derechos de autor cobra una especial significación.

Las medidas tecnológicas habrán de ampliar su protección para la garantía del respeto del derecho a la intimidad de las personas. Ha de tenerse también en cuenta que los propios sistemas de protección e identificación pueden procesar datos personales sobre los hábitos de consumo de los usuarios individuales con relación a los trabajos protegidos, lo que permitiría tener un seguimiento del comportamiento en línea que pudiera entrar en colisión con tal derecho.

Todos los Estados miembros reconocen un derecho exclusivo de reproducción, aunque las normativas nacionales difieran en cuanto al ámbito del derecho, ya que algunas consideran protegidos actos tales como la digitalización de una obra o la carga o descarga

desde la memoria de un ordenador y otras se centran en la reproducción material sin determinar claramente si quedan incluidos los actos de reproducción electrónica en general. Existe también una gran diversidad si quedan incluidos los actos de reproducción electrónica en general. Existe también una gran diversidad en cuanto a las excepciones y limitaciones aplicables al mismo, desde los fair uses del derecho anglosajón que establecen en su legislación una excepción general para los usos con fines de investigación, estudios privados, crítica, revisión e información sobre acontecimientos recientes, hasta otros ordenamientos jurídicos en que las limitaciones son mayores. En muchos Estados las excepciones relativas a la utilización de obras protegidas en bibliotecas y archivos no se aplican específicamente al material digitalizado.

Respecto a las copias para uso privado de material sonoro y audiovisual existe una práctica unanimidad entre los Estados miembros que establecen en su normativa una excepción al derecho exclusivo de reproducción a su favor. Es importante resaltar como las legislaciones de los Estados miembros no establecen distinción entre la tecnología analógica y la digital en lo que a materia de copias privadas se refiere y sólo Dinamarca no concede una exención por las copias privadas de obras protegidas incorporadas en soportes digitales, con independencia de que el sistema utilizado para realizar la copia sea digital o analógico.

Según la opinión de los abogados Ribas y Rodríguez, el upload de la obra, consistente en un ejercicio del derecho de comunicación pública que integra el consentimiento del autor para visualizar y descargar la obra, es un uso permitido, y el download ha de considerarse inherente al uso de Internet y por lo tanto, como una reproducción autorizada, a no se que figure su expresa prohibición. A partir de este momento las reproducciones posteriores serán copias privadas siempre que se cumplan los siguientes requisitos: que sean para uso privado del copista, que este sea un usuario legítimo de la obra, que la copia no sea objeto de utilización colectiva ni lucrativa y que la obra reproducida no sea un programa de ordenador. Respecto a la definición del calificativo privado, ha de entenderse referido a cualquier copia que no sea cedida a terceros, es decir, que no trascienda del ámbito estrictamente privado del copista, considerado como persona

física. El uso legítimo de la obra quiere indicar el acceso a la misma mediante una cesión autorizada del titular.

9. ORGANIZACIÓN DE LOS RECURSOS ELECTRÓNICOS

9.1 TRATAMIENTO TRADICIONAL DE LOS RECURSOS ELECTRÓNICOS

Las bibliotecas que han optado por ofrecer a los documentos Internet el tratamiento calificado como tradicional, someten todos los materiales a los mismos procesos; es decir, se seleccionan los recursos más adecuados a la colección i a los usuarios, los catalogan de la misma manera y con las mismas herramientas que utilizan para catalogar los materiales locales que se integran físicamente en la colección, y entonces su descripción pasa a formar parte del catálogo local. La única diferencia es que los recursos Internet, a diferencia de los otros recursos, en general no forman físicamente parte de la colección local, ya que la política más generalizada entre las bibliotecas es la de no mantener copia de los recursos Internet, aunque algunas bibliotecas hacen excepción a esta norma para algunos tipos de materiales electrónicos, como las revistas, con el fin de asegurar el acceso, y por eso optan por almacenar localmente estos materiales.

El tratamiento tradicional de los recursos Internet se encuentra en una primera fase de desarrollo y, por tanto, todavía está poco consolidado. Fue a partir de Internet Resource Project de OCLC (1993) cuando se investigó la naturaleza de la información textual electrónica de acceso remoto y los problemas teóricos y prácticos que resultaban de la aplicación de las *Anglo-American cataloguing rules*, 2nd edition, 19988 revisión (AACR2R) y del formato USMAC en la creación de registros bibliográficos MARC para estos recursos. El proyecto consistía en experimentar de manera práctica la adecuación de estas herramientas a la catalogación de documentos Internet: 30 profesionales con experiencia en la catalogación de ficheros informáticos catalogaron, cada uno, 30 ficheros Internet de una colección preseleccionada de 300 ficheros. Como resultado del proyecto se hicieron una serie de propuestas para adecuar el formato USMARC a la descripción de estos materiales. A partir de este momento, el formato ha ido evolucionando a medida que se ha adquirido más experiencia en el tratamiento de ficheros remotos.

El proyecto OCLC también fu el responsable del segundo impulso dado a la catalogación de recursos Internet. En el año 1995 este servicio bibliográfico inició otro proyecto – Internet Cataloging Project – con el objetivo práctico de crear una base de datos

– Intercat – con registros de recursos Internet que posteriormente se integrarían al catálogo colectivo de la institución. Cuando transcurrió un año de iniciarse el proyecto, a finales de abril de 1996, ya participaban 225 bibliotecas, muchas más de las que se habían previsto en un principio, y el número de registros de la base de datos Intercat era de 4046. Entre los recursos Internet que ya forman parte del catálogo de OCLC existen muchos textos electrónicos que procedían de proyectos de digitalización concretos, como el Project Gutenberg, o Columbia University Bartleby Library, directorios y otras herramientas de búsqueda en Internet, colecciones de material de archivo, páginas Web de sistemas de información universitaria, publicaciones periódicas electrónicas, etc. La realización del proyecto fue un estímulo para que las bibliotecas iniciasen la catalogación de estos materiales; también supuso el enriquecimiento de la base de datos de OCLC y de los catálogos locales de muchas bibliotecas americanas con registros que representan recursos Internet. Una vez terminado el proyecto como tal, se supone que muchas bibliotecas continuaran con la política de catalogar recursos Internet con los métodos tradicionales.

¿Cuáles son las ventajas de integrar la descripción de recursos Internet en catálogos locales y nacionales? En primer lugar, con esta práctica se asegura que estos materiales pasen por el mismo proceso de selección por el que pasaron todos los documentos de la biblioteca. La política de desarrollo de la colección y la selección de la cual es objeto el fondo de una biblioteca le ofrece unidad y asegura la calidad y la adecuación a las necesidades de los usuarios. Si esto es necesario en un entorno en el que editoriales y editores hacen una primera selección de todo lo que llega a publicarse, todavía es más necesario en el entorno Internet, donde no hay ningún tipo de filtro a la publicación libre de cualquier documento. Con esta selección previa, la biblioteca presta un servicio básico y de los más importantes para los usuarios. En la actualidad, algunas bibliotecas ya han empezado a elaborar políticas específicas para la selección de recursos electrónicos en general y de recursos Internet en concreto; sin embargo, la evaluación y la selección de estos últimos materiales todavía no están tan sistematizadas como la de otros recursos. En general, se puede deducir de algunas bibliotecas que han participado en el proyecto de OCLC es que en esta fase inicial las bibliotecas son todavía bastante prudentes, y que a la hora de seleccionar recursos para integrarlos en el catálogo local prefieren:

- Obras de referencia, como enciclopedias, diccionarios, directorios, etc
- Motores de búsqueda en Internet, como Yahoo, WebCrawler, Lycos, Infoseek, etc.
- Revistas electrónicas
- Textos electrónicos de autores literarios
- Materiales de interés para la investigación académica
- Páginas Web relacionadas con la misma institución o centro, por ejemplo, las páginas de los sistemas de información de la propia Universidad
- Herramientas útiles para el trabajo del personal interno

La segunda ventaja que resulta de integrar la descripción de recursos Internet en el catálogo local, es que el usuario encuentra en un solo sitio todos los materiales que la biblioteca ha considerado adecuados para la comunidad a la cual sirve, los encuentra procesados de la misma manera y además, los puede recuperar a través de un catálogo automatizado que, en la mayoría de los casos, tiene muchas capacidades de consulta y de recuperación. Las bibliotecas que han dispuesto de más de un catálogo de sus fondos, conocen el inconveniente que esto representa para el lector. Así pues, si a la dificultad de tener que consultar en dos lugares diferentes, se añade la gran cantidad de materiales que conforman el mundo de Internet, la falta de algún tipo de control en la red, la gran cantidad de motores disponibles en Internet para la búsqueda y la recuperación de recursos, etc. , es obvio que el servicio que se presta al usuario haciendo una selección previa de estos recursos e integrando su descripción en el mismo catálogo donde se pueden recuperar todos los otros documentos que conforman la colección. A medida que las bibliotecas disponen de catálogos con software Web, los beneficios todavía serán más evidentes, ya que el lector podrá recuperar los recursos Internet desde el mismo registro que consulta en el catálogo y, a partir de este, podrá navegar por todo el sistema. Si, además, estos registros elaborados localmente se utilizan para alimentar bases de datos de ámbito nacional, serán muchas las bibliotecas que los podrán copiar, a medida que los necesitan, a un coste más bajo que si los tiene que catalogar de nuevo.

Como se ha podido comprobar, para llevar a término este tratamiento bibliotecario de los recursos Internet, se ha tenido que modificar la normativa catalográfica para que acogiera la descripción de estos recursos. A continuación se examina de qué manera las AACR2R, el formato USMARC y el borrador de la segunda edición de la ISBD (CF) han dado lugar a la descripción y al acceso de estos materiales.

9.2 MODIFICACIONES INTRODUCIDAS EN LAS HERRAMIENTAS DE CATALOGACIÓN PARA ACOGER LOS RECURSOS INTERNET

9.2.1 AACR2

La edición de 1978 de las AACR2 dedicaba el capítulo 9 a la descripción de ficheros legibles por ordenador; con anterioridad, des de 1967, una serie de textos ofrecían una directrices iniciales para organizarlos. Dentro del programa ISBD de la FLA no hubo una norma destinada a la descripción de los ficheros informáticos hasta el año 1990 en el que se publicó la ISBD (CF): internacional Standard bibliographic description for computer files. Mientras se había publicado la revisión de 1988 de las AACR2 – las AACR2R – que, en el capítulo 9 reflejaba la experiencia catalográfica que se iba adquiriendo a medida que los ficheros informáticos se hacían más habituales en bibliotecas y se conocían mejor las características y los problemas que presentaba su descripción, así como su evolución y diversificación – se ha de tener en cuenta que entre una edición y otra ocurrieron hechos tan importantes como la introducción de los microordenadores y los CD-ROM. La ISBD (CF) pudo aprovecharse de esta experiencia y se introdujeron cambios destacables en las AACR2R como los siguientes:

El título del capítulo pasa a ser “computer files” (ficheros informáticos) en vez de “machina readable data files” (ficheros de datos legibles por máquina). La ISBD (CF) adopta la nueva terminología.

La fuente principal de información es la pantalla de título y no la etiqueta interna como en la edición de 1978. También se modifica el orden de preferencia de las fuentes y las fuentes prescritas de información de todas las zonas menos las de las zonas 6, 7 y 8.

En la zona 1 se modifica la designación general de material – de ficheros de datos legibles por máquina pasa a fichero informático -, y en nota aparecerá siempre la fuente de donde se ha tomado el título propio; esto permite conocer si en la base de datos ya existe un registro que describe el fichero informático. La nota sobre la fuente también es obligatoria para la zona 2 cuando la mención de edición se obtiene de una fuente diferente a la del propio título.

Se añade la zona 3 donde se rellenaran las características del fichero. En la edición de 1978 la zona quedaba vacía y esta información se anotaba en la zona 5, referente a la descripción física.

En la zona 4, referente a la publicación, la distribución se añaden cambios menores. Actualmente se ofrecen instrucciones para tratar documentos no publicados, en este caso ficheros informáticos. También se prevé anotar opcionalmente los datos de la fabricación del fichero, que no se contemplaban en la edición de 1978.

La zona 5, la de la descripción física, está totalmente renovada: la extensión y la designación específica de material corresponden a características del soporte físico, como en el resto de materiales, y no a las características internas, como cuando se anotaban el número y el tipo de ficheros y su extensión lógica; se añaden las dimensiones del soporte; cambia el concepto de material de acompañamiento y se menciona explícitamente que la zona no se anota para los ficheros remotos.

En la zona 7 se añaden las dos primeras notas sobre los requerimientos del sistema y la modalidad de acceso; también se modifican las notas en las zonas 3 y 5 para que se correspondan con los cambios introducidos en las zonas respectivas, y se suprimen las notas sobre la versión del programa y la forma de uso, que corresponde parcialmente a los requerimientos del sistema.

Tal como demuestra su aplicación, las AACR2R se pueden emplear para describir ficheros informáticos de acceso remoto y otros materiales nuevos como los multimedia interactivo; ahora bien, la adecuación de las reglas a estos recursos no es del todo satisfactoria. Cuando se publicó la edición de 1998 de las AACR2, los materiales

multimedia interactivos eran inexistentes y, aunque había bases de datos de acceso remoto consultables en línea, el uso de los cuales se ofrecía como servicio a los lectores, nunca se había pretendido que su descripción formara parte del catálogo. Más tarde, cuando se publicaron las AACR2R en el año 1993, en el capítulo 9 sólo se introdujeron cambios menores que no contemplaban la reciente diversificación de los ficheros informáticos. Como son unas reglas de ámbito internacional, en las cuales colaboran comités de distintos países, las modificaciones no se podían introducir de manera automática sin que hubiesen propuestas i acuerdos previos de todos los implicados; además, también se intentaba que no hubiesen grandes divergencias entre as reglas ISBD. Por eso, las revisiones de las AACR” eran relativamente lentas, sobretodo cuando se tenían que considerar materiales nuevos que están en constante evolución, y de los cuales se tiene poca experiencia.

Con lo que respecta a los ficheros informáticos de acceso remoto, en los últimos años la comunidad bibliotecaria americana ha discutido ampliamente la catalogación de estos materiales. Los congresos y seminarios sobre la organización de documentaos digitales y recursos Internet, los grupos de trabajo que se han formado para examinar la cuestión, y sobretodo la larga lista de “discusión papers” y “proposals” del comité MARBI (Machina Readable Bibliographic Information Comité) sobre aspectos concretos del tema son testimonio de la tarea realizada. Sin embargo, cuando se ha tenido que adoptar los cambios propuestos, éstos sólo se han podido introducir en el formato USMARC, que es nacional y por tanto no está sujeto a acuerdos con comités de otros países.

Como consecuencia de las deficiencias de las AACR” por lo que respecta a los materiales más nuevos y de la imposibilidad de modificarlas unilateralmente, los materiales multimedia interactivos, también de desarrollo reciente y en algunos casos disponibles por Internet, han sido objeto de un tratamiento provisional especial en el seno de la comunidad bibliotecaria americana. En el año 1994 se publicaron las Guidelines for bibliographic description of interactive multimedia, redactadas por un grupo de trabajo formado por miembros del Comité on Cataloging: Description and Access, de la Association for Library Collections and Technical Services. Las directrices tienen como objetivo la descripción de los materiales multimedia interactivos que integran diversas tecnologías – de ficheros informáticos con datos textuales y gráficos, re registros sonoros y de vídeo, del cine de

animación, etc. Como ya ha ocurrido con otros materiales, estas directrices, que siguen el esquema de descripción establecido en el capítulo 1 de las AACR2R y que han sido elaboradas por expertos en la materia, servirán para que los catalogadores adquieran experiencia en la catalogación de estos materiales; también servirán como un documento de trabajo del Joint Steering Comité for Revision of AACR. Con el tiempo pasarán a formar parte de las reglas de catalogación y se integrarán probablemente en su capítulo 9, de ficheros informáticos. En la última edición del USMARC, estos materiales ya se codificaban como ficheros informáticos, y el borrador de la segunda edición de la ISBD (CF), que recoge la experiencia norte americana, incluye prescripciones para la descripción de los materiales multimedia interactivos y los menciona explícitamente en la misma definición de ficheros informáticos.

9.2.2 FORMATO USMARC

El formato USMARC ha acogido desde la publicación de las AACR2 de 1978 la codificación de la descripción de los ficheros informáticos. Además, como no está comprometida a acuerdos internacionales, y ofrece información codificada complementaria a la información textual reglamentada por las AACR2, acepta con más facilidad que las reglas, adiciones, supresiones y borradores – siempre que no contradigan lo que establece la normativa catalográfica. Por eso, el formato USMARC ha podido acoger con facilidad algunos de los datos necesarios para representar los diferentes tipos de recursos Internet.

A la posición 6 de la cabecera de los registros USMARC, los ficheros informáticos se codifican con el código “m”. Esta posición rige la codificación del campo 007 – de longitud fija, con datos codificados sobre la descripción física y similar a los campos 03X del CATMARC – y la del 008 – también con datos codificados sobre el registro. Tanto el campo 007 como el 008 contienen información muy útil para la recuperación del registro y del documento.

El campo 007 tiene 23 posiciones fijas. En el caso de los ficheros informáticos, sólo se definen las seis primeras posiciones; quedan codificadas las siguientes características físicas:

Posición	Característica
007/00	Categoría de material
007/01	Designación específica de material
007/02	Fichero original/reproducción
007/03	Color
007/04	Dimensiones
007/05	Sonido

De esta lista y de los códigos que se anotan en cada posición hay que destacar los que corresponden a la designación específica de material, ya que se trata de una enumeración exhaustiva de los soportes físicos que pueden alojar los ficheros informáticos. Se ofrecen las siguientes designaciones, más amplias y precisas de las que se presentan en la zona 5 del capítulo 9 de las AACR2R: cinta en cartucho, chip en cartucho, disco óptico en cartucho, cinta en casete, bobina, disco magnético, disco óptico-magnético, disco óptico y fichero remoto.

El campo 008 tiene 40 posiciones fijas, un subconjunto de las cuales se define de la misma manera para todos los tipos de materiales; los valores de las posiciones 18-34 son específicos de cada tipo. En el caso de los ficheros informáticos, contienen la información codificada siguiente:

Posición	Característica
008/18-21	Sin definir
008/22	Destinatarios del documento
008/23-25	Sin definir
008/26	Tipo de fichero informático

008/27	Sin definir
008/28	Publicación de gobierno
008/29-34	Sin definir

En la posición 26 se codifican los tipos de ficheros informáticos. Esta es la tipología más amplia, sistematizada y actualizada que se ha encontrado en la bibliografía sobre el tema; quedan incluidos los siguientes ficheros informáticos con las definiciones correspondientes: datos numéricos, programa de ordenador, fichero representacional, documento (texto), datos bibliográficos, fuentes (tipografía), juego, fichero de sonido, multimedia interactivo, sistema o servicio en línea, y las combinaciones de dos o más categorías de las anteriores.

Por lo que respecta a los campos de longitud variable, el formato USMARC contempla los siguientes campos de uso exclusivo para los ficheros informáticos:

036	Número asignado por el productor del fichero informático
256	Características del fichero informático (contiene la zona 3 de la descripción bibliográfica)
307	Horas, etc. (se aplica a sistemas y servicios en línea e informa de la disponibilidad)
516	Nota sobre tipo de fichero informático o de datos de ordenador
538	Nota sobre datos del sistema
753	Acceso a datos del sistema
856	Localización y acceso electrónico

El campo 516 contiene información textual similar a la que se ofrece en la posición 26 del campo 008 en forma codificada. No existe una lista normalizada de tipo de datos o ficheros informáticos, pero como esta información no es un punto de acceso, la

normalización no es tan necesaria. Se anota un descriptor general (por ejemplo, texto, datos numéricos, programa de ordenador) que puede ir seguido de información más específica entre paréntesis, como la forma o el género del material textual (por ejemplo, biografía, diccionarios, índice). Algunas bibliotecas que catalogan recursos Internet han introducido el término “Web Page” (página Web) para designar este tipo de recurso. El campo no tiene un paralelo con las notas previstas de las AACR2 y, de momento, tiene la función de precisar, mientras no se modifiquen las reglas, la información tan general que se ofrece en la zona 3 de la descripción: datos y/o programas de ordenador.

El campo 538 contiene las dos primeras notas previstas en el capítulo 9 de las AACR2: requerimientos del sistema y modalidad de acceso. En un primer campo 538 se ofrecen las características físicas del fichero – como la densidad del registro – el lenguaje de programación y los requerimientos del sistema en el caso del software – modelo de ordenador, sistema operativo, memoria - ; también se pueden anotar los periféricos que se van a necesitar. Otro campo 538 incluye la modalidad de acceso, por ejemplo:

538 Mode of access: Internet

El campo 753 tiene el objetivo de proporcionar puntos de acceso sobre datos relativos al sistema. El campo consta de tres subcampos principales donde se registran la marca y el modelo de ordenador, el lenguaje de programación y el sistema operativo. Lo cierto es que este campo lo pueden utilizar las bibliotecas que catalogan ficheros informáticos; la misma información se puede ofrecer como nota en el campo 538 y, por eso, no se considera tan necesario que también constituya un punto de acceso al registro.

El campo 856 está situado entre los campos con información local (8XX) que conforma el registro de fondo, o de ejemplares, y localizaciones, conectado al registro bibliográfico. Este campo, de creación reciente, contiene la información necesaria para localizar un recurso electrónico; también ofrece información suficiente para poder transferir ficheros, para subscribirse a una revista electrónica o para conectarse a un recurso electrónico remoto. Como el recurso puede estar situado en diversas localizaciones, se acordó que su dirección iría en un campo del registro de fondo y localizaciones – que es donde va también la signatura topográfica de los otros materiales - ; por otra parte, la

definición del campo, su estructura y las directrices para aplicar los designadores de contenido se publican dentro del formato bibliográfico.

Campo 856, Localización y acceso electrónico

Indicadores

Primero: Método de acceso

0 Correo electrónico

1 FTP

2 Conexión remota (Telnet)

3 Conexión telefónica

7 Método especificado en el subcampo \$2

Segundo: Sin definir

Códigos de subcampo

\$a Nombre del host o dominio (R)

\$b Número de acceso (NR)

\$c Información de compresión del fichero (R)

\$d Directorio y subdirectorios (path) (R)

\$f Nombre electrónico (R)

\$g Nombre electrónico final de la serie (R)

\$h Procesador de la demanda (username) (NR)

\$i Instrucción (R)

\$j	Bits por segundo (NR)
\$K	Password (NR)
\$l	Conexión/desconexión (NR)
\$m	Persona de contacto para ayuda (R)
\$n	Localización del host del subcampo-a (NR)
\$o	Sistema operativo (NR)
\$p	Puerto (NR)
\$q	Forma de transferencia del fichero (NR)
\$r	Settings (NR) (número de bits por carácter; número de bits que indican el final de un byte, y paridad)
\$s	Dimensiones del fichero (R)
\$t	Emulación del terminal (R)
\$u	URL (R)
\$v	Horas en las que está disponible el método de acceso (R)
\$w	Número de control del registro (R)
\$x	Nota interna (R)
\$z	Nota pública (R)
\$2	Método de acceso (NR)
\$3	Especificación de la parte del recurso en el que se aplica en campo (NR)

(R) = Subcampo repetible

(NR) = Subcampo no repetible

Figura 1: Esquema del campo 856 del USMARC, localización y acceso electrónico, con un resumen de sus indicadores y subcampos ⁽³⁾

El campo 856 es muy potente; cuando se utiliza con catálogos con recursos Web tiene la capacidad de establecer una hiperconexión con el mismo recurso. Es decir, en recuperar un registro bibliográfico de un recurso Internet en un catálogo Web, mediante un clic en la dirección del recurso Internet, que aparece en el subcampo \$u del campo 856, se accede directamente al recurso; además este subcampo no tiene que ser visible para el usuario.

El desarrollo del campo 856 ha durado diversos años desde que se inició en 1993 con el Internet Resource Project de OCLC. De este proyecto surgió el Proposal 93-4, que planteaba la ampliación de los códigos de la posición 26 del campo 008, tipo de fichero informático, y la adición del campo 856 con información sobre localización y acceso electrónico; posteriormente, a medida que los recursos Internet se han ido multiplicando y diversificando, el campo se ha ido perfilando a través de otros “proposals” y “discussion papers” hasta tener la forma actual.

Ni las AACR2 ni el formato USMARC han ofrecido todavía solución al problema de los recursos que tienen un contenido similar pero que se encuentran en soportes diferentes. Sería el caso, por ejemplo, de una revista que está en formato electrónico y en formato impreso, o de una base de datos que es accesible en línea, que tiene una versión en CD-ROM y otra versión en forma impresa, y siempre con el mismo contenido. La lectura literal de las reglas de catalogación obligaría, en muchas circunstancias, a la creación de un registro para cada caso. Casos de este tipo son cada vez más habituales, y el resultado de trabajar siguiendo las normas significa la creación de muchos registros casi iguales que pueden desconcertar al usuario. SE ha discutido mucho sobre el tema, y parece ser que la tendencia es la de crear un solo registro para estos casos y ofrecer la información de los otros soportes en notas; por eso, si existen diversas versiones en formato electrónico

accesibles en línea, se puede repetir el campo 856, o en lagunas ocasiones, el subcampo \$u con la URL, tantas veces como sea necesario. Ahora bien, la solución definitiva del problema sólo puede ser resultado de una revisión profunda de la normativa catalográfica que revise los mismos principios en los que se basa.

9.2.3 ISBD (CF)

La primera edición de la ISBD (CF) (1990) recogía la experiencia de la aplicación del capítulo 9, ficheros informáticos, de las AACR2, de las cuales ya se había publicado la revisión de 1988. Las diferencias que resultan de la aplicación de los dos textos son mínimas. El texto propuesto refleja las modificaciones que en los últimos años se han hecho del formato USMARC para adecuarlo a todos los tipos de ficheros informáticos y las discusiones que han tenido lugar entre la comunidad bibliotecaria americana respecto al tema. En este sentido, hay que decir que los miembros del grupo de trabajo y del grupo de asesores predominan los representantes angloamericanos, de los Estados Unidos, el Reino Unido y Canadá.

La ISBD (CF) contempla los materiales que son manipulados por ordenador y los que se usan con los periféricos, y de manera explícita se mencionan las obras multimedia interactivas como uno de los objetos de la norma. También se distingue entre los ficheros informáticos de acceso local y los de acceso remoto, ya que el tratamiento que reciben en su descripción es diferente. Sin embargo, la definición de cada uno de estos dos tipos de ficheros es ambigua: el acceso local implica que exista un soporte físico que pueda ser descrito y que el usuario haya de insertar en un ordenador o en un periférico; un fichero es de acceso remoto cuando el usuario no puede manejar ningún soporte físico. La descripción se tiene que poder aplicar a todos los ejemplares de la misma edición, pero la distinción anterior entre los ficheros de acceso remoto y los de acceso local obligaría a tratar una misma base de datos, por ejemplo, de manera diferente en función si estuviera instalada en una red o si no lo estuviera, lo cual contradice los principios en los que se basa la misma norma ISBD. La ISBD (CF) puntualiza una cuestión respecto a la cual existen distintas posturas diferenciadas y que la comunidad norteamericana ha resuelto de la misma manera: todos los ficheros informáticos remotos se tienen que considerar documentos publicados.

La ISBD (CF) es mucho más precisa que las AACR2 a la hora de ofrecer indicaciones sobre un punto que provoca muchos problemas de interpretación: que es lo que supone una nueva edición de una obra. El texto de la IFLA intenta ofrecer soluciones enumerando qué cambios constituyen una nueva edición, modificaciones en el contenido del fichero, cambios en el lenguaje de programación, actualizaciones o mejoras en la eficacia de los ficheros, etc. y cuáles no, el soporte físico o sus dimensiones, los formatos relacionados con la impresión o la visualización, etc.

La influencia de los textos americanos en la ISBD (CF) es notable en los cambios propuestos para la zona 3, tipos de ficheros y extensión. No se modifican los elementos que conforman la zona (designación del fichero y extensión), pero sí los términos que designan el fichero. Si bien la zona 3 en el texto de las AACR2 no ha sufrido ninguna modificación, en concordancia con la nota correspondiente al campo 516 del USMARC (nota del tipo de fichero informático o datos del ordenador), se mantienen las designaciones generales del tipo de fichero, datos del ordenador, programas de ordenador, y datos y programas de ordenador, pero se anotan acompañadas de calificaciones. Por ejemplo: “computer numeric data” (datos numéricos de ordenador), “computer census data” (datos de censo de ordenador), “computer images” (imágenes de ordenador), “computer application programs” (programas de aplicación de ordenador), “computer operating system program” (programa de sistema operativo de ordenador), etc. Como en la versión anterior, estas designaciones pueden complementarse con la adición de la extensión del fichero entre paréntesis.

La zona 5 no presenta ninguna diferencia importante respecto a los textos angloamericanos; sin embargo, se propone una lista de designaciones específicas de material diferente y más breve que las sugeridas en las AACR2 y en el formato USMARC. Ahora bien, tal y como estos textos prescriben, la zona sólo se utiliza para ficheros informáticos locales. En la nota introductoria a la zona se hace una precisión que puede resultar ambigua e incluso confusa: se deja a criterio del catalogador decidir si se hace una descripción bibliográfica separada cuando el documento se presenta en soportes físicos diferentes, o bien, alternativamente, repetir la zona 5 para cada soporte físico. Obviamente, esta opción de hacer una descripción bibliográfica separada puede ser contradictoria con la afirmación que se hace en el capítulo 2 según la cual si sólo hay un cambio en el soporte

físico, no se puede considerar una edición separada y , por tanto, no requiere una descripción bibliográfica a parte.

Las notas previstas en la ISBD (CF) son paralelas a las establecidas en la edición de 1988 de las AACR2; la única diferencia apreciada está en el orden en que se enumeran. Las AACR2, más sistemáticas, ofrecen las notas en el orden previsto en el registro; la ISBD (CF) no está tan trabajada en este sentido. Sin embargo, el resultado es el mismo, y las dos primeras notas son, en las dos normas, las que hacen referencia a los requerimientos del sistema y a la modalidad de acceso. Precisamente, esta nota sobre el método de acceso es el único lugar establecido en la ISBD (CF) para proporcionar la localización de los ficheros informáticos de acceso remoto. La información que se incluye en el campo 856 del USMARC, que tiene la capacidad de localizar el recurso y recuperarlo cuando el registro se utiliza en catálogos Web, no se ha previsto en la ISBD (CF) seguramente porque se considera que son datos más cercanos al registro de fondo que al registro bibliográfico.

9.2.4 ANÁLISIS COMPARATIVO DE LAS NORMAS ISBD (CF) E ISBD (ER)

Las normas internacionales para la descripción bibliográfica de los recursos electrónicos, ISBD(ER), se ocupan de normalizar los registros bibliográficos que describen los documentos electrónicos. Nacieron como continuación de las ISBD(CF), normas para la descripción de archivos de ordenador, expresión que se quedó antigua antes de su nacimiento.

ISBD(CF): Descripción Bibliográfica Internacional Normalizada para Archivos de Ordenador

ISBD(ER): Descripció Bibliogràfica Normalitzada Internacional per a Recursos Electrònics.

INTRODUCCIÓN	
 ISBD (CF)	 ISBD (ER)
 INTRODUCCIÓN La ISBD surge de la resolución tomada en la Reunión Internacional de Expertos en Catalogación, organizada en Copenhague en 1969 por el Comité de Catalogación de la IFLA, en la que se decidió que se debía normalizar la forma y el contenido de la descripción bibliográfica. La Descripción Bibliográfica Internacional Normalizada para Publicaciones Monográficas fue la primera de las ISBDs aparecidas como resultado del mandato de 1969. El primer texto de las ISBD(M) se publicó en 1971 bajo la forma de un conjunto de recomendaciones. Ya en 1973 cierto número de bibliografías nacionales habían adoptado este texto y, gracias a las traducciones del texto original inglés a otras lenguas, se habían tenido en cuenta por cierto número de Comités de Catalogación a la hora de volver a redactar el borrador de las reglas nacionales de descripción bibliográfica. Ya en ese momento se había percibido que la palabra impresa no es más que uno de los procedimientos de la transmisión documental	 INTRODUCCIÓN La ISBD surge de la resolución tomada en la Reunión Internacional de Expertos en Catalogación, organizada en Copenhague en 1969 por el Comité de Catalogación de la IFLA, en la que se decidió que se debía normalizar la forma y el contenido de la descripción bibliográfica. La Descripción Bibliográfica Internacional Normalizada para Publicaciones Monográficas fue la primera de las ISBDs aparecidas como resultado del mandato de 1969. El primer texto de las ISBD(M) se publicó en 1971 bajo la forma de un conjunto de recomendaciones. Ya en 1973 cierto número de bibliografías nacionales habían adoptado este texto y, gracias a las traducciones del texto original inglés a otras lenguas, se habían tenido en cuenta por cierto número de Comités de Catalogación a la hora de volver a redactar el borrador de las reglas nacionales de descripción bibliográfica. Ya en ese momento se había percibido que la palabra impresa no es más que uno de los procedimientos de la transmisión documental

a través de la cual se satisfacen las necesidades de comunicación de individuos e instituciones, y que, por lo tanto, el programa de la IFLA para la Descripción Bibliográfica Internacional Normalizada debería dirigir la preparación de una estructura descriptiva normalizada para materiales no librarios. Como consecuencia de ello, se preparó y publicó en 1977 la primera edición de las ISBD(NBM), Descripción Bibliográfica Internacional Normalizada para Materiales no Librarios. Estas ISBDs recogían disposiciones referidas a archivos de datos legibles por ordenador. Sin embargo, cuando el Comité de Revisión de las ISBD, formado por la IFLA en 1981, empezó a trabajar en las ISBD(NBM) en paralelo con las ISBD(CM), ISBD(M) y ISBD(S), se decidió que debía dedicarse una especial atención a la necesidad, cada vez más urgente, de preparar unas ISBDs separadas para los archivos de ordenador. Con el desarrollo de programas y ficheros de datos para ordenadores más pequeños, la naturaleza del soporte se hizo más compleja; además, este cambio en la producción fue muy similar al de otros materiales que se incorporaban ampliamente a las colecciones de las bibliotecas y que precisaban, asimismo, un control bibliográfico. Como resultado se estableció el Grupo de Trabajo de las	a través de la cual se satisfacen las necesidades de comunicación de individuos e instituciones, y que, por lo tanto, el programa de la IFLA para la Descripción Bibliográfica Internacional Normalizada debería dirigir la preparación de una estructura descriptiva normalizada para materiales no librarios. Como consecuencia de ello, se preparó y publicó en 1977 la primera edición de las ISBD(NBM), Descripción Bibliográfica Internacional Normalizada para Materiales no Librarios. Estas ISBDs recogían disposiciones referidas a archivos de datos legibles por ordenador. Sin embargo, cuando el Comité de Revisión de las ISBD, formado por la IFLA en 1981, empezó a trabajar en las ISBD(NBM) en paralelo con las ISBD(CM), ISBD(M) y ISBD(S), se decidió que debía dedicarse una especial atención a la necesidad, cada vez más urgente, de preparar unas ISBDs separadas para los archivos de ordenador. Con el desarrollo de programas y ficheros de datos para ordenadores más pequeños, la naturaleza del soporte se hizo más compleja; además, este cambio en la producción fue muy similar al de otros materiales que se incorporaban ampliamente a las colecciones de las bibliotecas y que precisaban, asimismo, un control bibliográfico. Como resultado se estableció el Grupo de Trabajo de las
---	---

ISBD(CF) que se reunió por primera vez en marzo de 1986. Gracias a los comentarios de una amplia variedad de usuarios de archivos de ordenador se fueron mejorando paulatinamente los borradores preliminares basados en la revisión de las ISBD(NBM); este y otros esfuerzos llevados a cabo por el Grupo de Trabajo culminaron a finales de 1988 en un borrador final de las ISBD(CF). Como característica clave de este trabajo destaca la armonización del texto de las ISBD(CF) con los textos revisados de las cuatro ISBD que se mencionan más arriba y que se habían publicado en 1987-88.	ISBD(CF) que se reunió por primera vez en marzo de 1986. Gracias a los comentarios de una amplia variedad de usuarios de archivos de ordenador se fueron mejorando paulatinamente los borradores preliminares basados en la revisión de las ISBD(NBM); este y otros esfuerzos llevados a cabo por el Grupo de Trabajo culminaron a finales de 1988 en un borrador final de las ISBD(CF). Como característica clave de este trabajo destaca la armonización del texto de las ISBD(CF) con los textos revisados de las cuatro ISBD que se mencionan más arriba y que se habían publicado en 1987-88..
Las ISBD(CF) incluyen un índice y cinco apéndices. El primer apéndice presenta un precepto para la técnica especial de la descripción en varios niveles. El segundo apéndice proporciona un breve esbozo del modo mediante el cual deben transcribirse los datos cuando un texto se lee parcialmente de derecha a izquierda y parcialmente de izquierda a derecha. El tercer apéndice contiene unas listas de denominaciones específicas y generales de materiales. El cuarto apéndice proporciona abreviaturas recomendadas para su uso en registros en inglés. El apéndice final lleva los ejemplos formulados con el fin de ofrecer una muestra de; resultado de la aplicación de la normativa	Los recursos electrónicos son producto de tecnología volátil que continúa generando cambios a un ritmo muy rápido. Entre los avances recientes se encuentran los siguientes: aparición de los multimedia interactivos; desarrollo de la tecnología óptica; disponibilidad de recursos electrónicos remotos en Internet y en WWW (World Wide Web), y reproducciones de recursos electrónicos. Como resultado de estas consideraciones y en reconocimiento de la importancia que las ISBD (CF) han de ofrecer a las implicaciones bibliográficas de estas tecnologías, el Standing Committee de la Cataloging Section de la IFLA decidió iniciar un examen y la revisión formal de las

ISBD a todas las áreas del registro. La Sección de Catalogación de la IFLA aprobó esta primera edición de las ISBD(CF).	ISBD(CF) en colaboración con la Section on Information Technology. A finales de 1994 se creó un Grupo de Trabajo formado por miembros activos y asesores de ambas secciones. Con el patrocinio de la IFLA y el soporte económico del Research Libraries Group, el grupo de trabajo se reunió en abril de 1995 y dispuso las bases para la preparación del borrador de la segunda edición. Este borrador se distribuyó para su revisión mundial y, como respuesta, se recibieron más de treinta comentarios que sumaban más de ciento diez páginas de texto procedentes de lectores particulares, de asociaciones de bibliotecarios y de bibliotecas nacionales. Como resultado de esta revisión se han hecho muchas mejoras en el texto, incluido el reconocimiento de la necesidad de un término nuevo que caracterice el material del que se trata: este término es "recurso electrónico", el cual, a causa del conjunto de materiales que esta ISBD pretende abarcar, se considera más propio que el término "archivo de ordenador" que se utilizaba previamente. La ISBD(ER) incluye un índice y cinco apéndices. El Apéndice A ofrece una disposición para la técnica especial de la descripción en más de un nivel. El Apéndice B ofrece un breve esquema de la forma en que
---	--

	se han de transcribir los datos cuando están parcialmente en una escritura que se lee de derecha a izquierda y parcialmente en una escritura que se lee de izquierda a derecha. El Apéndice C enumera la designación general de la clase de material, las designaciones del recurso y las designaciones específicas de material recomendadas, con sus definiciones. El Apéndice D ofrece las abreviaturas recomendadas por los registros en cada lengua. El Apéndice E da ejemplos formulados para ilustrar el resultado de aplicar las disposiciones de la ISBD en todas las áreas de un registro
--	---

<http://www.geocities.com/zaguan2000/electronic> 1

NOTAS PRELIMINARES					
0.1. Alcance, objetivos y uso					
● ISBD (CF)			ISBD (ER)●		
● 0.1.1 Alcance			0.1.1 Alcance●		
La	Descripción	Bibliográfica	La	Descripción	Bibliográfica
Internacional	Normalizada	para Archivos de	Internacional	Normalizada	para Recursos
Ordenador	-denominada	a partir de ahora	Electrónicos	-denominada	a partir de ahora
ISBD(CF)-	especifica	los elementos	ISBD(ER)-	especifica	los elementos
necesarios	para	la descripción e	necesarios	para	la descripción e

identificación de publicaciones monográficas legibles por ordenador, asigna un orden a los elementos de la descripción y especifica un sistema de puntuación para la misma. Sus especificaciones se refieren, en primer lugar, a los registros bibliográficos producidos por los centros bibliográficos nacionales (en las ediciones de las bibliografías nacionales impresas, en otros registros impresos y en los ficheros asociados de datos legibles por ordenador) y, en segundo lugar, a registros bibliográficos de otros centros catalogadores, tanto en forma impresa o legible por ordenador. (En el caso de datos bibliográficos almacenados en soporte legible por ordenador, las ISBDs determinan la forma de la presentación de los datos tanto en línea como en productos impresos y no la forma, el orden, etc., de los ficheros). Los archivos de ordenador monográficos son materiales codificados para su manipulación por medio de un ordenador. Esto comprende archivos de dos tipos distintos: archivos de datos (información en forma de números, letras, símbolos o combinación de todos ellos) y programas (instrucciones o rutinas para la ejecución de ciertas tareas que incluyen el proceso de los datos).	identificación de estos documentos, les asigna un orden y especifica un sistema de puntuación para la descripción. Sus disposiciones se refieren, en primer lugar, a los registros bibliográficos producidos por los centros bibliográficos nacionales (en las ediciones de las bibliografías nacionales impresas, en otros registros impresos y en los recursos asociados de datos legibles por ordenador) y, en segundo lugar, a registros bibliográficos de otros centros catalogadores, ya sea en forma electrónica o impresa. Los recursos electrónicos son materiales controlados por ordenador, incluso los que requieren la utilización de un periférico (por ejemplo, un lector de CD-ROM) conectado a un ordenador; los documentos se pueden utilizar de manera interactiva o no. Se incluyen dos clases de recursos: de datos (información en forma de números, letras, gráficos, imágenes y sonido o su combinación) y programas (instrucciones o rutinas para la ejecución de ciertas tareas que incluyen el proceso de los datos). Además, pueden combinarse para incluir datos y programas electrónicos (por ejemplo, servicios en línea, multimedia interactivos) Por los objetivos de la catalogación,
---	--

Por razones catalográficas, los archivos de ordenador están tratados en las ISBD(CF) de dos maneras distintas según que el acceso a los mismos sea local o remoto. Se utiliza acceso local para significar que puede describirse un soporte físico. El usuario debe insertar dicho soporte (por ejemplo, un disco, casete o cartucho) en un ordenador, generalmente un microordenador. Se entiende por acceso remoto que el usuario no debe manejar un soporte físico -de forma característica sólo puede proporcionarse acceso mediante el uso de un dispositivo de entrada y salida conectado a un sistema informático (por ejemplo, un fichero en una red o en un entorno de tiempo compartido) o por el uso de archivos almacenados en un disco duro. Se toma esta definición ya que es aplicable a la mayoría de los archivos que están generalmente disponibles, incluyendo aquellos a los que se accede vía una red [local] o vía telecomunicaciones. Se incluyen igualmente los ficheros producidos y/o generados para distribución limitada, para demanda previo pago de tasa o hechos a la medida. Un fichero residente en la memoria permanente de un ordenador (ROM) o en el soporte lógico se entiende que forma parte del dispositivo en el cual está almacenado y,	los recursos electrónicos están tratados en las ISBD(ER) de dos maneras distintas según el acceso sea local o remoto. El acceso es local cuando puede describirse el soporte físico. El usuario debe insertar dicho soporte (por ejemplo, un disco magnético u óptico, casete o cartucho) en un ordenador, o en un periférico conectado a un ordenador -generalmente un microordenador. El acceso es remoto cuando el usuario no puede manipular un soporte físico -en general, el acceso se hace a partir de un dispositivo de entrada y salida de datos (por ejemplo, un terminal) conectado a un sistema informático (por ejemplo, un recurso en una red) o por el uso de archivos almacenados en un disco duro en otros dispositivos de almacenaje. Esta definición se aplica a la mayoría de los recursos que están normalmente disponibles por los procedimientos habituales, incluyendo los multimedia interactivo , o los recursos disponibles en redes o accesibles por telecomunicaciones. Se incluyen igualmente los ficheros producidos o generados para distribución limitada, para demanda previo pago de tasa o por encargo. Un recurso residente en la memoria permanente de un ordenador (ROM) se considera que forma parte del dispositivo en el cual está almacenado y, si se
---	--

si se cataloga, debe considerarse como un fichero que requiere acceso remoto. Juguetes programados, calculadoras y otros objetos programados se considera que están fuera del alcance de las ISBDs.	cataloga, debe considerarse como un recurso que requiere acceso remoto. Juguetes programados, calculadoras y otros objetos programados se considera que están fuera del alcance de las ISBD(ER)
La ISBD(CF) es una de las ISBDs ya publicadas. Las restantes se refieren a material no librario (ISBD(NBM)), publicaciones seriadas (ISBD(S)), monografías (ISBD (M), material cartográfico (ISBD(CM)), monografías anteriores a 1801 (ISBD(A)) y música impresa (ISBD (PM)). Cada ISBD pretende incorporar un conjunto coherente de especificaciones para su propio tipo de publicación, pero no ha habido ningún intento de hacer una ISBD excluyente.	La ISBD(ER) es una de las diversas ISBDs ya publicadas. Las restantes cubren el material no librario (ISBD(NBM)), publicaciones seriadas (ISBD(S)), monografías (ISBD (M), material cartográfico (ISBD(CM)), monografías anteriores a 1801 (ISBD(A)) y música impresa (ISBD (PM)). Cada ISBD pretende incorporar un conjunto coherente de especificaciones para su propio tipo de publicación, pero no se ha tratado de que ninguna de las ISBD sea excluyente.
Los usuarios necesitarán en ocasiones referirse a varias ISBD cuando, por ejemplo, la 'publicación que se va a describir presente características descritas en otras ISBDs como, por ejemplo, una publicación informática legible por ordenador publicado como una serie o con una monografía aneja. Todas las ISBDs se basan en las ISBD(G).	Cuando un recurso electrónico tiene características recogidas en otras ISBD (por ejemplo, una publicación en serie electrónica, un mapa digitalizado), se recomienda que el centro bibliográfico utilice en primer lugar las disposiciones de la ISBD(ER) y, cuando sea necesario, aplique las disposiciones de otras ISBD. No obstante, algunos centros bibliográficos preferirán describir estos recursos utilizando otra ISBD apropiada al material y complementara con la aplicación de las ISBD(ER). Todas las ISBD se basan
La ISBD(CF) está fundamentalmente dedicada a solventar las necesidades actuales de las bibliotecas, las agencias bibliográficas nacionales y los centros de investigación. Sin	en la ISBD general (ISBD(G))

embargo, puede precisar cierto grado de elaboración antes de aplicarse a categorías anticuadas de material. Al mismo tiempo, y ya que los ficheros descritos en las ISBD(CF) son el resultado de tecnologías volátiles, las estipulaciones específicas de estas ISBDs, particularmente las del área 3 (área del tipo y dimensiones del fichero) y del área 5 necesitarán corregirse para que sean apropiadas para manejar adecuadamente características adicionales de los ficheros o formas recientemente desarrolladas de material.	en la ISBD general (ISBD(G)). Los recursos electrónicos tienen cada vez más ediciones diferentes. Se considera una nueva edición cuando se determina que existen diferencias significativas en el contenido intelectual o artístico del documento. En estos casos, se creará un registro bibliográfico separado. Un documento que no presenta diferencias significativas con respecto a otro no justifica un registro bibliográfico separado; en todo caso, un centro bibliográfico puede optar por crear registros bibliográficos múltiples. En el caso de recursos de acceso remoto (por ejemplo, servicios en línea) que se actualizan a menudo, se recomienda que el centro bibliográfico omita la mención de edición en el área 2 y ofrezca una nota o notas apropiadas en el área 7 Cuando un documento esté disponible en diferentes tipos de soportes físicos, o en soportes de salidas o formatos de visualización diferentes, se pueden seguir uno de los siguientes métodos de descripción: (1) Los diferentes soportes físicos se pueden describir en el mismo registro bibliográfico, en una línea aparte para cada soporte o agrupados en una línea continua de descripción; o
---	--

	(2) cada soporte físico diferente se puede describir en soportes físicos separados. De todos modos, en el caso de un documento multimedia interactivo que se presente en dos o más soportes físicos, cada uno de ellos se describirá siguiendo la opción del método (1), según la cual, cada soporte ocupa una línea aparte de la descripción en el mismo registro bibliográfico Las ISBD(ER) se ocupa sobre todo de las necesidades actuales de los centros bibliográficos nacionales, de las bibliotecas y de los centros de recursos. Puede ser necesaria alguna adaptación previa antes de aplicarla a categorías obsoletas de material. Al mismo tiempo, como los recursos descritos en las ISBD(ER) son producto de tecnologías volátiles, las especificaciones de estas ISBDs, particularmente las del área 3 (tipo y extensión del recurso) y del área 5 (descripción física) se habrán de corregir adecuadamente para adaptarlas a las características adicionales de los recursos o a nuevas formas de material.
ISBD (CF)	ISBD (ER) ●
● 0.1.2 Objetivo El objetivo fundamental de las ISBDs	0.1.2 Objetivo ● El objetivo fundamental de las ISBDs

es proporcionar especificaciones para una catalogación descriptiva compatible en todo el mundo, con el fin de facilitar el intercambio internacional de registros bibliográficos entre centros bibliográficos nacionales y la comunidad bibliotecaria y del mundo de la información. Al especificar los elementos que comprende la descripción bibliográfica y al prescribir el orden en el que se deben presentar esos elementos y la puntuación que los delimita, las ISBDs aspiran a (a) lograr que los registros procedentes de distintas fuentes puedan intercambiarse de tal manera que los registros producidos en un país puedan incorporarse fácilmente a los catálogos de otras bibliotecas o en las listas bibliográficas de otros países; (b) facilitar la interpretación de los registros frente a las barreras lingüísticas de modo que registros producidos para usuarios de una lengua puedan ser interpretados por usuarios de otra; y (c) facilitar la conversión de los registros bibliográficos a soporte legible por ordenador.	es proporcionar especificaciones para una catalogación descriptiva compatible en todo el mundo, con el fin de facilitar el intercambio internacional de registros bibliográficos entre centros bibliográficos nacionales y, en general, en toda la comunidad internacional bibliotecaria y documental. Al especificar los elementos que comprende la descripción bibliográfica y al prescribir el orden en el que se deben presentar esos elementos y la puntuación que los delimita, las ISBDs aspiran a (A) lograr que los registros procedentes de distintas fuentes sean intercambiables, de tal manera que los registros procedentes de distintas fuentes puedan incorporarse fácilmente a los catálogos de bibliotecas o en bibliografías de cualquier otro país; (B) facilitar la interpretación de los registros frente a las barreras lingüísticas de modo que registros producidos para usuarios de una lengua puedan ser interpretados por usuarios de otra; y (c) facilitar la conversión de los registros bibliográficos a formato electrónico
---	--

0.3. Esquema comparativo de la ISBD(G) y de la ISBD(CF)						
	Generales		Archivos de ordenador		Recursos electrónicos	
AREA	Puntuación	ELEMENTOS	Puntuación	ELEMENTOS	Puntuación	ELEMENTOS
Nota: Cada área, excepto la primera, irá precedida de punto, espacio, raya, espacio (.-)						
1. Área de título y mención de responsabilid		1.1 Título propiaamente dicho		1.1 Título propiaamente dicho		1.1 Título propiaamente dicho
	[]	1.2 Designación general de 1a clase de material	[]	1.2 Designación general de la clase de material (opcional)	[]	1.2 Designación general de la clase de material (opcional)
	=	1.3 Título paralelo	=	*1.3 Título paralelo	=	*1.3 Título paralelo

	:	1.4 Información complementaria del título	:	*1.4 Subtítulo y/o información complementaria del título	:	*1.4 Subtítulo y/o información complementaria del título
	/ ;	1.5 Menciones de responsabilidad Primera mención Segunda y subsiguientes menciones	/ ;	1.5 Mención de responsabilidad Primera mención *Segunda y subsiguientes menciones	/ ;	1.5 Mención de responsabilidad Primera mención *Segunda y subsiguientes menciones
2. Área de edición		2.1 Mención de edición		2.1 Mención de edición		2.1 Mención de edición
	=	2.2 Mención paralela de edición	=	*2.2 Mención paralela de edición (opcional)	=	*2.2 Mención paralela de edición (opcional)

	/ ;	2.3 Menciones de responsabilidad relativas a la edición Primera mención Segunda y subsiguientes menciones	/ ;	2.3 Menciones de responsabilidad relativas a la edición Primera mención *Segunda y subsiguientes menciones	/ ;	2.3 Menciones de responsabilidad relativas a la edición Primera mención *Segunda y subsiguientes menciones
	,	2.4 Mención adicional de edición	,	*2.4 Mención adicional de edición	,	*2.4 Mención adicional de edición

	/ ;	2.5 Menciones de responsabilidad relativas a una mención de edición adicional Primera mención Segunda y subsiguientes menciones	/ ;	2.5 Menciones de responsabilidad relativas a una mención adicional de edición Primera mención *Segunda y subsiguientes menciones	/ ;	2.5 Menciones de responsabilidad relativas a una mención adicional de edición Primera mención *Segunda y subsiguientes menciones
3. Área de datos específicos de la clase de documento				3.1 Designación específica de la clase de archivo		3.1 Designación del recurso
				3.2.Extensión del archivo (opcional)		3.2.Extensión del recurso (opcional)

4. Área de publicación, distribución, etc	;	4.1. Lugar de publicación y/o distribución Primer lugar Segundo u subsiguientes lugares	;	4.1. Lugar de publicación y/o distribución Primer lugar Segundo u subsiguientes lugares	;	4.1 Lugar de publicación y/o distribución Primer lugar Segundo u subsiguientes lugares
	:	4.2. Nombre del editor y/o distribuidor	:	4.2. Nombre del editor y/o distribuidor	:	4.2. Nombre del editor y/o distribuidor
	[]	4.3 Mención de la función del editor y/o distribuidor	[]	*4.3 Mención de la función del editor y/o distribuidor	[]	*4.3 Mención de la función del editor y/o distribuidor

	,	4.4 Fecha de publicación y/o distribución	,	4.4 Fecha de publicación y/o distribución	,	4.4 Fecha de publicación y/o distribución
	(	4.5 Lugar de impresión	(	*4.5 Lugar de fabricación	(	*4.5 Lugar de fabricación
	:	4.6 Nombre del impresor	:	*4.6 Nombre de fabricante	:	*4.6 Nombre de fabricante
	,)	4.7 Fecha de la impresión	,)	4.7 Fecha de fabricación	,)	4.7 Fecha de fabricación
5. Área de descripción física		5.1 Extensión y designación específica del material		5.1 Extensión y designación específica del material		5.1 Extensión y designación específica del material

	:	5.2 Mención de otras característica físicas	:	5.2 Mención de otras características físicas	:	5.2 Mención de otras características físicas
	;	5.3 Dimensiones	;	5.3 Dimensiones	;	5.3 Dimensiones
	+	5.4 Material anejo	+	*5.4 Material anejo (opcional)	+	*5.4 Material anejo (opcional)
6. Área de serie		6.1. Título propiamente dicho de la serie		6.1. Título propiamente dicho de la serie o subserie		6.1. Título propiamente dicho de la serie o subserie
	=	6.2 Títulos paralelos de la serie	=	*6.2 Títulos paralelos de la serie o subserie	=	*6.2 Títulos paralelos de la serie o subserie

	:	6.3 Información complementaria del título	:	*6.3 Información complementaria del título de la serie o subserie (opcional)	:	*6.3 Información complementaria del título de la serie o subserie (opcional)
	/ ;	6.4 Menciones de responsabilidad relativas a la serie Primera mención Segunda y subsiguientes menciones	/ ;	6.4 Menciones de responsabilidad relativas a la serie o subserie Primera mención Segunda y subsiguientes menciones	/ ;	6.4 Menciones de responsabilidad relativas a la serie o subserie Primera mención Segunda y subsiguientes menciones

	,	6.5 ISSN (Número Internacional Normalizado de Publicaciones Seriadas)	,	6.5 ISSN (Número Internacional Normalizado de Publicaciones Seriadas) de la serie o subserie	,	6.5 ISSN (Número Internacional Normalizado de Publicaciones Seriadas) de la serie o subserie
	;	6.6 Número de serie	;	6.6 Número de serie o subserie	;	6.6 Número de serie o subserie
	.	6.7 Número y/o título de la subserie				
	=	6.8 Títulos paralelos de la subserie				

	:	6.9 Subtítulo y/o información complementaria del título de la subserie				
	/ ;	6.10 Menciones de responsabilidad relativas a la subserie Primera mención Segunda y subsiguientes menciones				
	,	6.11 ISSN de la subserie				

	;	6.12 Número de la subserie				
7. Área de notas		7. Área de notas		7. Área de notas		
8. Área del numero normalizado (o alternativo) y de las condiciones de adquisición		8.1 Número normalizado (o alternativo)		8.1 Número normalizado (o alternativo)		8.1 Número normalizado (o alternativo)
	=	8.2 Título clave	=	8.2 Título clave	=	8.2 Título clave
	:	8.3 Condiciones de adquisición y/o precio	:	8.3 Condiciones de adquisición y/o precio	:	8.3 Condiciones de adquisición y/o precio

	()	8.4 Aclaraciones (en diferentes posiciones)				
--	-----	--	--	--	--	--

● Notas generales sobre el esquema de la ISBD(CF):	● Notas generales sobre el esquema de la ISBD(ER):
A. Los elementos opcionales se indican como tales	A. Los elementos opcionales se indican como tales
B. Los elementos precedidos de asterisco se pueden repetir si es necesario.	B. Los elementos precedidos de asterisco se pueden repetir si es necesario.
C. Las áreas 6 (Serie), 7 (Notas) y 8 (Número normalizado, etc.) se pueden repetir si es necesario. Además, el área 5 (Descripción física) se puede repetir en algunos casos (véase área 5, nota introductoria).	C. Las áreas 6 (Serie), 7 (Notas) y 8 (Número normalizado, etc.) se pueden repetir si es necesario. Además, el área 5 (Descripción física) se puede repetir en algunos casos (véase área 5, nota introductoria).
D. En el esquema anterior, los términos "primera mención ..." "segunda y subsiguientes menciones..." y similares, indican el orden en el que estas menciones aparecen en la descripción y no tienen otra connotación.	D. En el esquema anterior, los términos "primera mención ..." "segunda y subsiguientes menciones..." y similares, indican el orden en el que estas menciones aparecen en la descripción y no tienen otro significado.
E. En las ISBD(CF) no están previstas especificaciones para el elemento 8.2 del esquema de la ISBD (G) (Título clave). Las especificaciones relativas a las aclaraciones al número normalizado (o alternativo) o la mención de las condiciones de adquisición y/o	E. En las ISBD(ER) no se incluyen especificaciones para el elemento 8.2 del esquema de la ISBD (G) (Título clave). Las especificaciones relativas a las aclaraciones al número normalizado (o alternativo) o la mención de las condiciones de adquisición y/o

precio (elemento 8.4 de la ISBD (G)) se incluyen en los elementos S. 1 y 8.3, respectivamente, en lugar de especificarse como si se tratase de un elemento independiente.	precio (elemento 8.4 de la ISBD (G)) se incluyen en los elementos S. 1 y 8.3, respectivamente, en lugar de especificarse como si se tratase de un elemento independiente.
F. Siempre que la información normalmente asociada a un área o elemento aparezca en la publicación gramaticalmente vinculada como parte integrante de otro elemento o área, se transcribe como tal	F. Siempre que la información normalmente asociada a un área o elemento aparezca en la publicación gramaticalmente vinculada como parte integrante de otro elemento o área, se transcribe como tal

9.3 PROPUESTAS PARA MEJORAR EL ACCESO A RECURSOS ELECTRÓNICOS Y FACILITAR SU ORGANIZACIÓN

9.3.1 URI (UNIFORM RESOURCE IDENTIFIERS)

Para mejorar la búsqueda de recursos Internet y la localización y acceso a partir de motores de búsqueda y de directorios, existen una serie de iniciativas, algunas internacionales, que, a medio plazo, pueden colaborar en organizar Internet. En primer lugar, hay que citar los trabajos de Internet Engineering Task Force (IETF) para desarrollar diferentes “Uniform Resource Identifiers” (URIs). El primero de estos identificadores ya hace tiempo que funciona: “Uniform Resource Locator” (URL) ha permitido normalizar la dirección de los recursos Internet de acuerdo con un esquema estándar. Sin embargo, este localizador no es operativo cuando el recurso cambia de lugar, lo cual es muy frecuente en la red. Para solucionar este problema existen otros identificadores del recurso que sean independientes de su localización. El IETF trabaja en dos identificadores más: el “Uniform Resource Name” (URN), que funcionara de manera similar al ISBN de los libros y ofrecerá un nombre de manera unívoca a todos los objetos Internet, y el “Uniform Resource Characteristic” (URC), que será como una cita del recurso. Por otro lado OCLC ha desarrollado el PURL “Persistent Uniform Resource Locator” para controlar las

localizaciones de los recursos Internet de su catálogo colectivo, desarrollando la misma función que teóricamente tiene que hacer el URN. Evidentemente, URL y URN serán útiles no solamente para potenciar y mejorar las capacidades de búsqueda y acceso de los motores y directorios, sino que con la inclusión de estos datos también mejorarán los registros MARC de recursos Internet cargados en catálogos locales y nacionales, sobretodo si estos catálogos funcionan en entorno Web.

En segundo lugar, habría que citar otra línea de trabajo que tiene como objetivo la creación de metadatos para recursos Internet, pero este tema será tratado en la segunda parte del trabajo.

10. FORMATOS PARA LA DESCRIPCIÓN DE RECURSOS ELECTRÓNICO: ESTÁNDARES DE METADATOS

Los metadatos son definidos, generalmente, como datos sobre los datos. Más concretamente, son un conjunto estructurado de elementos que describen un “paquete de información” con el propósito de poder identificar, recuperar y usar la información. Un paquete de información se define como una instancia de información registrada, como un libro, un artículo, una revista electrónica, etc.

Los metadatos para la descripción son solo un tipo de metadatos, aunque uno de los más populares en el mundo bibliotecario. Existen metadatos para valorar y seleccionar contenidos como PICS (Platform for Internet Content Selection), para la gestión de derechos y la propiedad intelectual (INDECS) y para cuestiones administrativas (A-Core). Los metadatos descriptivos van desde los muy ricos y altamente técnicos como el formato MARC y las cabeceras TEI, hasta los relativamente simples elementos que componen Dublin Core.

Los metadatos pueden estar incluidos en el propio documento al cual describen o estar fuera de él.

10.1 TEXT ENCODING INITIATIVE. TEI

Todo documento que forme parte de una colección bibliotecaria, en forma electrónica o de cualquier otro modo, debe cumplir unos requisitos de estabilidad, fiabilidad e información bibliográfica para que pueda ser útil para la comunidad de usuarios, no solo para la persona que haya creado el documento. En el ámbito académico de las humanidades han estado interesados durante mucho tiempo en las posibilidades de almacenar y recuperar material textual en formato electrónico, pero ninguno de los documentos electrónicos originalmente codificados satisfacía estos requisitos. En concreto, los esfuerzos por digitalizar materiales raros y viejos manuscritos, en la mayoría de los casos no eran útiles para la comunidad académica porque estos documentos electrónicos contenían errores ortográficos (como resultado del proceso de conversión de imagen a texto), omitían la información sobre la publicación y la edición del texto impreso que era usado como base del documento electrónico.

Una de las aplicaciones de los textos electrónicos en humanidades es la posibilidad de compilar automáticamente variantes textuales para nuevas ediciones de trabajos antiguos, y otra es hacer accesible a la comunidad académica material que, debido a la localización del documento, sería muy difícil para ellos poder consultarlo físicamente. Por lo tanto, la precisión en la transcripción a los materiales electrónicos, y la presencia de detallada y precisa información bibliográfica tienen una gran importancia. La información bibliográfica que requiere la comunidad académica sobre los materiales antiguos debe ser más descriptiva que una referencia bibliográfica para fines bibliotecarios.

Además de la información bibliográfica proporcionada por la catalogación propia de un texto electrónico, la comunidad académica en humanidades necesita una descripción más detallada de los aspectos físicos de los libros o manuscritos que estudian. Por ejemplo, información sobre en que lugar se encuentra una página rota en las ediciones originales o manuscritos del mismo trabajo, el aspecto del título de la página y como las líneas de los versos o de la prosa están ordenados en una página de una obra de teatro o poema, estos son solo algunas de los aspectos imprescindibles para alguien que trate de compilar una nueva edición de un trabajo antiguo. A pesar de que este tipo de metainformación no esta entre los

objetivos de la catalogación, es necesaria, por ejemplo, para que sea útil una versión electrónica de una obra de Shakespeare. Por tanto, los textos electrónicos en humanidades crean problemas sobre la cantidad y el tipo de metainformación deben proporcionar.

En el contexto de esta necesidad específica de metainformación precisa y detallada que debe acompañar a los textos electrónicos, se estableció en 1987 el Text Encoding Initiative (TEI), un proyecto conjunto de la Association for Computers and the Humanities, la Association for Computational Linguistics y la Association for Literary and Linguistic Computing. El objetivo del proyecto TEI la definición de una guía genérica para la representación de materiales textuales en formato electrónico.

10.1.1 SGML Y TEI

TEI especifica un número de elementos requeridos para cada texto, en concreto un encabezamiento formado por cuatro elementos obligatorios que contienen la información bibliográfica y un cuerpo que puede contener texto, imágenes y otros objetos. Con el fin de proporcionar reglas para el codificado e intercambio de textos electrónicos, el esquema de TEI utiliza SGML (Standard Generalized Mark-up Language) y su conjunto de etiquetas de marcado para codificar el material textual. SGML es un standard internacional (ISO 8879) para codificar información electrónica que define los métodos independientes de los dispositivos y sistemas para representar textos y otros objetos y está más relacionado con el contenido y su ordenación que con su formato o apariencia. SGML está constituido por varios subconjuntos de etiquetas, llamados Document-Type Definition (DTD), que especifican los requerimientos de contenido y conjuntos de etiquetas para cada tipo de documento. HTML (Hypertext Mark-up Language), que es un tipo de DTD de SGML, es el subconjunto de etiquetas más conocido de SGML. TEI es otro subconjunto.

Igual que otras aplicaciones de SGML, TEI es independiente de plataformas, sistemas, aplicaciones y dispositivos y se ajusta a los protocolos de red para el intercambio de información. Cualquier documento codificado en SGML incluye tres partes: la declaración de SGML (por ejemplo, una sentencia declarando que se trata de un documento SGML), un DTD (por ejemplo, el conjunto de etiquetas de SGML utilizado para un documento concreto) y la instancia del documento (por ejemplo, un documento codificado

de acuerdo a las reglas declaradas en el DTD). Con SGML, y por lo tanto con TEI, cada documento electrónico lleva consigo su propia metainformación o metadatos.

El esquema TEI define como escribir clases específicas de DTDs de SGML que especifican como cada documento electrónico es codificado en SGML. De acuerdo con TEI, los requisitos de contenido especifican requisitos mínimos de marcado para un nivel bajo de codificación, pero también los debe especificar para codificaciones muy complejas apropiadas para el marcado detallado necesario en las investigaciones de humanidades. Cada documento electrónico codificado de acuerdo con TEI también contiene una cabecera TEI además del cuerpo del documento. Las cabeceras TEI contienen metainformación sobre el documento electrónico, contiene información bibliográfica utilizable por la bibliotecas, y que han sido diseñadas con el asesoramiento de bibliotecarios.

Aunque las normas TEI estaban relacionadas inicialmente con textos impresos como los estudiados en literatura, lingüística e historia, y diseñadas como un esquema de codificación para estructuras textuales complejas, los textos electrónicos deben ser entendidos de un modo amplio: TEI abarca recursos textuales y no textuales como los contenidos en una base de datos de investigación o los componentes de las publicaciones que no se encuentran en soporte papel.

A medida que el proyecto iba desarrollándose, cada vez iban apareciendo más textos electrónicos en colecciones como las del Electronic Text Center (ETC) en la Universidad de Virginia, o la del Oxford Text Archive (OTA) de la Universidad de Oxford y la del Center for Electronic Text in the Humanities (CETH), en las Universidades de Princeton y Rutgers. Estaba claro que la comunidad informática en humanidades podía beneficiarse de la experiencia de los catalogadores y de las reglas de catalogación que se habían utilizado en las bibliotecas con el propósito de compartir registros.

Los catalogadores deberían aplicar también la experiencia obtenida por la demanda de textos electrónicos en humanidades: los esfuerzos de catalogar y codificar metainformación en textos electrónicos de humanidades proporciona directrices para la catalogación de materiales electrónicos, particularmente recursos disponibles en Internet.

10.1.2 EL ENCABEZADO DE TEI

El encabezado TEI está adjunto a un documento electrónico, es una etiqueta que contiene directrices sobre la estructura lógica del documento. También debe ser independiente del texto al que pertenezca, como un vehículo de metainformación sobre el material electrónico. Es esta capacidad de producir registros independientes de metainformación, siguiendo unas reglas, la que convierten al encabezado TEI en algo interesante para los catalogadores: en efecto, las reglas diseñadas por TEI sirven como normas para utilizar SGML en la descripción de documentos electrónicos de un modo adecuado tanto para los requerimientos de los formatos electrónicos como para la catalogación más tradicional.

El encabezado TEI incluye cuatro partes, siendo obligatoria solo una de ellas, pero todas pertenecen a temas específicos de la catalogación de formatos electrónicos.

Descripción del archivo y del formato fuente que contiene el texto (obligatorio)

Sistema de codificación aplicado (opcional)

Descripción: clasificación, palabras clave, encabezamientos de materia (opcional)

Revisión (opcional)

Aunque solo es obligatoria la descripción, es recomendable la utilización de los otros apartados, especialmente en el caso de encabezados TEI independientes porque pueden contener información difícil de describir utilizando las reglas AACR2 para ficheros de ordenador.

10.1.3 ANÁLISIS DE UN ENCABEZADO TEI

Consideraciones generales

Los elementos descripción del archivo y la descripción (clasificación, palabras clave, encabezamientos de materia) son los más familiares para los catalogadores ya que contienen los metadatos que constituyen la base de un registro de catálogo. La descripción del archivo contiene información descriptiva tal como el título o información sobre la

publicación. También contiene elementos que permiten codificar información sobre el texto o formato fuente a partir del cual se ha generado el texto electrónico. La descripción también contiene encabezamientos de materia y esquemas de clasificación según los cuales se le asignará a los textos electrónicos una signatura topográfica o código de acceso. El elemento de la descripción debe contener también la información necesaria para los estudios humanísticos que puede ir desde el idioma en que ha sido escrito el texto hasta información sobre el contexto social.

Encabezamientos de materia y clasificación

No todos los documentos codificados con TEI tienen esta información debido a que el elemento de la descripción del perfil no es obligatorio, aunque si es altamente recomendado de acuerdo con el esquema TEI.

La presencia o ausencia de encabezamientos de materia en un documento codificado con TEI no debe, sin embargo, ser un criterio para valorar la utilidad del esquema TEI. La descripción permite incluir encabezamientos de materia, y su inclusión depende de los propósitos con que ha sido creado el documento electrónico.

Descripción del archivo

Este elemento incluye la siguiente información:

Título

Edición

Extensión

Publicación

Serie

Notas

Descripción de la fuente

10.2 PICS

El World Wide Web es un sistema abierto y dinámico con un enorme potencial para la edición, el comercio y la comunicación. Su carácter abierto y universal son algunas de sus cualidades, pero también son la causa de inestabilidad y caos. Proporcionar servicios de bibliotecas digitales a través del web requiere una infraestructura para permitir una política de decisiones, requerimientos funcionales y facilidades de producción de una auténtica biblioteca.

Entre los requisitos funcionales de una biblioteca se encuentran la selección y organización de recursos dirigida a satisfacer las necesidades de información de los usuarios de la biblioteca. Actualmente, el web es un espacio de información indiferenciado caracterizado por un rápido y caótico crecimiento.

El Platform for Internet Content Selection (PICS) es una parte de esta infraestructura. PICS, una tecnología que soporta la asociación de etiquetas descriptivas con recursos de Internet, es una iniciativa del W3C. OCLC es un miembro y participante activo.

En las primeras versiones de las especificaciones de PICS, las etiquetas eran atributos (violencia, por ejemplo) con una escala numérica asociada (de 0 a 10, por ejemplo). Este era un sistema útil para obtener un filtro fiable pero no era especialmente útil para un gran número de metadatos (datos sobre datos, e.g registros catalográficos)

Las especificaciones de PICS están sufriendo actualmente importantes revisiones que le conferirán una riqueza expresiva suficiente para captar descripciones cercanas a la noción de catalogación existente en la comunidad bibliotecaria. Cuando esas revisiones hayan finalizado, los vendedores de software podrán proporcionar herramientas para la creación, manipulación y mantenimiento de metadatos para recursos electrónicos.

Semántica vs Sintaxis

PICS no define la semántica de ningún dato contenido en el standard. Más bien define la sintaxis de transmisión de los datos, la especificación formal de cómo es

expresada la semántica de los datos contenidos en el standard. Esta distinción es similar a la existente entre las AACR2 y el formato MARC. Las AACR2 definen la semántica o significado del contenido de un registro bibliográfico, mientras que MARC especifica la estructura del registro que contiene el significado.

Una de las características más importantes de PICS es que la semántica de los metadatos contenidos en el standard son definidos por otros, por lo tanto cualquier comunidad puede crear metadatos con un significado determinado. El Dublín Core, por ejemplo, es un conjunto de elementos creados para obtener mejores descripciones de recursos en Internet. Uno de los mayores impedimentos para su utilización como modelo de descripción ha sido la falta de una sintaxis ampliamente aceptada. PICS proporciona una arquitectura para soportar tal sintaxis.

10.3 DUBLIN CORE METADATA INITIATIVE

10.3.1 DUBLIN CORE

El Dublin Core es un conjunto de metadatos formado por 15 elementos dirigidos a facilitar la recuperación de recursos electrónicos. Fue concebido originalmente para que fuese el propio autor del recurso Web, quien realizase la descripción del mismo. También ha atraído la atención de comunidades que realizan descripciones formales de recursos tales como museos y bibliotecas.

Los grupos de trabajo implicados en el proyecto Dublín Core han reunido a expertos del mundo bibliotecario, comunidades que investigan sobre bibliotecas digitales y a gran variedad de especialistas en contenidos. La construcción de un consenso internacional e interdisciplinario alrededor de un conjunto de elementos es el principal logro de la evolución de Dublín Core. Este progreso representa el saber y la experiencia colectiva de los distintos grupos que trabajan en el área de la descripción de recursos.

Dublín Core es, sin duda, un fuerte candidato para la descripción de recursos electrónicos, debido a las siguientes razones:

- **Simplicidad**

Dublín Core se creó para que pudiese ser utilizado por personas no expertas en el ámbito de la catalogación. Por lo tanto, lo que se pretende es que los creadores de recursos web o los webmasters, desconocedores de la catalogación, puedan utilizar Dublín Core para describir recursos, y conseguir de esta manera que los recursos sean más “visibles” para los motores de búsqueda y sistemas de recuperación de información.

La mayor parte de los 15 elementos que componen Dublín Core poseen un significado fácilmente comprensible y representan lo que debería ser un denominador común en la descripción de recursos (sería muy similar, en este sentido, a una ficha de catálogo). El objetivo de Dublín Core no es sustituir a modelos de descripción más ricos y complejos como pudiera ser la

catalogación siguiendo el formato MARC, más bien lo que se intenta es proporcionar un conjunto de elementos de descripción que pueden ser usados tanto por catalogadores como cualquier persona para describir recursos de una forma simple.

- **Interoperabilidad semántica**

En Internet existen distintos modelos de descripción, los que dificulta la búsqueda de información a través de disciplinas cuyas fronteras no están claramente delimitadas. Por ejemplo, las bibliotecas, museos y en el ámbito de los sistemas de información geográfica, se utilizan estándares diferentes en la descripción de recursos. Esto refleja las distintas necesidades de descripción que existen en estos tres ámbitos y pone de manifiesto el hecho de que estos estándares han evolucionado de forma independiente.

Llegados a un cierto nivel de descripción, los conjuntos de elementos son distintos porque se están describiendo cosas distintas. La mayor parte de los autores no asocian un atributo “nubosidad” a sus documentos, sin embargo, si estamos describiendo imágenes de satélite, este sería un aspecto importante a tener en cuenta.

La cuestión es que la mayor parte de los recursos comparten un conjunto de atributos que son similares aunque estemos hablando de disciplinas diferentes, pero tienen nombres distintos simplemente porque se han desarrollado de forma independiente y en momentos más o menos distantes en el tiempo.

La promoción de un conjunto de elementos con un significado común, mejoraría las perspectivas en la búsqueda interdisciplinar de información, unificando los atributos relacionados. Por ejemplo, tanto el autor como el creador pueden ser considerados un solo atributo con el objetivo de localizar un recurso. Dublin Core pretende ser este conjunto de elementos común.

- **Consenso internacional**

Es fundamental tener conciencia del alcance internacional de la localización de recursos en Internet para el desarrollo de una adecuada infraestructura. Dublín Core se ha beneficiado de la participación activa de diferentes países alrededor del mundo.

- **Flexibilidad**

A pesar de que Dublín Core surge ante la necesidad de que la descripción de los recursos sean realizadas por sus propios autores, Dublín Core también ha dirigido su atención hacia los ámbitos de descripción formal de documentos. A medida que aumenta la diversidad y el volumen de los recursos en el Web, las bibliotecas y museos, reconocen el valor de la presencia de metadatos en los propios recursos.

Dublín Core, en manos de catalogadores expertos, puede representar una alternativa a modelos de descripción más elaborados tales como el formato MARC. Dublín Core proporciona suficiente flexibilidad para codificar estructuras adicionales para tales aplicaciones.

- **Modularidad de metadatos en el Web**

La gran diversidad de metadatos necesarios en el Web requieren un entorno que soporte la coexistencia de paquetes de metadatos desarrollados y mantenidos de manera independiente. Dublín Core está dirigido específicamente a la localización de recursos, pero podría tener otras posibles aplicaciones (metadatos administrativos, etc).

El reconocimiento de la conveniencia de este tipo de modularidad ha conducido al desarrollo de Dublín Core desde el Warwick Workshop, siendo formalizado como Warwick Framework.

10.3.2 ELEMENTOS DE DUBLIN CORE:

Los elementos poseen nombres descriptivos que pretenden transmitir un significado semántico a los mismos. Para promover una interoperabilidad global, una descripción del valor de algunos elementos podrá ser asociada a vocabularios controlados. Se asume que otros vocabularios controlados serán desarrollados para asegurar esta interoperabilidad en dominios específicos.

Cada elemento es opcional y puede repetirse. Además, los elementos pueden aparecer en cualquier orden.

Aunque algunos entornos, como HTML, no diferencian entre mayúsculas y minúsculas, es recomendable escribir correctamente cada metadato según su definición para evitar conflictos con otros entornos, como XML.

Los elementos se pueden clasificar en tres grupos, según la clase o el ámbito de la información que se guarda en ellos:

- Elementos relacionados con el contenido del recurso
- Elementos relacionados con el recurso cuando es visto como una propiedad intelectual
- Elementos relacionados con la instanciación del recurso

Contenido	Propiedad Intelectual	Instanciación
Title	Creator	Date
Subject	Publisher	Type
Description	Contributor	Format
Source	Rights	Identifier
Language		
Relation		
Coverage		

DC.Title (Título): Nombre dado a un recurso, usualmente por el autor.

DC.Creator (Autor o creador): Persona u organización responsable de la creación del contenido intelectual del recurso. Por ejemplo, los autores en el caso de documentos escritos, artistas, fotógrafos e ilustradores en el caso de recursos visuales.

DC.Subject (Materias): Temática del recurso. Típicamente, Subject expresará las palabras clave o frases que describen el título o el contenido del recurso. Se fomentará el uso de vocabularios controlados y de sistemas de clasificación formales.

DC.Description (Descripción): Descripción textual del recurso, tal como un resumen en el caso de un documento o una descripción del contenido en el caso de un documento visual.

DC.Publisher (Editor): Entidad responsable de hacer que el recurso se encuentre disponible en la red en su formato actual, por ejemplo la empresa editora, un departamento universitario u otro tipo de organización.

DC.Contributor (Otros colaboradores): Persona u organización que haya tenido una contribución intelectual significativa en la creación del recurso pero cuyas contribuciones son secundarias en comparación a las de las personas u organizaciones especificadas en el elemento Creator (por ejemplo, editor, ilustrador y traductor)

DC.Date (Fecha): Fecha en la que el recurso se puso a disposición del usuario en su forma actual. Esta fecha no ha de confundirse con la que pertenece al elemento Coverage, que sería asociada con el recurso sólo en la medida en que el contenido intelectual está de algún modo relacionado con esa fecha.

DC.Type (Tipo de recurso): Categoría del recurso, por ejemplo página personal, poema, diccionario. Para asegurar la interoperabilidad, Type debería ser seleccionado de entre una lista de valores que actualmente se encuentra bajo desarrollo en un grupo de trabajo.

DC.Format (Formato): Formato de datos de un recurso, usado para identificar el software y posiblemente, el hardware que se necesitaría para mostrar el recurso. Para asegurar la interoperabilidad, los valores de Format deberían ser seleccionados de entre una lista de valores que actualmente se encuentran bajo desarrollo en un grupo de trabajo.

DC.Identifier (Identificador del recurso): Secuencia de caracteres usados para identificar unívocamente un recurso. Ejemplos para recursos en línea pueden ser URLs y URNs (cuando están implementados). Para otros recursos pueden ser usados otros formatos de identificadores, como por ejemplo el ISBN.

DC.Source (Fuente): Secuencia de caracteres utilizado para identificar unívocamente un trabajo a partir del cual proviene el recurso actual. Por ejemplo, es posible usar Source con la fecha de 1603 como descripción de una película basada en una obra de Shakespeare, pero es preferible, en ese caso, usar Relation “IsBasedOn” con una referencia a un recurso distinto cuya descripción contenga el elemento Date con valor 1603.

DC.Language (Idioma): Idioma/s del contenido intelectual del recurso.

DC.Relation (Relación): Identificador de un segundo recurso y su relación con el recurso actual. Este elemento permite enlazar los recursos relacionados y las descripciones de los recursos.

- IsVersionOf: Incluye la edición de un trabajo
- IsBasedOn: Traducción de un trabajo
- IsPartOf: Capítulo de un libro
- IsFormatOf: Mecanismo de transformación de un conjunto de datos en una imagen

Para asegurar la interoperabilidad, las relaciones deberían ser seleccionadas de una lista de elementos que actualmente se encuentra bajo desarrollo en un grupo de trabajo.

DC.Coverage (Cobertura): Característica de cobertura espacial y/o temporal del contenido intelectual del recurso.

La cobertura espacial se refiere a una región física (por ejemplo, sector celestial); uso de coordenadas (por ejemplo, longitud y latitud) o nombres de lugares extraídos de una lista controlada.

La cobertura temporal se refiere al contenido del recurso en vez de a cuando fue creado o puesto accesible ya que este último pertenece al elemento Date.

DC.Rights (Derechos): Una referencia (URL, por ejemplo) para una nota sobre derechos de autor, para un servicio que dará información sobre términos y condiciones de acceso a un recurso. Una especificación formal del elemento Rights se encuentra actualmente en discusión y por lo tanto su uso se considera experimental.

10.4 PROYECTOS DE CATALOGACIÓN DE RECURSOS ELECTRÓNICOS

10.4.1 EL PROYECTO CORC

Uno de los servicios de valor añadido que proporciona cualquier biblioteca del mundo es un catálogo para ayudar a los usuarios a descubrir y tener acceso a los recursos disponibles en la biblioteca.

Durante siglos esto ha implicado un catálogo físico (de alguna forma, la ficha de catálogo ha sido el tipo de catálogo más ampliamente usado durante el último siglo) que contiene información sobre objetos físicos alojados en la biblioteca. Y la biblioteca ha servido tradicionalmente a sus usuarios como el punto de acceso más accesible a la información sobre información (metadatos) además de ser un punto de entrega de recursos informativos, ya sean entregados instantáneamente o tras un periodo de tiempo, por estar alojados en otro lugar, a través del préstamo interbibliotecario u otros medios.

Los rápidos cambios en la tecnología experimentados en la última parte del siglo XX han redefinido las fronteras en lo que respecta a la adquisición, el almacenamiento, la descripción y los puntos de acceso en la cadena de entrega de información. Para las bibliotecas, el World Wide Web representa una gran oportunidad de expandir la amplitud y el alcance de sus colecciones y servicios. Los usuarios se encuentran cada vez más cómodos buscando y usando información del Web.

Por lo tanto, no sorprende que el avance de la era digital haya llevado a la digitalización de algunas bibliotecas y al diseño de proyectos sobre selección y acceso a recursos vía Web. A pesar de que algunos productos de estos esfuerzos han sido y continuarán siendo útiles, no existía un catálogo o registro estandarizado, global y cooperativo, de recursos Web seleccionado por las bibliotecas. Y si existía alguno, la mayor parte de metadatos empleados para describirlos, no siempre se habían creado de forma que pudiesen ser ampliamente compartidos.

El hecho de proporcionar mejores métodos para localizar y organizar recursos disponibles en el Web no había sido una prioridad para los bibliotecarios. Se han producido

mayores esfuerzos por parte de algunas comunidades y agencias para ayudar al desarrollo de estándares y tecnologías emergentes sobre metadatos. Hay iniciativas prometedoras como, por ejemplo, Dublin Core y el desarrollo de estándares como el HTML, RDF, XML, por parte del consorcio World Wide Web (W3C).

Teniendo en cuenta todo lo anterior, OCLC se ha asociado a un grupo de bibliotecas para iniciar un nuevo acercamiento en la construcción de un catálogo de recursos Web seleccionados por bibliotecas, el proyecto CORC es la respuesta a la necesidad de las bibliotecas de llevar sus conocimientos y servicios hacia la cooperación. OCLC advirtió que era un momento apropiado de tratar el universo de recursos electrónicos en su medio natural, el World Wide Web.

Desarrollo del proyecto CORC

OCLC ya había investigado y había trabajado en el desarrollo de prometedoras tecnologías que podrían ser aplicadas para apoyar a las bibliotecas en su esfuerzo por organizar el acceso a los recursos de Internet que estaban siendo seleccionados por estas bibliotecas. Una vez que se decidió optar por un modelo de catálogo cooperativo para desarrollar CORC, OCLC desarrolló un borrador para identificar e incorporar los objetivos que deberían conseguirse para que el proyecto fuese un éxito. Esto, principalmente, incluía: la catalogación cooperativa de recursos web, uso de metadatos tanto de uso local como compartidos, la obtención de un catálogo de metadatos para recursos físicos y digitales, el control de autoridades para los puntos de acceso, la importación y exportación RDF/XML, la integración de Dublin Core y MARC en un único sistema, la recuperación flexible de recursos, la posible utilización del sistema Unicode, asignación asistida de los encabezamientos de materia y los epígrafes clasificatorios, la extracción automática de palabras clave y el mantenimiento de los enlaces.

El uso de Dublin Core era un hecho, así como el uso de Mantis, un paquete de herramientas para construir catálogos en entornos web. Para los recursos web, Dublin Core ha adquirido un interés general, y Mantis es un sistema excelente sobre el cual trabajar.

Con el objetivo de desarrollar un servicio que resultase útil para las bibliotecas en la descripción y el acceso a los recursos de Internet, en agosto de 1998, se presentó el proyecto CORC ante el Research Advisory Committee, un grupo de asesores externo.

El primer borrador se desarrolló y se desarrolló un prototipo en la OCLC Office of Research, invitando a participar a las bibliotecas haciendo uso del sistema que se acababa de redefinir. Aunque el sistema Mantis estaba funcionando, la mayor parte de las prestaciones que se pretendían proporcionar, en particular el deseo de proporcionar un sistema que trabajase tanto con Dublin Core como con la catalogación en MARC, hacían de CORC un inusual proyecto de investigación en varios sentidos. Se comenzó el desarrollo de lo que podría ser uno de los sistemas más importantes para OCLC. Pero era bastante claro que la Office Research no conseguiría llevar a cabo el proyecto por si sola, tanto por las necesidades de mano de obra como por la experiencia y el consejo necesario del resto de OCLC. Estas necesidades desembocaron en la asignación de personal al proyecto desde los departamentos de Marketing, Desarrollo, Calidad y Documentación.

El desarrollo del proyecto comenzó en 1998, con la intención de tener usuarios reales en enero de 1999. La Office Research podía trabajar durante un año, añadiendo nuevas funciones al sistema o refinándolo conforme lo considerase oportuno. Durante este periodo de tiempo, OCLC podía también establecer planes con el objetivo de considerar CORC como un servicio propio. El funcionamiento de CORC se hizo efectivo el 15 de enero, aunque en un primer momento el acceso estaba restringido a las bibliotecas participantes en el proyecto.

Más allá de las características del sistema, la participación de un gran número de bibliotecas ha sido un hecho decisivo para el éxito de CORC. Se ha sabido construir un sistema de metadatos. Se ha construido un sistema y se disponen de las tecnologías que pueden ser implementadas en diferentes sistemas. El truco ha sido la construcción de un sistema que satisface las necesidades de las bibliotecas, cosa imposible de llevar a cabo sin la participación activa de bibliotecas que aplican el sistema a tareas reales.

Además, se estableció contacto, a través de la Universidad de Michigan, con el Committee on Institutional Cooperation (CIC). Varios miembros de CORC asistieron al

congreso del CIC sobre metadatos, celebrada en 1998. La asistencia a este congreso brindó la posibilidad de conocer a personal del CIC y a otros pioneros en la aplicación de metadatos en el marco bibliotecario. Las bibliotecas del CIC decidieron participar también en CORC.

Desde la llamada inicial a la participación a finales de 1998, CORC ha suscitado un alto interés por parte de la comunidad bibliotecaria. Hasta la fecha, CORC ha atraído a una amplia variedad de instituciones interesadas en participar en el proyecto.. Las grandes bibliotecas académicas forman la lista más amplia de participantes, pero también participan bibliotecas públicas, pequeñas bibliotecas académicas, un museo, varias bibliotecas especializadas de la agencia federal de Estados Unidos, una biblioteca especializada en medicina, etc. Fuera de Estados Unidos, participan en CORC bibliotecas de Canadá, Europa, Oriente Medio, Australia y de algunos lugares de Asia.

10.4.2 BERKELEY DIGITAL LIBRARY

Página principal de la Berkeley Digital Library 1

La Berkeley Digital Library SunSITE será una biblioteca en el sentido clásico, no solo una colección, sino también unos servicios con valor añadido proporcionados por personal cualificado. Los materiales serán seleccionados por su contribución al propósito de la colección y su posible utilidad para los usuarios. La colección será organizada y dispuesta para el fácil acceso por parte de aquellos que la necesiten, y se proporcionará asistencia cuando esta sea requerida.

Pero la biblioteca digital de Berkeley también va más allá del papel de una biblioteca tradicional, para funcionar como catalizador para el desarrollo de bibliotecas digitales en todo el mundo. Allí se puede encontrar información sobre algunas de las últimas tecnologías para las bibliotecas digitales (como java), o descubrir qué discusiones electrónicas tratan estos temas o que publicaciones se deben seguir. Es un buen lugar donde encontrar información sobre copyright y la propiedad intelectual, las imágenes digitales o la preservación de materiales. Además se puede encontrar información sobre SGML o Z.39.50.

La biblioteca digital de Berkeley está utilizando cinco estrategias para apoyar el desarrollo de bibliotecas digitales, museos y archivos:

- Reunir y publicar información sobre proyectos de bibliotecas digitales
- Reunir y proporcionar acceso a contenidos digitales
- Proporcionar una plataforma para la investigación y el desarrollo digital.
- Promover debates sobre temas relacionados a las bibliotecas digitales, museos y archivos
- Proporcionar servicios actualizados.

Socios

La asociación estratégica con otros actores en el mismo campo es una clave estratégica que se está empleando para conseguir una colección y unos servicios mejores que los que se podrían crear estando solos. Estas asociaciones actualmente incluyen a bibliotecas públicas, asociaciones profesionales, universidades y compañías comerciales. Algunos ejemplos de estas asociaciones son:

Sun Microsystems, Inc.: Sun es, obviamente, un socio en la biblioteca digital SunSITE, pero son vistos más como un socio estratégico que les ayuda a expandir y mejorar la biblioteca digital a través de iniciativas tecnológicas.

Electronic Book Technologies (EBT): Los proyectos en SGML han sido apoyados por generosas donaciones de software por parte de Electronic Book Technologies, una empresa que desarrolla software con SGML. La biblioteca digital de Berkeley actualmente utiliza su software DinaWeb para convertir ficheros codificados con SGML a HTML.

University of California Press: La University of California Press, la biblioteca de Berkeley y la Library Automation están cooperando en un proyecto de la Mellon Foundation para colocar en Internet un conjunto de revistas de humanidades usando el Web y SGML. El proyecto también está investigando temas de política electrónica tales como modelos de coste en la recuperación.

InFoPeople Project: La biblioteca del estado de California creó un proyecto para colocar ordenadores en las bibliotecas públicas de todo el estado y formar al personal bibliotecario en el uso de Internet tanto para acceder a la información como para publicarla. La biblioteca de Berkeley ha colaborado con el proyecto, viendo que este proyecto era una parte importante de su estrategia para hacer que las colecciones de su biblioteca digital fuesen fácil y ampliamente accesibles para los ciudadanos del estado.

Association of Research Libraries: La biblioteca digital de Berkeley SunSITE estableció una relación de trabajo con la ARL a través del etiquetado de publicaciones de ARL relacionadas con las bibliotecas digitales y situándolas en la SunSITE.

Desarrollo de estándares

Una parte importante en el desarrollo de las aplicaciones en una biblioteca digital es un compromiso de crear y mejorar los estándares. La biblioteca digital de Berkeley SunSITE enlaza a recursos sobre estándares existentes o sus borradores, ya sean estándares de facto o no. Actualmente, la biblioteca digital SunSITE es una fuente autorizada para el borrador del estándar ANSI/NISO del Z39.56, tanto en formato HTML como Adobe Acrobat.

Programa de bibliotecarios digitales

Las bibliotecas digitales no pueden ser construidas, gestionadas y mantenidas sin bibliotecarios digitales. Desafortunadamente, no hay un gran número de gente a los que se les pueda llamar “bibliotecarios digitales”. Parte de la misión de SunSITE es buscar a quienes puedan ser llamados bibliotecarios digitales y ayudarles a que abran un camino que puedan seguir todos los demás. Se buscan bibliotecarios que hayan mostrado habilidades técnicas e imaginación para utilizar las tecnologías digitales para los servicios de las bibliotecas tradicionales. A estas personas se les ofrece una cuenta en la SunSITE y un nivel básico de apoyo en la infraestructura.

Apoyo a los proyectos de bibliotecas digitales

La biblioteca digital SunSITE busca activamente investigaciones sobre bibliotecas digitales y desarrollos de proyectos para albergarlos en SunSITE.

Recursos y colecciones de SunSITE

<http://sunsite.berkeley.edu/ARL/>

La biblioteca digital SunSITE alberga una colección de documentos de la ARL relacionados con las bibliotecas digitales.

Current Cites

Más de cinco años de bibliografía comentada sobre artículos seleccionados, libros y documentos electrónicos sobre tecnologías de la información. Más de dos años de la colección están en HTML, por ello es posible enlazar con el recurso que está siendo citado.

Finding Aids for Archival Collections

Acceder al contenido de colecciones archivísticas puede ser difícil a no ser que se viaje al lugar donde esté albergado un archivo en el que se esté interesado. Un proyecto de la UC Berkeley, pero que incluye otras instituciones, ha codificado cientos de finding aids de archivos en SGML y los ha hecho accesibles en la biblioteca digital SunSITE. Algunos de los finding aids también proporcionan acceso a representaciones digitalizadas de los propios documentos, tales como fotografías históricas.

Libweb: Library Servers via WWW

Se trata de una lista de servidores de bibliotecas de todo el mundo, mantenida por Tomas Dowling, uno de los bibliotecarios digitales de SunSTIE

Nineteenth-Century Literature

Una publicación trimestral de la University of California Press. Actualmente los números anteriores a 1990 están disponibles a texto completo.

The Online Medieval and Classical Library

Una colección de textos medievales y clásicos están disponibles, desde Chaucer y Homer hasta las sagas clásicas de autores anónimos.

Los fondos futuros

La única constante es el cambio. Lo que es cierto hoy podría ser falso mañana. El reto es trabajar unidos y apoyar y gestionar el proceso del cambio. Más que construir monumentos estáticos de ideas y tecnologías fallidas, se debe aceptar el reto de preparar un futuro que no se puede predecir. La biblioteca digital de Berkeley se esfuerza por anticiparse a los cambios que transformarán a las bibliotecas, museos y archivos de todo el mundo.

11 RDF

Resource Description Framework (RDF) es una infraestructura que permite el codificado, intercambio y la reutilización de metadatos estructurados. RDF permite la utilización de vocabularios comprensibles y utilizables tanto por máquinas como por humanos y está diseñado para la reutilización de los metadatos entre distintas comunidades de información. RDF trata de ser un marco para la codificación y el intercambio consistente de paquetes de metadatos, con las definiciones y significados de los metadatos procedentes de las comunidades de información que los han desarrollado.

11.1 ANTECEDENTES

RDF es un producto del World Wide Web Consortium (W3C). Antes del desarrollo de RDF, Internet Engineering Task Force (IETF) había desarrollado un mecanismo conocido como PICS (Platform for Internet Content Selection). PICS era un mecanismo para comunicar las valoraciones de las páginas web que iban desde los servidores a los clientes. Estas valoraciones, o sistemas de etiquetado, fueron definidas por varias comunidades que querían hacer uso de ellas. PICS estaba enfocado a ser un mecanismo general para crear sistemas de valoración. Distintas organizaciones podían evaluar los contenidos en base a sus propias metas y objetivos. A pesar de pretender ser un mecanismo general, las primeras motivaciones para el desarrollo de PICS estaban ligadas a la presencia de pornografía en la red y el deseo de la comunidad de Internet de desarrollar un sistema propio de regulación para prevenir el esfuerzo de los gobiernos para imponer restricciones sobre contenidos.

Se identificaron algunas limitaciones en la arquitectura de PICS, y se desarrollaron nuevas funcionalidades necesarias para asociar los metadatos descriptivos con los recursos a los cuales describían. Fue entonces cuando se vio que la infraestructura definida para PICS podía ser útil en otros dominios. El W3C combinó los trabajos que se estaban realizando hasta entonces en el campo de los metadatos y los llevó al grupo de trabajo de RDF.

RDF es el resultado de la reunión de una serie de comunidades que trabajaban sobre metadatos, sus necesidades y requerimientos para proporcionar una robusta arquitectura que soportara metadatos en una amplia gama de aplicaciones. A pesar de la gran influencia de PICS, y siendo un marco para la utilización de metadatos y un mecanismo de representación de conocimiento, RDF no es producto de ninguna influencia concreta sino que es el resultado de las contribuciones sobre los estudios sobre metadatos que se habían venido realizando. También ha sido influenciado por Dublín Core y Warwick Framework

11.2 MODELO DE DATOS DE RDF

RDF proporciona un modelo para definir recursos. Los recursos poseen propiedades. En RDF, un recurso es un objeto que puede ser identificado unívocamente por un Uniform Resource Identifier (URI). Las propiedades asociadas con los recursos están identificadas por tipos de propiedades, y los tipos de propiedades tienen unos valores. Los tipos de propiedades expresan las relaciones de los valores asociados con los recursos. En RDF, los valores deben ser atómicos (ya sean cadenas de caracteres o valores numéricos) u otros recursos que, a su vez, deben tener sus propias propiedades. El conjunto de propiedades que se refieren a un recursos se llama descripción. El modelo RDF proporciona un mecanismo de sintaxis independiente para representar los recursos y sus descripciones asociadas. La sintaxis de RDF, expresada en XML, proporciona un mecanismo para codificar este modelo.

Por ejemplo, imaginemos que tenemos un documento que representa el artículo de NEEDLEMAN, M.H⁽¹⁵⁾ y que es enlazado por el URI ficticio <http://www.dra.com/SR/rdfarticle.com> . Este recurso tiene algunas propiedades, entre ellas la propiedad “autor”. La propiedad autor tiene el valor “Mark Needleman” (si hubiese más de un autor, habría ocurrencias múltiples de la propiedad autor). El recurso tiene otra propiedad llamada, “título”, que tiene el valor “RDF:Te Resource Description Framework”. Se podría expandirla descripción mediante metadatos del recurso asociando otras propiedades, tales como el idioma (que tendría el valor de “inglés”; habría múltiples instancias de la propiedad idioma si el artículo estuviese disponible en otros idiomas), formato (que debería tener el valor de “MS Word” o “HTML”), y longitud (que debería

describir el recurso en términos de número de bytes, número de palabras, etc.). Cada uno puede definir estos recursos, con un vocabulario particular, o hacer uso de definiciones creadas por otras comunidades de metadatos.

En el ejemplo que acabamos de comentar, todas las propiedades que han sido definidas poseen valores simples (atómicos) consistentes en cadenas de caracteres. Pero el valor de autor, en lugar de ser la cadena “Mark Needleman” podría haber sido el URI que llevase a este recurso

11.3 LA SINTAXIS DE RDF

El modelo de datos RDF proporciona un sistema de sintaxis independiente para describir los recursos (usando el concepto de RDF de propiedades con valores asociadas con recursos). Pero para que los metadatos puedan intercambiarse, sin embargo, deben expresarse con una sintaxis concreta. Para RDF, la sintaxis es XML. XML es un producto del W3C, se trata de un lenguaje flexible para definir conjuntos de etiquetas y se está convirtiendo en una de las tecnologías más ampliamente utilizadas en Internet. Era obvio utilizar este lenguaje como sintaxis para RDF. RDF impone algunas estructuras formales en el uso de XML para permitir la representación consistente de la semántica. Una de las cosas más importantes que hace RDF es el uso de Namespaces de XML. Los Namespaces son una forma de asociar una autoridad con una etiqueta y, por tanto, eliminar la ambigüedad de esa etiqueta. En el ejemplo que hemos comentado, ¿cual es el significado de la propiedad “Autor”? ¿es el autor tal y como se define en las AACR2 y MARC21, o es un significado particular que le hemos dado a la propiedad? Un usuario (y una máquina) no tiene forma de saberlo y, en el caso de reconocer la sintaxis de la propiedad, debería intentar asignar algún significado a esa propiedad basándose en su propio conocimiento, lo cual no se correspondería con lo que nosotros queremos. Con la utilización de los Namespaces de XML, se puede asociar una autoridad con una propiedad y distinguir entre propiedades que puedan tener una sintaxis similar pero significados distintos.

El conjunto de metadatos de Dublín Core (DC), define un conjunto de metadatos para la descripción de recursos.

15. CONCLUSIONES

En la actualidad es posiblemente un tópico pretender organizar y ordenar todo el universo Internet de manera uniforme: la inmensa cantidad de materiales existentes en la red, su crecimiento en progresión geométrica, y la diversidad de procedencia y, sobretodo, la diversidad de intereses de los organismos e individuos que publican, dificultaran la adopción de un sistema común de metadatos. No obstante, es de suponer, que al menos por lo que respecta a los recursos que puedan interesar a comunidades concretas de usuarios, sea más fácil llegar a acuerdos que contribuyan a controlar bibliográficamente determinados espacios Internet. En este sentido, existen muchas posibilidades abiertas para que se pueda mejorar el control y acceso de los documentos Internet, ya será desde la misma red a partir de motores de los que dispone o de la elaboración de directorios y guías de recursos, ya sea con la integración de la descripción de estos recursos en catálogos de bibliotecas. Los profesionales de la información son los que tienen que decidir qué método o métodos son los más adecuados para satisfacer las necesidades de información de los usuarios.

BIBLIOGRAFÍA ISO 690 E ISO 690-2 (RECURSOS ELECTRÓNICOS)

1. BARÓ I QUERALT, J.; MARTÍN, F. Los motores de búsqueda en Internet: características básicas. [En línea]. [Citado en enero 2002]. Disponible en: <http://camelot.upf.es/digital/biblioteca.htm>
2. BORREGO HUERTA, A. Bibliotecarios reales para bibliotecas virtuales: la necesidad de asistencia humana en el entorno digital. *Métodos de Información*, 2001, vol. 8, nº 48, pp. 47-55
3. ESTIVILL RIUS, A. Organització dels recursos Internet. *Item*, 1996, nº 18, pp. 42-74
4. GARCÍA CAMARERO, E.; GARCÍA MELERO, L.A. *La Biblioteca Digital*. Madrid: Arco/Libros, 2001
5. GIMENO MONTOSO, M.J.; BARRUECO CRUZ, J.M.; GARCÍA TESTAL. Catalogación de recursos electrónicos accesibles en Internet: revisión de propuestas para una normativa. *Fesabid 98*, VI Jornadas Españolas de Documentación, 1998
6. HICKEY, TB; CHILDRESS, E; WATSON, BC. The genesis and development of CORC as an OCLC Office of Research project. Mayo-junio 1999 [Citado en junio de 2002]. Disponible en: <http://www.oclc.org/oclc/new/n239/feature/02feature.htm>
7. HÍPOLA, P.; VARGAS-QUESADA, B.; SENSO, J.A. Bibliotecas digitales: situación actual y problemas. *El profesional de la información*, 2000, vol 9, nº4, pp. 4-13
8. KUANG-HWEI (JANET) LEE-SMELTZER. Finding the needle: controlled vocabularies, resource discovery, and Dublin Core. *Library Collections, Acquisitions & Technical Services*, 2000, vol. 24, pp. 205-215
9. KUNY, T; CLEVELAND, G. La biblioteca digital: mitos y retos. *Métodos de Información*, 2001, vol. 8, nº 48, pp. 72-82

10. Libro verde sobre derechos de autor y derechos afines en la sociedad de la información. COM, 1995
11. Libro verde sobre derechos de autor y derechos afines en la Sociedad de la Información. COM, 1996
12. LLORET, N. Cómo plantear un proyecto para el desarrollo de una biblioteca digital. JADOC'99, 2as Jornadas andaluzas de Documentación, 1999, pp. 307-317
13. MALDONADO MARTÍNEZ, A; FERNÁNDEZ SÁNCHEZ, E. Evaluación de los principales "buscadores" desde un punto de vista documental: recogida, análisis y recuperación de recursos de información. Fesabid 98, VI Jornadas Españolas de Documentación, 1998
14. MALDONADO MARTÍNEZ, A; FERNÁNDEZ SÁNCHEZ, E. Análisis comparativo de buscadores en Internet. El profesional de la información, 2000, vol. 9, nº 3, pp. 40-46
15. NEEDLEMAN, M.H. RDF: The Resource Description Framework. Serials Review, 2001, vol. 27, nº 1, pp. 58-61
16. OLVERA LOBO, M.D. Métodos y técnicas para la indización y recuperación de los recursos de la World Wide Web. Boletín de la asociación Andaluza de Bibliotecarios, 1999, nº 57, pp. 11-22
17. ORTIZ HERRERA, S. La obra digital y el derecho de reproducción especial referencia a la copia privada. Fesabid 98, VI Jornadas Españolas de Documentación, 1998
18. PAZ YANES, C. Bibliotecas digitales: algo viejo, algo nuevo y algo prestado. Métodos de Información, 2001, vol. 8, nº 48, pp. 56-64
19. POUCHARD, L. Cataloging for digital libraries: the TEI scheme and the TEI header. 1998 [citado en junio de 2002]. Disponible en: <http://alexia.lis.uiuc.edu/review/6/pouchard.html>

20. RIBAS, X. Usos permitidos en Internet. [En línea]. [Citado en marzo de 2002]. Disponible en: <http://www.onnet.es/01005006.htm>
21. RIBAS, X.; RODRIGUEZ. La copia privada en Internet. [En línea]. [Citado en marzo de 2002]. Disponible en: <http://www.onnet.es/01005007.htm>
22. TALAVERA IBARRA, A.M.. Una biblioteca virtual: nace o se hace. VIII Conferencia Internacional de Bibliotecología “Biblioteca Digital: Perspectivas, Experiencias y Oportunidades”, Santiago de Chile, 24-26 octubre 2001
23. TENNANT, R. The Berkeley digital library SunSITE. Febrero 1996 [Citado en junio de 2002]. Disponible en: <http://www.dlib.org/dlib/february96/ucb/02tennant.html>
24. Tratados adoptados en la conferencia diplomática sobre determinadas cuestiones relativas a los derechos de autor y derechos afines, celebrada el 20 de diciembre de 1996 en Ginebra bajo los auspicios de la Organización Mundial de la Propiedad Intelectual (OMPI)
25. WEIBEL, S. The Dublin Core: A simple content description model for electronic resources. [En línea]. Octubre-noviembre 1997 [citado en junio de 2002]. Disponible en: <http://www.asis.org/Bulletin/Oct-97/weibel.htm>
26. <http://sunsite.berkeley.edu> [Citando en enero de 2002]
27. <http://exlibris.usal.es/merlo/escritos/redcbd1.htm> [Citado en febrero de 2002]
28. <http://www.geocities.com/zaguan2000/electronic> [Citado en mayo de 2002]