

**Universidad Central de Venezuela
Facultad de Ciencias
Escuela de Computación**

Lecturas en Ciencias de la Computación
ISSN 1316-6239

Notas sobre Bibliotecas Digitales

Jaime Blanco.

RT 2007-07

Centro IOMMA
Caracas, Septiembre, 2007.

**Universidad Central de Venezuela
Facultad de Ciencias
Escuela de Computación**

Lecturas en Ciencias de la Computación
ISSN 1316-6239

Notas sobre Bibliotecas digitales

Jaime Blanco
RT 2007-07

Centro IOMMA
Caracas, Septiembre, 2007

Notas sobre Bibliotecas digitales

Jaime Blanco

jblanco@ciens.ucv.ve

Universidad Central de Venezuela. Facultad de Ciencias. Escuela de
Computación.
Centro de Investigación de Operaciones y Modelos Matemáticos (CIOMMA)
Venezuela. Caracas
Apdo. 47002, 1041-A.

RT 2007-07

ABSTRACT

En este estudio de las bibliotecas digitales tratamos de encontrar una definición, así como mostrar la naturaleza y gestión de las bibliotecas digitales, en la transformación y cambio de las bibliotecas tradicionales al nuevo concepto de biblioteca digital. Con la ayuda del Computador es más sencillo y rápido encontrar la información, estos abren una amplia gama de posibilidades y diferentes áreas de investigación. Con la llegada de Internet las bibliotecas comenzaron a crear sus sitios Web, incluyendo información sobre su servicios, horarios de atención y acceso en línea a sus catálogos públicos; pero llegar a la Biblioteca Digital implica un trabajo de digitalización de la Información, en la cual hace falta definir estándares para el intercambio de información.

Keywords: Bibliotecas Digitales.

Septiembre, 2007

TABLA DE CONTENIDO

Abstract.....	2
1. <i>Definiciones De Bibliotecas Digitales</i>	4
2. <i>El camino de la biblioteca tradicional a la biblioteca digital</i>	7
3. <i>Campos de acción relacionados con las BIBLIOTECAS DIGITALES</i>	11
4. <i>Las bibliotecas digitales hoy y proyectos en curso</i>	14
5. <i>Estándares en la Industria de Bibliotecas</i>	18
6. <i>Conclusiones</i>	27
7. <i>Bibliografía</i>	28

Bibliotecas digitales

1. DEFINICIONES DE BIBLIOTECAS DIGITALES

En el estudio de las bibliotecas digitales convergen muchas áreas y nichos diferentes de investigación con el objetivo de resolver algunas de las dificultades involucradas en la definición, naturaleza y gestión de las bibliotecas digitales, en la transformación y cambio de las bibliotecas tradicionales al nuevo concepto de biblioteca digital.

Es por esto mismo que las posibles definiciones de biblioteca digital varían dependiendo del alcance y apertura del concepto y de la problemática que se desee solucionar a través de ellas.

Nuestra definición es el resultado del material consultado y del estado actual y futuro de las bibliotecas digitales.

Para mayor cobertura se incluyen algunas definiciones obtenidas y recolectadas en [Fox01] que a su vez reúne definiciones de fuentes diferentes:

Las nuevas bibliotecas digitales tendrán características no posibles en las bibliotecas tradicionales, extendiendo el concepto de biblioteca más allá de sus límites físicos, tendrán recursos y servicios innovadores
(2001 PITAC Report, "Digital Libraries: Universal Access to Human Knowledge", p. 3)

Las bibliotecas digitales son organizaciones que proporcionan recursos, para seleccionar, estructurar, interpretar, distribuir, y garantizar a través del tiempo la integridad y persistencia de material digital para hacerlo asequible, legible y económicamente disponible y utilizable por una comunidad.
(Digital Library Federation)

Las bibliotecas digitales son sistemas de información complejos conformados por datos, información y conocimiento, para ayudar a satisfacer las necesidades de usuarios y sociedades, proporcionando servicios de información (escenarios), una organización útil de la información (estructuras), capacidad para administrar la ubicación y localización de la información (espacios) y comunicar la información a los usuarios y agentes.
(Edward A. Fox, July 1999, according to 5S Framework)

Una biblioteca digital sucede en el contexto del diseño de un espacio complejo en cuatro dimensiones: comunidad, tecnología, servicios y contenido.
(Gary Marchionini and Edward A. Fox, "Progress toward digital libraries: augmentation through integration", pp. 219-225)

El campo de las bibliotecas digitales, se refiere a la aplicación de la tecnología digital a los problemas de la información, diseccionados por instituciones como bibliotecas, archivos, museos, colegios, publicadores, y agencias de información. El trabajo en bibliotecas digitales se refiere a integrar servicios para servir mejor las necesidades humanas, a través de un tratamiento holístico, de la interfase, ubicación, tiempo, lenguaje y sistema.

(Gary Marchionini and Edward A. Fox, "Progress toward digital libraries: augmentation through integration", Information Processing & Management, 35(3):219-225, May 1999.)

Una biblioteca digital es una colección de servicios y objetos de información, que ayuda a los usuarios en el manejo, organización, presentación y disponibilidad directa o indirecta de objetos de información por vía digital.

(The Scope of the Digital Library, Draft preparado por Barry M. Leiner para el DLib Working Group on Digital Library Metrics, 1998)

Una biblioteca digital es una base de datos organizada de objetos digitales de información en formatos variados, mantenidos para proporcionar un acceso fácil orientado a una comunidad, con las siguientes características adicionales: herramienta genérica de acceso (catálogo) con capacidad de búsqueda y retorno de resultados de todo el universo de la base de datos y con procedimientos organizados técnicamente para la administración de los objetos en la base de datos (ingreso, retiro).

(Peter Graham, Rutgers University Libraries, 1997)

Las bibliotecas digitales son un conjunto de recursos electrónicos y capacidades técnicas para crear, buscar, y usar información. Son una extensión y mejora a los sistemas de almacenamiento y recuperación de información, que manipulan datos digitales en cualquier formato (texto, imágenes, sonido) y que existen en redes distribuidas. El contenido de las bibliotecas digitales incluye datos, metadatos (datos que describen aspectos de los datos y relaciones a otros datos o metadatos).

(1996 UCLA-NSF Social Aspects of Digital Libraries Workshop)

Las bibliotecas digitales son construídas, coleccionadas y organizadas por una comunidad de usuarios, y su funcionalidad soporta las necesidades de información y usos de esta comunidad. Son un componente de las comunidades mismas, en las cuales los individuos y grupos interactúan unos con otros, usando recursos y sistemas de datos, información y conocimiento. En este sentido son una extensión de una variedad de instituciones de información donde los recursos se seleccionan, coleccionan, organizan, preservan y se ponen al acceso de una comunidad de usuarios.

(1996 UCLA-NSF Social Aspects of Digital Libraries Workshop)

Las bibliotecas digitales son sistemas que proporcionan a una comunidad de usuarios un acceso coherente a un vasto repositorio organizado de información y conocimiento.

(Clifford Lynch, 1995)

Las bibliotecas digitales son sistemas que proporcionan a una comunidad de usuarios un acceso coherente a un vasto repositorio organizado de información y conocimiento. Esta organización de información se caracteriza por la inexistencia de conocimiento previo requerido para los usuarios de dicha información. La habilidad del usuario de acceder, reorganizar y utilizar el repositorio, se enriquece por la tecnología digital disponible.

Interoperability, Scaling, and the Digital Libraries Research Agenda, report of the 1995 IITA DL Workshop)

Una biblioteca que ha sido extendida y mejorada por la aplicación de tecnología digital. Las extensiones se refieren a: organización y administración de las colecciones, acceso a los ítems de la biblioteca y a la información contenida en ella y a la capacidad de comunicación de información sobre cada ítem.

(Terry Smith, UCSB, 1995).

Biblioteca digital es un término genérico para definir estructuras federadas que proporcionan a los humanos acceso intelectual y físico a una cantidad inmensa de redes de información, codificada en formatos digitales multimedia.

(The University of Michigan Digital Library: This Is Not Your Father's Library, Bill Birmingham, 1994)

Una biblioteca digital es un ambiente tecnológico distribuido que reduce de manera dramática las barreras de la creación, disseminación, manipulación, almacenamiento, integración y reuso de la información.

(Edward A. Fox, editor, Source Book on Digital Libraries, 1993, pg. 65)

Otras definiciones

Un nuevo tipo de bibliotecas que permiten crear servicios novedosos para sus clientes, un ambiente para descubrir información, construir conocimiento, soportar colecciones y documentos dinámicos, apoyar la colaboración y el aprendizaje de individuos y grupos. Proporcionar una arquitectura que soporta una apropiada organización de recursos con varios formatos, y en diferentes ubicaciones, demostrando un uso efectivo de los recursos [Rob02]

Sun Microsystems define en [Sun02] una biblioteca digital como *la extensión de funciones electrónicas realizadas por los usuarios de manera típica, para acceder a una biblioteca tradicional. Estos recursos de información pueden ser traducidos en forma digital, almacenados en repositorios multimedia y hacerlos asequibles a través de servicios basados en web. La biblioteca digital debe evolucionar para satisfacer el nuevo paradigma educacional de educación a distancia y acceso remoto a la información.*

Nuestra definición

Una biblioteca digital, es una organización que utiliza la tecnología de la computación y telecomunicaciones con el fin de proveer servicios de información a una comunidad con intereses comunes. Su identidad está diseminada tanto en la biblioteca misma como en la comunidad de información y en las otras entidades que participan e interactúan con ella. El objetivo de la biblioteca digital debe sobrepasar el ofrecer información sobre su colección y sus ítems digitalizados para lograr apoyar los procesos de aprendizaje de sus usuarios, de otras entidades de la comunidad de información y de ella misma, así como la adquisición, organización, categorización, búsqueda, despliegue, diseminación y preservación de la información y conocimiento generado durante su ciclo de existencia y evolución.

2. EL CAMINO DE LA BIBLIOTECA TRADICIONAL A LA BIBLIOTECA DIGITAL

Las bibliotecas han evolucionado al igual que muchos modelos de negocio con la participación y en paralelo a la tecnología de la computación en búsqueda de la eficiencia, eficacia y logro de los objetivos específicos inherentes a una área que interactúa a la vez, con la búsqueda del conocimiento, las necesidades y evolución de instituciones académicas y comerciales y un universo de usuarios cada vez con nuevas necesidades.

Las bibliotecas tradicionales han existido para servir a una comunidad más o menos definida y delimitada con procesos básicos como la selección, organización, almacenamiento, diseminación, y preservación de la información normalmente dentro de unas limitaciones físicas de una institución.

Antes del advenimiento de la computación y telecomunicaciones, la localización física de la comunidad era muy importante para que la biblioteca pudiera prestar sus servicios a tiempo y de manera eficiente. Era obligatorio para la biblioteca seleccionar, organizar, almacenar, diseminar y preservar la información dentro de los límites físicos de la institución. Sin embargo, la noción de bibliotecas digitales está basada en el hecho de que ya la información y sus servicios no tienen que estar localizados ni delimitados en un espacio físico donde se encuentren los usuarios de la biblioteca

Los procesos, alcance y propósitos de la existencia de una biblioteca deben y pueden ser amplios, unos abarcando una función educacional y otras comerciales, otras para ofrecer sólo información (datos procesados y desplegados en forma útil) y otras para ofrecer más que información, conocimiento (información transformada en significado). Otras para disminuir costos de publicación y distribución, facilitar la preservación, crear alianzas para reutilizar esfuerzos y conocimiento adquirido por otras entidades, dinamizarse en un tiempo y lugar específicos. **[Rob02]**

El alcance y propósito de la biblioteca en su misión hacia la satisfacción de necesidades de su comunidad, mas los avances tecnológicos, son quienes abren una amplia gama de posibilidades y diferentes áreas de investigación y aplicaciones para que surjan bibliotecas digitales en sus amplias y variadas definiciones.

En los inicios de 1970 las bibliotecas tradicionales empezaron a adoptar aplicaciones de software para permitirles realizar su operativa y funciones de manera más eficiente. Este tipo de aplicaciones fueron en su mayoría desarrolladas por las mismas universidades o entidades de investigación directamente involucradas con las bibliotecas. Estos sistemas de información para bibliotecas se volvieron cada vez mas comunes en su desarrollo y utilización, se incrementaron sus desarrolladores y proveedores, sus beneficios fueron más evidentes, y adicionalmente el poder de cómputo aumentó y disminuyeron los costos. **[Sun03]**

La siguiente fase consistió en sistemas de biblioteca integrados, que permitía a los bibliotecarios realizar casi todas sus funciones en línea. (Sistemas bibliotecarios Integrados - Integrated Library Systems - ILS). Se aumentaron las posibilidades de consulta y utilización de los datos ingresados a la biblioteca, igual que la eficiencia y exactitud, sin embargo la mayoría de estas aplicaciones eran host-terminal, con hardware, sistemas operativos y aplicaciones propietarias que no permitían la interacción con otros sistemas o interoperabilidad con otras instituciones y usuarios.

Durante la transición de tecnologías de computación de gran escala a los microcomputadores a mediados de los 80's, se cambió el panorama y muchos de estos sistemas fueron reinventados, para permitir la reutilización de recursos y redes. Luego el advenimiento de Internet y de las posibilidades de proporcionar acceso a recursos locales o externos a través de los estándares de comunicación TCP/IP se permite la interconectividad de recursos de computación, servidores, y terminales.

Los 90's implicaron nuevas infraestructuras de comunicación para crear y almacenar información en las bibliotecas, se repiensen los diferentes modelos de "entrega de la información a los usuario finales". Las bibliotecas se convierten en partes vitales e integrales de la infraestructura tecnológica de las universidades que compiten de igual a igual con el presupuesto educativo y de las demás áreas de la institución. **[Sun03]**

Durante los 90s, **[Sun02]** los proyectos de bibliotecas digitales eran principalmente experimentales, y comenzaron algunos avances en las técnicas relacionadas con las bibliotecas digitales a través de la investigación patrocinada por la U.S. National Science Foundation (NSF) el U.K. Joint Information Systems Committee (JISC).

Las bibliotecas comenzaron a crear sus sitios Web, incluyendo información sobre su servicios, horarios de atención y acceso en línea a sus catálogos públicos - OPACs – On-Line Public Access Catalogs -. Es un hecho que la computación ha cambiado el trabajo de los bibliotecarios y usuarios de las bibliotecas, los catálogos son más fáciles de manejar, libres de acceso y proporcionan consultas mejoradas. Han potenciado creación de catálogos unidos por región, país o tópicos de interés para una comunidad específica y han permitido su acceso rebasando los límites físicos del campus universitario.

El primer sitio Web **[Leb99]**

de bibliotecas digitales fue Helsinki City Library de Finlandia que abrió en 1994, luego surgió *Gabriel* – acrónimo de Gateway and Bridge to Europe's National Libraries, el sitio Web para las bibliotecas nacionales representadas en la Conferencia Europea Nacional de Bibliotecarios (CENL), con la participación de 3 bibliotecas: Koninklijke Bibliotheek, British Library and Helsinki University Library, al cual otras bibliotecas se unieron y crearon en 1995 un piloto funcional basado en los servicios de cada una de ellas y sus colecciones.

En 1999 **[Sun02]** los proyectos de investigación de bibliotecas digitales se expandieron internacionalmente cuando la NSF unió su programa de investigación con actividades similares adelantadas por JISC, resultando así la iniciativa JISC-NSF International Digital Library Initiative, cuyos objetivos a 3 años eran:

- Ensamblar colecciones de información que no eran asequibles ni utilizables por las barreras técnicas, distancia, tamaño, fragmentación y otras limitaciones
- Crear nueva tecnología y entendimiento para que un conjunto de usuarios distribuidos pudieran encontrar, y explotar esta información
- Evaluar el impacto de esta nueva tecnología y sus beneficios internacionales

Desde entonces muchos otros grupos han estado involucrados en la expansión de las tecnologías y técnicas relacionada con bibliotecas digitales, incluyendo la European Union, Association for Computing Machinery (ACM), el Instituto de Ingenieros Eléctricos y Electrónicos (IEEE), la Federación Internacional de Asociaciones de Bibliotecas (International Federation of Library Associations (IFLA), La asociación americana de bibliotecas (American Library Association (ALA), la Coalition for Networked Information (CNI), y la Federación de Bibliotecas Digitales (DLF). **[Sun02]**

Hay hoy miles de bibliotecas digitales en la Web, unas con un alcance y cubrimiento de funcionalidad mayor, otras que solo publican sus catálogos y colecciones ofreciendo diferentes niveles de servicios para diferentes audiencias. Se amplía el panorama de la definición de la biblioteca digital y los avances y proyectos actuales tienen como objetivo aumentar la disponibilidad de los recursos bibliográficos entre miembros de una comunidad, facilitar la interconexión y el uso de recursos bibliográficos digitalizados.

Entre las funciones básicas que se puede distinguir en las bibliotecas digitales, y heredar de las bibliotecas tradicionales, se encuentran:

- Adquisiciones: Compra y adquisición de materiales
 - Catalogación: Creación de catálogos
 - Circulación: Entradas y Salidas del material (prestamos,)
 - Catálogo Público para acceso en línea (OPAC – Online public access catalog): Catálogo electrónico que contiene la información de los ítems de la biblioteca
 - Suscripciones: Mantenimiento y renovación de suscripciones seriales
- [Sun03]**

En nuestro medio es válido el concepto aún de iniciar el camino hacia la biblioteca digital con la implementación del servicio referencial de colecciones de información en formato digital asequibles y organizadas. Pero la tendencia es clara para evolucionar hacia la digitalización del material, la interconexión que borre fronteras y la generación de conocimiento tanto en la comunidad que usa la biblioteca como la de la biblioteca misma.

Muchas bibliotecas ya han comenzado a desarrollar sus colecciones digitales, transformado información física en formatos digitales diferentes (textos, imágenes, video, sonido), para permitir a grandes audiencias el acceso a documentos especializados, antiguos y valiosos, que previamente era imposible acceder por su interés de preservación, fragilidad, valor, horario reducido de atención, condiciones de uso y poca disponibilidad por demanda. Sin embargo los problemas técnicos relacionados con el almacenamiento, transformación de material y generación de conocimiento e información semántica, son todavía inmensos (anchos de banda, almacenamiento y direccionamiento, compresión, procesamiento del lenguaje natural y su ambigüedad).

Igualmente existen hoy bases de datos digitales de publicaciones seriales como revistas, journals que están ahora disponibles en formato digital, primero en CD-ROM y ahora a través de servicios en-línea, para ser integrados y reutilizados por las bibliotecas entre sí, creando el concepto de servicios globales, transformando las bibliotecas en recursos disponibles por demanda vía las más avanzadas tecnologías de red.

La creación y acercamiento a las diferentes formas de bibliotecas digitales ha evolucionado a medida que se encuentran soluciones a cada uno de los desafíos que implica el habilitar funcionalidades como incorporar catálogos digitales, búsqueda, despliegue y todavía se encuentran en desarrollo muchas de las técnicas para la traducción de material en formatos digitales y búsquedas que proporcionen y creen conocimiento tanto en los usuarios como en la biblioteca misma. Es así como el concepto de bibliotecas digitales va desde la creación de un catálogo en línea, como la dotación y mantenimiento de material digital, y la sistematización de los procesos operativos y de administración de las bibliotecas, la disseminación de la información y la creación y conservación del conocimiento

todo en busca de objetivos para implementar servicios innovadores, para preservar el material y posibilitar su acceso y comunicación.

3. CAMPOS DE ACCIÓN RELACIONADOS CON LAS BIBLIOTECAS DIGITALES

El campo de investigación en las bibliotecas digitales debe ser visto como una variedad de dominios combinados o diferentes participantes que realmente reflejan su verdadero potencial. En el campo de las bibliotecas digitales se involucran las perspectivas de las bases de datos, la federación, hipertexto, la Web, métodos de transformación o traducción de información física a digital, nuevas formas de información (no tangible en las bibliotecas tradicionales), almacenamiento, bodegas de datos y minería, preservación de la información en el tiempo, estándares de catalogación y metadatos, búsquedas, patrones, semántica, agentes recuperadores de información y demás estándares abiertos del desarrollo de software. **[Nur95]**

Las bibliotecas hoy enfrentan un panorama de evolución frente a muchas áreas de trabajo diferentes y tienen el reto de proporcionar mayores niveles de acceso a la información y servicios mejorados, a la vez que deben asimilar y avanzar con el ritmo del cambio tecnológico, y el manejo del presupuesto.

Internet mismo se ha convertido en una competencia para las bibliotecas digitales, ya que si estas no logran difundir efectivamente su información y permitir interfaces Web efectivas, se sub-utilizarían los esfuerzos y sitios Web de las bibliotecas especializadas.

La gran cantidad de nuevas tecnologías y herramientas disponibles hoy, la demanda de sistemas basados en estándares abiertos y fáciles de integrar, la enorme inversión que han realizado las librerías comerciales en Internet, para mejorar su personalización, disponibilidad del contenido, y funcionalidad, generan presión y pueden ser desalentadoras para abordar un proyecto de bibliotecas digitales, y aunque los beneficios pueden ser significativos, muchas bibliotecas no tienen la capacidad de absorber, mantener y constantemente actualizar sus diferentes ambientes en el tiempo **[Sun03]**

Digitalización y administración de la información digitalizada

Con el advenimiento del Internet, las expectativas para el acceso a la información han aumentado dramáticamente. La necesidad hoy es el acceso inmediato a todos los recursos de la información, sin importar desde donde y el momento (any time – any where). Esta tendencia ha propulsado la digitalización de material y su almacenamiento en sistemas multimedia

Las bibliotecas que quieran digitalizar sus colecciones deben entrar en la utilización de tecnología de administración de media digital

- Convertir el contenido físico a formato digital

- Extraer o crear metadata del material convertido, para facilitar su búsqueda y descubrimiento.
- Almacenar y administrar el contenido digital y metadata en un repositorio multimedia.
- Entregar contenido vía transferencia de archivos o “streaming media”

Catálogo Público de Acceso en Línea (OPAC - Online public access catalog catalogo)

Los OPAC's se han evolucionado en su accesibilidad a través de las interfaces Web, permitiendo el acceso al catálogo de la biblioteca de cualquier ubicación a través de un browser estándar. Los catálogos digitales de acceso público mas sofisticados proporcionan contenido enriquecido relacionado con el material de la biblioteca, como imágenes de las portadas, tablas de contenido, resúmenes y comentarios sobre el material (revisiones) o vínculos que generan valor agregado, relacionados con el material.

Algunos OPAC's también proporcionan acceso a colecciones de otros sitios Web y con el apoyo en los estándares para intercomunicación o generación de unión de catálogos.

Búsqueda Difundida - Broadcast Searching – (Meta search, federated search, single search)

El crecimiento en el número y volumen de recursos electrónicos, ha traído consigo un nuevo problema: ¿como logra el usuario promedio de una biblioteca identificar los mejores recursos para encontrar la información deseada, aprender a navegar y realizar sus consultas en interfases diferentes?

Es mas, hay muchas bibliotecas que han invertido grandes cantidades de dinero para lograr sus catálogos de acceso público en línea y aún así han experimentado una muy baja utilización de sus recursos. El usuario cuando se enfrenta a muchas fuentes de información e interfaces potenciales, encuentra algunas veces más fácil realizar su búsqueda a través de los motores de búsqueda mas conocidos, aún sabiendo que los resultados obtenidos pueden ser menos relevantes que si buscara en bases de datos de bibliotecas especializadas.

Las nuevas herramientas de búsqueda difundida eliminan al usuario la necesidad de realizar búsquedas repetidas en diferentes bases de datos y presentan al usuario una única interfaz para todas. Con un solo query se puede realizar la búsqueda en diferentes fuentes simultáneamente.

Estas herramientas pueden ser uni-protocolo o multi-protocolo:

Uni-protocolo: Traducen las consultas en un formato estándar como z39.50

Multi-protocolo: Traducen las consultas en el lenguaje y protocolo nativo esperado por cada fuente de información

Algunos de estos productos ofrecen características adicionales para filtrado de los resultados, ordenamiento, categorización de los resultados, almacenamiento y la reejecución de las búsquedas.

Ejemplo: Sirsi SingleSearch Tm

<http://www.sirsi.com/Sirsinews/20030124singlesearch.html>

<http://www.sirsi.com/Sirsiproducts/broadcastsearch.html>

Algunas bibliotecas han implementado sus propias búsquedas difundidas de manera que sirven como portales para todos los recursos de la biblioteca.

Enlace de Referencia (Reference Linking)

El software basado en enlaces de referencia proporciona un medio inteligente de navegar de un pedazo de información a otra información relacionada, lo cual implica ahorro de tiempo para el usuario.

El enlace de referencia es muy utilizado por ejemplo en las referencias bibliográficas, para buscar otros trabajos del mismo autor.

El estándar de OpenURL (NISO) esta evolucionando independiente de los proveedores, como un estándar neutral para enlazar recursos en línea.

Un número de compañías han anunciado su intención de ofrecer soluciones compatibles con OpenURL.

Descubrimiento de Conocimiento, Recuperación de la Información, Colaboración

Hasta ahora el descubrimiento de información se ha realizado por búsquedas directas. De manera que para obtener información de una biblioteca es necesario ir y explícitamente realizar la búsqueda. La convergencia de las telecomunicaciones, la computación la inteligencia artificial, pueden mejorar significativamente los servicios de las bibliotecas tradicionales.

Las áreas de Descubrimiento del Conocimiento en bases de datos (KDD), los patrones del comportamiento del individuo frente a la búsqueda y los resultados obtenidos, la recuperación y acceso a la información en un esquema multi-lenguaje, las iteraciones del usuario en el proceso de búsqueda interactiva, su nivel sociocultural representan un área relacionada directamente con la colaboración, interacción para la recuperación de la información y generación del conocimiento.

La posibilidad de comunicar información al usuario, sin necesidad de una acción explícita de su parte, la capacidad de proporcionar sensibilidad al contexto, las bodegas de datos y la minería de datos para garantizar la persistencia de la información y las predicciones basadas en patrones a través de modelaje, redes neuronales artificiales, modelos de markov, etc., son áreas importantes hoy en las bibliotecas digitales. **[Rob02]**

4. LAS BIBLIOTECAS DIGITALES HOY Y PROYECTOS EN CURSO

Aunque aún existe mucha diferencia en la aplicación de las diferentes definiciones y alcances de bibliotecas digitales, y muchas de ellas solo ofrecen un front end de catálogos digitales con motores de búsqueda, sin utilizar todavía en pleno las posibilidades que se abren con la capacidad de la computación y las comunicaciones, hay varios proyectos y ejemplos interesantes de bibliotecas digitales o servicios relacionados que pueden revelar las posibilidades que se abren en su evolución.

Proyecto Electronic Beowulf

Este proyecto es un ejemplo de imágenes digitalizadas del manuscrito Beowulf, un poema heroico anglo sajón muy antiguo. La biblioteca británica trata con este proyecto de mostrar un ejemplo de organización y disseminación en la Web de material antiguo y frágil como lo es el manuscrito Beowulf.

El manuscrito original tiene cientos de años, y está en una condición física extrema y frágil.

<http://www.uky.edu/~kiernan/eBeowulf/guide.htm>

JSTOR

Un proyecto cuya misión es proporcionar acceso digital a journals (scholarly journal articles) para la visualización impresión y descarga en varios formatos. En este servicio, las entidades educativas, individuos u otras entidades se subscriben a jstor, pagando una suma por el ingreso, de esta forma las bibliotecas reducen sus costos operativos porque no tienen que subscribirse almacenarlos, o digitalizarlos. Cuando los artículos son requeridos JSTOR cobra a las bibliotecas los cargos de derechos de autor, y cargos adicionales.

<http://www.jstor.org/>

OpenCourseWare (OCW)

Este proyecto, del instituto Tecnológico de Massachussets (Massachussets Institute of Technology), en el cual se compromete a poner el material libre de todos sus cursos en la Web, iniciativa que refleja la tendencia de la creación y disseminación de contenido colaborativo, y el efecto de esto en las tendencias en la educación y cursos.

<http://ocw.mit.edu/index.html>

Cornell Center of Excellence for Digital Libraries

La investigación en bibliotecas es el centro de una revolución en la creación, publicación y distribución de información procesada digitalmente. Acceso incremental, administración de miles de publicaciones electrónicas y servicios de información.

La biblioteca de la Universidad Cornell ha sido pionera en estas iniciativas por más de una década. Con la asociación con Endeavor Association Systems y su producto ENCompass, crean colecciones digitales basadas en sus propios activos y colecciones virtuales combinando materiales únicos en todo el mundo.

Endeavor hace posible a las bibliotecas digitalizar sus colecciones a través de un paquete comercial estandarizado.

<http://moa.cit.cornell.edu/>

Lockss

Proyecto LOCKSS (“Lots Of Copies Keep Stuff Safe”), entre Sun Microsystems, Universidad de Stanford, la fundación Andrew W.Mellon y la Fundación National Science Foundation, se pretende desarrollar un sistema seguro y confiable que salvaguarde y preserve el acceso a publicaciones digitales con suscripción.

Servicios Personalizados en el Museo J. Paul Getty

Sun y el museo J. Paul Getty están desarrollando una aplicación inalámbrica para que el visitante del museo le sirva como guía para navegar y como recurso para explorar las colecciones del museo, arquitectura y jardines.

Los museos han invertido tiempo y dinero para poner sus recursos en formato digital y en la Web, sin embargo este es el primer museo que permitirá un manejo personalizado de la información, antes, durante y después de la visita, para mejorar la experiencia de sus visitantes **[Sun03]**

Colaboración es otra funcionalidad que puede implementarse en las bibliotecas digitales – Un proyecto que está direccionando esto es DEBORA (Digital Access to Books Of the Renaissance). Estos investigadores utilizan métodos etnográficos para conocer como las personas interactúan con las biblioteca tradicionales y encontraron que la interacción entre de los usuarios de la biblioteca contribuyen al aprendizaje y localización del materia.

<http://www.comp.lancs.ac.uk/computing/research/cseg/projects/debora/set.html>

Proyectos compatibles OAI:

Open Digital Libraries

ODL es un framework para construir bibliotecas digitales como redes de componentes independientes, con comunicación de inter-componentes basadas en OAI-PMH para adoptar las filosofías y mejores prácticas del OAI

<http://oai.dlib.vt.edu/odl/>

Virginia Tech Electronic Thesis and Dissertation Collection

Es una de las pocas instituciones en el mundo que exigen las tesis y disertaciones en formato electrónico. Esta colección es hoy accesible a través de una interfaz que proporciona datos OAI-compatible.

<http://www.dlib.vt.edu/projects/OAI/vttd/vttd.html>

NDLTD – Networked digital library of theses an Dissertations

Este es un proyecto internacional que soporta la creación, archivo e intercambio de versiones digitales de tesis y disertaciones. Se utilizan protocolos OAI, para crear una biblioteca digital federada.

<http://www.theses.org/>

Electronic Thesis/Dissertation - OAI Union Catalog

Un servicio construido recolectando metadatos de Archivos abiertos de tesis y disertaciones electrónicas.

La tecnología utilizada esta basada en Open Archives con datos recolectados de fuentes y almacenados en un Catalogo Unido. El catalogo unido se le implementó una interfaz y motor de búsqueda, pero los datos son igualmente accesibles de otros proveedores de servicio tanto locales como remotos.

<http://oai.dlib.vt.edu/~etdunion/cgi-bin/index.pl>

Repository Explorer

Un test de conformidad técnica para archivos de datos que cumplan con la interfaz que proporciona datos OAI (OAI's data provider interfase).

<http://re.cs.uct.ac.za/>

MARC XML DTD

US-MARC fue incluido como uno de los formatos iniciales de metadatos, por su amplio uso en bibliotecas y sistemas relacionados con ellas. Como el sistema Marian (Múltiple Access Retrieval of library Information with Anotations - desarrollado en Virginia Tech). Un formato de transporte XML ha sido definido para MARC de manera que los registros MARC puedan ser intercambiados utilizando los protocolos OAI.

Un esquema XML que permita el soporte a la distribución de registros MARC, en la comunidad OA.

Web Characterization Repository

The W3C Web Characterization Repository es una base de datos en línea de recursos metadata en el campo de caracterización en Web, Esto incluye links a herramientas de publicación, y archivos de datados. El objetivo del repositorio además permite que cualquier recurso se pueda agregar al repositorio.

<http://www.w3.org/WCA/>

Computer Science Teaching Center -CSTC

The Computer Science Teaching Center es una biblioteca digital para recursos de enseñanza para educadores de la ciencia computacional. Los profesores pueden ir al CSTC y requerir sus recursos, un editor asigna revisores para el documento requerido y finalmente se decides si se publica como material oficial.

Otros profesores pueden ir al CSTC para encontrar recursos que puedan usar en sus clases. Se puede navegar en el material oficial, por varias formas, autor, fecha, tema, etc.

<http://www.cstc.org/>

[Oai03]

Phronesis

Por otro lado se encuentra el desarrollo hecho por el ITESM Campus-Monterrey, denominado Phronesis, el cual fue diseñado para crear bibliotecas digitales distribuidas en Internet, también permite el envío, búsqueda, recuperación de documentos, así como la administración de una biblioteca digital vía Web. Es importante mencionar que pueden instalarse varios servidores en Internet, siendo cada uno autónomo.

Phronesis utiliza a Dublin Core como estándar para el formato de metadatos

Marian (Multiple Access Retrieval of library Information with Anotations)

Así mismo el Tecnológico de Virginia (Virginia Tech) desarrollo el sistema MARIAN, que puede definirse como una sistema multiusuario diseñado como parte de una infraestructura de bibliotecas digitales la cual puede soportar varias sesiones para la realización de consultas alternadas con navegación y despliegue de documentos. También ha sido utilizado como un OPAC (On-line Public Catalog).

Proporciona la búsqueda por autor, título

Búsqueda por palabras claves (tema, título y otros campos además de palabras dentro del texto)

Marian utiliza a MARC como formato de metadatos

The Simple Digital Library Interoperability Protocol

Protocolo simple para la interoperabilidad de bibliotecas digitales (Simple Digital Library Interoperability Protocol (SDLIP) protocolo para integrar fuentes de información heterogéneas

Fue desarrollado en conjunto con las universidades Stanford University, UC Berkeley, UC Santa Barbara, the San Diego Supercomputer Center, and the California Digital Library Project.

Los clientes utilizan SDLIP para solicitar búsquedas a fuentes de información. El resultado de los documentos es retornado sincrónicamente. Se pueden construir implementaciones sobre transporte HTTP o CORBA

SDLIP contiene un “wrapper”, que habla por un lado con los servicios externos a través de protocolos de mas alto nivel requeridos en cada servicio. Y en el frente,

se soporta SDLIP si la fuente de información habla SDLIP, el cliente puede interactuar directamente con la fuente.

La interfaz de búsqueda define las operaciones necesarias para iniciar las búsquedas, La interfaz de Metadata permite a los clientes explorar las capacidades de información de la fuente, Las interfaces de entrega y resultados son facilidades opcionales para acceso a información asincrónicamente.

La IETF (Internet Engineering Task Force) está definiendo un estándar para buscar repositorios de documentos en la Web. Esto es parte de la iniciativa Web-based Distributed Authoring and Versioning (WebDAV), y es llamada Autoría, Localización y Búsqueda Distribuida (Distributed Authoring, Searching and Locating -DASL). DASL es un protocolo basado en HTTP. Define como deben realizarse las solicitudes de búsqueda y como deben entregarse los resultados.

5. ESTÁNDARES EN LA INDUSTRIA DE BIBLIOTECAS

Es indispensable determinar en la actualidad el estado y avance en los estándares relacionados con el dominio del problema de bibliotecas digitales para los cuales hoy ya existen muchos de ellos, que han implicado a muchas entidades su adaptación y utilización para posibilitar el intercambio de información y el concepto de bibliotecas federadas.

A continuación realizamos una recopilación de los estándares actuales basados en información de extraída de **[Sun03]** y del los sitios específicos de cada uno de los estándares.

La estandarización es un elemento clave en las estrategias de catalogación e intercambio de información bibliográfica.

Metadatos

Si bien el concepto de metadatos (“datos sobre los datos”) es bastante más antiguo que el termino mismo, en los últimos años se ha producido una explosión en su utilización y referenciación, y por consiguiente en la generación de metadatos por la gran extensión en el uso de Internet y su consecuente proliferación de información y recursos disponibles para ser accedidos, intercambiados y usados. **[Red99]**

Esto hizo que fuera necesario encontrar mecanismos para referenciar dichos recursos de información así como su intercambio. En relación a la creación y definición estándar de metadatos se han producido un sin numero de iniciativas y proyectos, y obviamente otros tantos para proporcionar la interoperabilidad e integración de las aplicaciones o servicios que manejen dichos metadatos, lo cual ha dado lugar a diferentes movimientos de grupos y consorcios por generar esquemas estándares y una forma mas transparente para su intercambio. **[Red99]**

Definiciones de Metadata.?

- Información que comunica el significado de otra información [3]
- Datos estructurados sobre recursos, usados como soporte o referencia para la realización de operaciones
- Aquello que las herramientas, bases de datos, aplicaciones y otros servicios de información usan para definir la estructura y el significado de sus objetos, servicios y otros artefactos computacionales **[Red99]**

Se han desarrollado innumerable cantidad de “esquemas de metadata”, generales o específicos para ciertas áreas de aplicación. Todos ellos, o la mayoría, buscan convertirse en estándar de hecho y algunos de ellos han sido definido por consorcios especialmente formados para confeccionar esos estándares. De tal forma se busca mejorar la interoperabilidad y la integración de los mismos.

MARC

MAchine-Readable Cataloguing (MARC)

Es un formato estándar para la representación y comunicación de información bibliográfica en una forma legible por una maquina y para el almacenamiento e intercambio de registros bibliográficos. Diseñado por la biblioteca del Congreso (Library of Congress) a finales de 1960, gestionado por el Network Development and MARC Standards Office de dicha biblioteca. MARC permite a las bibliotecas convertir sus catálogos de tarjetas físicas en un formato digital para ser leído interpretado, usado y compartido: Define como estructurar la información bibliográfica y sus elementos de datos que son la base para la mayoría de catálogos bibliográficos.

Todos los estándares MARC conformaron la documentación ISO 2709:1996 “formato para intercambio de información”

Las ventajas de computarizar los catálogos fue adquirida rápidamente, y hoy varias versiones de MARC son utilizadas por todo tipo de bibliotecas alrededor del mundo para proporcionar acceso computarizado a sus colecciones.

A pesar de la cooperación internacional, diferentes versiones de MARC evolucionaron, incluidas UKMARC, INTERMARC, y USMARC, que divergen en las prácticas de catalogación y requerimientos.

En 1977 la Federación Internacional de Asociaciones e Institutos de Bibliotecas (Federation of Library Associations and Institutes -IFLA) publicó UNIMARC: Universal MARC format, el cual pretendía facilitar el intercambio internacional de datos entre agencias bibliográficas.

Una Segunda edición de UNIMARC sale en 1980 y en1983 se publica un libro de UNIMARC con un enfoque principalmente para la catalogación de monografías y publicaciones seriales aprovechando el progreso de la estandarización de

información bibliográfica generada por el ISBDs (International Standard Bibliographic Descriptions (ISBDs)).

La Biblioteca del Congreso y la Biblioteca Nacional de Canadá anunciaron la armonización de los formatos USMARC y CAN/MARC a inicios de 1999, bajo el nombre de MARC 21.

MARC 21 ha sido elegido como el formato estratégico para la biblioteca británica (British Library)

Ejemplo de un registro MARC21 para un libro

```
LDR      *****nam##22*****#a#4500
001      <control number>
003      <control number identifier>
005      19920331092212.7
007      ta
008      820305 s1991#### nyu#### ##### #001#0# eng##
020 ## $a0845348116 :$c$29.95 (£19.50 U.K.)
020 ## $a0845348205 (pbk.)
040 ## $a<organization code>$c<organization code>
050 14 $aPN1992.8.S4$bT47 1991
082 04 $a791.45/75/0973$219
100 1# $aTerrace, Vincent,$d1948-
245 10 $aFifty years of television :$ba guide to series and
pilots, 1937-1988 /$cVincent Terrace.
246 1# $a50 years of television
260 ## $aNew York :$bCornwall Books,$cc1991.
300 ## $a864 p. ;$c24 cm.
500 ## $aIncludes index.
650 #0 $aTelevision pilot programs$zUnited States$vCatalogs.
650 #0 $aTelevision serials$zUnited States$vCatalogs.
```

<http://lcweb.loc.gov/marc/index.html>

Dublin Core – DCMI Core

El DCMI (anacronismo de Dublin Core Metadata Initiative), y más conocido solo por Dublin Core, es una iniciativa que define un conjunto de elementos para un descubrimiento de información en forma simple. Uno de sus objetivos principales, y el que lo hace más atractivo, es el intentar posicionarse como intermediario entre los formatos específicos desarrollados por las diferentes comunidades.

Un estándar de metadata descriptivo que emerge para objetos imágenes y multimedia, diseñado con un grupo de 15 ítems de información para ser simple de entender y usar.

Dublin Core fue diseñado para proporcionar un mecanismo ampliamente aceptado que permita descubrir, y además adaptarse, adicionando mas campos de

importancia particular para una determinada comunidad buscando soportar variados propósitos, dominios o modelos de negocio.

En esta forma, el mismo estándar base puede ser usado para un amplia variedad de propósitos y modelos de negocio.

Dublin Core es un forum abierto comprometido con el desarrollo de estándares de metadata para soportar un amplio rango de propósitos y modelos de negocio.

El taller original fue realizado en Dublin, Ohio en 1995, de donde el termino "Dublin Core" surgió. Desde entonces se han llevado a cabo nueve talleres en Inglaterra, Australia, Finlandia, Alemania, Canadá, Japón y Estados Unidos.

<http://dublincore.org/>

Encoded Archival Description

El desarrollo del estándar Encoded Archival Description (EAD) comenzó en 1993 con un proyecto de la biblioteca de la Universidad de California, Berkeley. El objetivo del proyecto era investigar la posibilidad y utilidad de desarrollar un estándar de codificación para hacer accesible a los usuarios los recursos archivados en variadas instituciones como inventarios, registros, índices, y otros documentos creados por archivos, bibliotecas, museos, y repositorios de manuscritos.

El proyecto creció de la necesidad de incluir información más allá de la que proporcionan los registros MARC. La participación aumentó y un estándar del prototipo fue lanzado en 1996.

El trabajo sobre EAD continúa. Como estándar internacional potencial, EAD se mantiene en la Oficina de Network Development and MARC Standards en la biblioteca del congreso (Library of Congress) en sociedad con la Society of American Archivists.

<http://www.loc.gov/ead/>

Z39.50

Z39.50 se refiere al estándar ISO 23950 (Z.39.50) Especificación de protocolo y servicio de aplicación para la recuperación de información "Information Retrieval Application Service Definition and Protocol Specification", y al estándar ANSI/NISO Z39.50. Ambos estándares son idénticos técnicamente, excepto algunas diferencias editoriales.

Una de las primeras dificultades de poder utilizar información en redes era la necesidad de entender la organización y estructura específica de la información en cada colección. El requerimiento de un protocolo común de búsqueda que resolviera este problema, era evidente para permitir a las bibliotecas compartir los recursos en red.

El Z39.50 define una manera estándar para comunicar dos computadores con el propósito de recuperar información. El Z39.50 facilita el uso de grandes cantidades de información en bases de datos para estandarizar los procedimientos y características de la búsqueda y recuperación de información. Soporta la recuperación de información en un ambiente distribuido, cliente servidor, donde un computador opera como cliente y realiza los requerimientos de búsqueda a otro computador que actúa como servidor de información.

El software en el servidor realiza la búsqueda en una o más bases de datos, y crea un conjunto de registros que cumplen el criterio de búsqueda como resultado. El servidor devuelve los registros del conjunto de resultados al cliente.

El poder de Z39.50 es que separa la interfaz del cliente de la información de los servidores, motores de búsqueda y bases de datos. Z39.50 proporciona una visión consistente de la información desde una amplia variedad de fuentes, y ofrece a los implementadores del lado cliente, la capacidad de integrar información de un amplio rango de bases de datos y servidores. [Umd96]

El estándar Z39.50 fue desarrollado como un protocolo cliente/servidor para buscar y obtener información de cualquier base de datos remota.

Z39.50 fue propuesto originalmente en 1984. Desde entonces, 3 versiones han sido liberadas en 1988, 1992 y 1995.

Basado en MARC (el estándar para registros bibliográficos), Z39.50 estandariza las funciones básicas de búsqueda y recuperación, abriendo así posibilidades de servicios extendidos para ordenar documentos, actualizar bases de datos, almacenar búsquedas y realizar aplicaciones como por ejemplo el préstamo inter bibliotecas. Los proveedores de software para bibliotecas han podido desarrollar muchas herramientas para explotar el valor de la información almacenada utilizando este estándar.

Open Archives Initiative

La iniciativa de archivos abiertos (OAI - Open Archives Initiative) nació de la comunidad e-print, que promueve y mantiene las colecciones asequibles vía Web de documentos académicos (scholarly papers).

La meta de estas colecciones eprint es reducir el tiempo del ciclo para arbitrar y publicar trabajos del estudiante, de manera que dicho material se pueda hacer accesible libremente y rápidamente a la comunidad académica. La meta inicial del OAI era desarrollar marcos de trabajo para la interoperabilidad y acceso compartido o federado a los archivos e-print.

Los conceptos del marco de interoperabilidad de OAI, tienen usos más amplios en el acceso que se abre a una gama del contenido digital. La arquitectura de OAI define un mecanismo por demanda para "harvesting" o recolección de metadatos formateados en XML.

OAI es un camino diferente al que proporciona el Z39.50 en su metodología de búsqueda distribuida de información, y su iniciativa pretende corregir algunas dificultades encontradas en la puesta en práctica de Z39.50; por ejemplo, diversos servidores que interpretan las consultas de diversas maneras, a falta de especificidad del estándar Z39.50, restricciones de rendimiento y problemas de escalamiento.

Hoy la OAI es una organización dedicada a resolver problemas de interoperabilidad en bibliotecas digitales, definiendo protocolos simples, aplicados al intercambio de metadatos, de manera que se permita pasar metadatos desde archivos fuentes a un archivo de destino.

Como un primer paso para la interoperabilidad, se desarrolló un protocolo de recolección de metadata que permitiera el paso de metadata de un repositorio a otro, y posteriormente para proporcionar al usuario servicios de despliegue, búsqueda y anotaciones.

A través del profesor Edward Fox del Tecnológico de Virginia quien se ha involucrado en este proceso desde el comienzo y con el fin de contribuir en el desarrollo y soporte de estándares y protocolos, el compromiso inicial fue el de integrar a la OAI el proyecto de NDLTD (red de bibliotecas de Tesis y Disertaciones -Networked Digital Library of Theses and Dissertations). Posteriormente el CSTC (Computer Science Teaching Center) y el W3C Web Characterization Repository se convirtieron en las dos primeras bibliotecas digitales OAI - compatibles. Muchos otros proyectos desde entonces se han unido al OAI. **[Vir03]**

OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting proporciona un marco de trabajo de interoperabilidad independiente de las aplicaciones basadas en la recolección de metadatos.

Hay 2 tipos de participantes en un marco de trabajo OAI-PMH:

Proveedores de Datos: Sistemas que soportan OAI-PMH como medio para exponer sus metadatos.

Proveedores de Servicios: Utilizan la metadata recolectada vía OAI-PMH como la base para construir servicios de valor agregado **[Oai03]**

OpenURL

OpenURL es un tipo específico del proceso de enlace de referencia que empaqueta metadata e identificadores que describen información de objetos y apuntan a un servidor o “resolver” (resolvedor) que especifica como realizar la resolución de vínculos para una institución específica.

El “resolver” puede aceptar este paquete de datos, combinarlo con su información institucional y convertir esta información en vínculos reales.

Usando la información en el OpenURL, el “resolver” crea un menú de servicios locales que es entregado al usuario final. La sensibilidad del contexto, permite que los vínculos estén basados en lo que está realmente disponible en cada institución individual.

El OpenURL es un protocolo abierto para interoperabilidad entre un recurso de información y un componente de servicio que ofrece servicios localizados en un ambiente de vínculos abiertos.

Es en efecto un URL que transporta metadata o claves para acceder a la metadata relacionada con el objeto para el cual se proporciona el OpenUrl

El destino de un OpenURL es le ISC (componente de servicio institucional), el resto del OpenURL transporta la metadata del objeto. El formato de especificación para el OpenURL puede ser encontrado en :

http://www.exlibrisgroup.com/sfx_openurl.htm

Para que un servicio de información sea un atento OpenURL, necesita implementar una forma de conocer la diferencia entre un usuario con acceso al ISC y otro sin acceso. El mecanismo exacto no es importante, y puede ser un cookie-pusher usado por el original SFX, o información contenida en un certificado digital similar al que se propone en el proyecto prototipo DLF de certificados digitales, o un perfil de usuario almacenado en un servicio de información o algo más.

Para los usuarios reconocidos a tener acceso al ISC, los servicios de información proporcionan un OpenURL para cada objeto pasado al usuario.

OpenURL esta basado en la idea que los links deberán conducir al usuario a los recursos apropiados. El ISC describe el contexto al usuario, donde se recibe el OpenURL como entrada y toma una acción con el.

Metadata Encoding and Transmission Standard

Un estándar importante que emerge para la interoperabilidad de colecciones digitales es el Metadata Encoding and Transmission Standard (METS), que proporciona un marco de trabajo uniforme para administrar y transmitir objetos digitales.

El proyecto Making of America II (MOA2) desarrolló un formato de codificación descriptivo para metadata estructural para trabajos textuales y basados en imágenes.

Soportado por la Digital Library Federation y la Library of Congress, METS construido sobre MOA2, proporciona un formato para codificar la metadata necesaria tanto para la administración de objetos digitales dentro de un repositorio y compartir estos objetos entre repositorios

Las principales bibliotecas académicas y de investigación están citando a METS como un estándar importante para la interoperabilidad de bibliotecas digitales y congregándose detrás de él.

Encabezados TEI (TEI Header)

TEI Header (anacronismo de Text Encoding Initiative) contienen información bibliográfica describiendo un trabajo codificado de tal manera que el texto en si, su fuente, su codificación y sus revisiones estén totalmente documentados. Es también iniciativa de la Biblioteca del Congreso de USA
<http://www.wwp.brown.edu/encoding/training/teiheader/>

VRA Core

La iniciativa VRA (anacronismo de Visual Resources Association) para recursos visuales consiste en un conjunto de elementos para describir trabajos de la cultura visual así como las imágenes que documentan esos trabajos. Se presenta como un punto de partida que facilite el compartir información sobre recursos visuales y colecciones.
<http://www.vraweb.org/>

RDF

El RDF (anacronismo de Resource Description Framework), recomendación del W3C de febrero del 99, es una estructura que permite la codificación, intercambio y reuso de metadatos estructuradas. Según la definición de su sitio oficial, RDF provee un sistema ontológico liviano para el soporte de intercambio de información en la WEB.
<http://www.w3.org/RDF/>

MOF

MOF (anacronismo de Meta Object Facility), iniciativa del Object Management Group, es un ambiente de trabajo, independiente de la plataforma, que permite definir, manipular e integrar metadatos y datos. Este estándar es usado para describir metamodelos y está basado en UML.
<http://www.omg.org/technology/documents/formal/mof.htm>

XMI

El XMI (anacronismo de Xml Metadata Interchange) es una iniciativa del Object Management Group. El propósito declarado para XMI (ver el reporte OMG XMI 1.1) es permitir el fácil intercambio de metadatos entre herramientas basadas en UML y repositorios basados en MOF en ambientes heterogéneos distribuidos, a través de estándar de W3C XML.
<http://www.omg.org/technology/documents/formal/xmi.htm>

JMI

La especificación JMI (anacronismo de Java Metadata interfase) define, basado en

MOF, una infraestructura dinámica, e independiente de la plataforma, que permite la creación, almacenamiento, acceso, búsqueda e intercambio de metadatos utilizando interfaces Java

<http://java.sun.com/products/jmi/>

CWM

La iniciativa del Object Management Group CWM (anacronismo de Common Warehouse Metamodel) es una interfase estándar que puede ser usada para un fácil intercambio de metadatos, de áreas tales como los Datawarehouses y Business Intelligence, entre herramientas, plataformas y repositorios ubicados en ambientes heterogéneos distribuidos.

<http://www.omg.org/technology/documents/formal/cwm.htm>

Sistemas y herramientas para recuperación de Información

- mg - from the Managing Gigabytes book.
- The Porter Stemming Algorithm
- Snowball: a language for creating stemmers
- Lancaster stemming algorithm: Paice Husk method.
- Z39.50 free software.
- xapian An open source search engine (formally Open Muscat, Omsee and Omseek).
- ht://Dig Open source IR software.
- Swish++ Open source IR software.
- Senga Open source IR software.
- OASIS Open source IR software.
- Lucene Open source IR software.
- Alkaline Open source IR software.
- Glimpse Open source IR software.
- Isearch Open source IR software.
- Ongobongo Open source IR software.
- Perfect Open source IR software.
- ROADS Open source IR software.
- SEARCH.PHP3 Open source IR software.
- YASE Open source IR software.
- Exist Open source IR software.
- PhpDig Open source IR software.
- NISs Open source IR software.
- Anarchivist Open source IR software.
- ASPSeek Open source IR software.
- HISS Open source IR software.
- Latente Open source IR software.
- NeatSeeker Open source IR software.
- Nose Open source IR software.
- OpenFTS Open source IR software.
- Oxyus Open source IR software.
- Ransacker Open source IR software.
- siteIndexer Open source IR software.

- SPINdex Open source IR software.
- Doc Fox Open source IR software.
- MPS Information Server Open source IR software.
- freeWAIS-sf Open source IR software.
- Harvest Web Indexing Open source IR Indexer.
- Zebra Open source IR software.
- Lemur Open source IR software.
- PLS Open source IR software.
- mnoGoSearch Open source IR software.
- Search Tools Web site.
- Inquiry
- SMART
- SMART on-line tutorial
- A list of dozens of free and commercial IR systems
- Recommender Systems amongst other things.
- CFEngine: an Open Source Collaborative Filtering Recommender System.

6. CONCLUSIONES

De acuerdo al levantamiento de información realizado y al estado de los estándares y proyectos en curso, podemos concluir que las bibliotecas digitales continúan en el proceso de evolución de extender sus servicios a un esquema que permita adicionar conocimiento a los documentos en el espacio de la biblioteca digital **Fen[01]**, migrando sus colecciones físicas a colecciones digitales, y cambiando la visión de repositorios para comportarse como compuertas de información, integradas a la comunidad, una biblioteca digital abierta y sensitiva a su ambiente.

Los estándares abiertos ocupan un papel clave a medida que nos movemos a la era de los WebServices, que permitirán tener lo que queremos, cuando lo queremos, sin importar donde estamos, a través de cualquier tipo de dispositivos. TCP/IP, Java y XML, la movilidad y redes inalámbricas, la preservación y transformación del contenido digital a disposición de los usuarios, la posibilidad de su distribución, la tendencia hacia los portales personalizados para integrar todos los recursos de una biblioteca bajo una sola presencia en Internet, con posibilidades de personalización de acuerdo a las necesidades de los usuarios son todas tendencias a seguir, que generarán confianza hacia la incertidumbre de los cambios tecnológicos y de las bibliotecas mismas.

El uso de estándares para intercambio de información, recursos y servicios se ha convertido en una necesidad en el escenario globalizado de Internet, y existe un desarrollo caótico, explosivo y poco coordinado de de esquemas y ambientes de metadata e intercambio de información.

Por otra parte se nota que XML se convierte en la base para definir la deseada intercomunicación entre los ambientes o esquemas, así como aparecen algunos

ambientes que se posicionan como estándares de hecho, por su uso, en ciertas áreas o dominios. (Dublín Core, OAI, y MARC) **[Red99]**

En este sentido es necesario trabajar para extender el verdadero potencial de una biblioteca digital, permitiendo así características como:

- Un solo punto de acceso y autenticación para validar el acceso, permisos y habilitar links a los recursos disponibles.
- Personalización para el usuario, con el fin de diseñar vistas para sus fuentes de información favoritas.
- Facilidades para guardar búsquedas o resultados obtenidos y refrescarlos posteriormente.
- Implementación de alertas para el usuario sobre nuevo material de su interés basado en el conocimiento adquirido del usuario a través de su interacción con el portal.
- Servicios virtuales de referencia: servicios que faciliten la conversación y consulta de servicios de referencia con un bibliotecario especializado. (El chat, mail, en tiempo real para permitir una entrevista con un bibliotecario de referencia para solicitar y obtener respuesta inmediata.)
- Transformación y almacenamiento de la información física a nuevos formatos.
- Administración, preservación y difusión del material de la biblioteca
- Ofrecimiento de los recursos de información, de manera distribuida y en múltiples formatos
- Habilitar las posibilidades de colaboración e interacción eliminando las barreras de la heterogeneidad de la información y servicios disponibles.

7. BIBLIOGRAFÍA

[Fen01]

Feng ,Ling. . Jeusfeld ,Manfred A. Hoppenbrouwers ,Jeroen.
Towards Knowledge-Based Digital Libraries
ACM SIGMOD Record archive
Volume 30 , Issue 1 (March 2001) Pages: 41 - 46
2001
ISSN:0163-5808
(Disponible en <http://infolab.uvt.nl/pub/fengl-2001-27.pdf>)

[Fox01]

Fox, Edward A. Gupta, Rajat: Digital Libraries: Virginia Tech Courseware
Copyright 1998-2001, (Disponible en <http://ei.cs.vt.edu/%7Edlib/def.htm>)

[Ism03]

International Standard Maintenance Agent Z39.50
<http://www.loc.gov/z3950/agency/>.
<http://lcweb.loc.gov/z3950/agency/markup/markup.html>

[Leb99]

Lebert, Marie: From the print Media to de Internet. 1999 (Disponible en <http://www.etudes-francaises.net/entretiens/printeng00.htm>)

[Nur95]

Nürnberg ,Peter J. Furuta, Richard. Leggett ,John J.
Digital Libraries: Issues and Architectures

The Second Annual Conference on the Theory and Practice of Digital Libraries -June 11-13, 1995 - Austin, Texas, USA

<http://csdl.tamu.edu/DL95/papers/nuernberg/nuernberg.html>

[Oai03]

Open Archives Initiative

Sitio Web localizado en <http://www.openarchives.org>.

The Open Archives Initiative Protocol for Metadata Harvesting
Protocol Version 2.0 of 2002-06-14

Document Version 2003/02/21T00:00:00Z

<http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>

[Red99]

Reddy, Raj. Ager, Tryg.

Digital Information Organization in Japan
International Technology Research Institute
World Technology (WTEC) Division

WTEC Panel Report on
February 1999

[Rob02]

Robins, Dave. Sochats, Ken: Digital Libraries: State of the Art and Future Directions, 2002 (Information Technology Conference, Pittsburgh). (Disponible en <http://ltl13.exp.sis.pitt.edu/Website/Webresume/DigitalLibrariesPaper/DigitalLibraries.htm>)

[Stan03]

STANFORD DIGITAL LIBRARIES TECHNOLOGIES

<http://www-diglib.stanford.edu/diglib/pub/index.shtml>

[Sun02]

Pasquinelli, Art

Digital Library Technology Trends

Sun Microsystems, Inc.

Agosto 2002

[Sun03]

Pasquinelli, Art.
Information Technology Advances in Libraries
Sun Microsystems, Inc.
Enero 2003

[Umd96]

UMDL - Digital Library Project, University of Michigan
Contributors to Term & Concept section: Judy Ahronheim, Ken Alexander, Gene Alloway, Kevin Butterfield, Ed Durfee, George Furnas, and Michael Wellman.
3 June 1996. (Disponible en <http://www.si.umich.edu/UMDL/glossary.html>)

[Vir03]

Virginia Tech Digital Library Research Laboratory
Open Archives Initiative -Virginia Tech DLRL Projects
(Disponible en <http://www.dlib.vt.edu/projects/OAI/>)